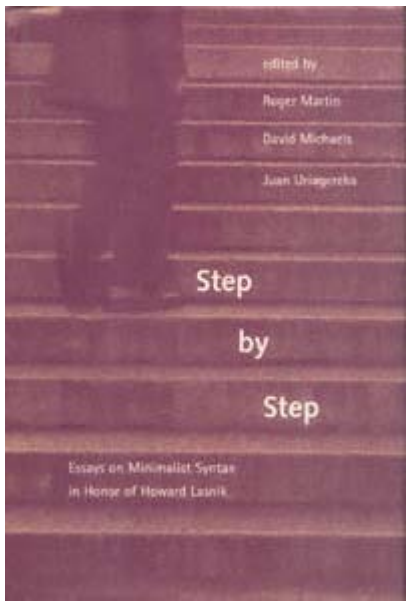


title:	Step By Step : Essays On Minimalist Syntax in Honor of Howard Lasnik
author:	Uriagereka, Juan.
publisher:	MIT Press
isbn10 | asin:	026213361X
print isbn13:	9780262133616
ebook isbn13:	9780585307428
language:	English
subject	Minimalist theory (Linguistics) , Grammar, Comparative and general--Syntax.
publication date:	2000
lcc:	P158.28.S74 2000eb
ddc:	415
subject:	Minimalist theory (Linguistics) , Grammar, Comparative and general--Syntax.

cover next page >

Step by Step

Essays on Minimalist Syntax in Honor of Howard Lasnik

Edited by
Roger Martin,
David Michaels, and
Juan Uriagereka

The MIT Press
Cambridge, Massachusetts
London, England

This book was set in Times New Roman by Asco Typesetters, Hong Kong and was printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Step by step : essays on minimalist syntax in honor of Howard Lasnik / edited
by Roger Martin, David Michaels, and Juan Uriagereka.
p. cm.
Includes bibliographical references and index.
ISBN 0-262-13361-X (hc : alk. paper)
1. Minimalist theory (Linguistics) 2. Grammar, Comparative and general
Syntax. I. Lasnik, Howard. II. Martin, Roger (Roger Andrew), 1967
III. Michaels, David, 1936 IV. Uriagereka, Juan.
P158.28.S74 2000
415dc21 99-054610

Contents

Contributors

Andrew Barss Department of Linguistics and Program in Cognitive Science, University of Arizona

Zeljko Boskovic * Department of Linguistics, University of Connecticut

Noam Chomsky Department of Linguistics and Philosophy, Massachusetts Institute of Technology

Hamida Demirdache Department of Linguistics, University of Nantes

Hiroto Hoshi Department of the Languages and Cultures of East Asia, School of Oriental and African Studies, University of London

Kyle Johnson Department of Linguistics, University of Massachusetts, Amherst

Roger Martin Department of Information and Cognitive Systems, Faculty of Education and Human Sciences, Yokohama National University

Keiko S. Murasugi Department of British and American Studies, Nanzan University

Javier Ormazabal Department of Basque Philology, University of the Basque Country, and Basque Center for Language Research

Mamoru Saito Department of Japanese, Nanzan University

Daiko Takahashi Faculty of Language and Culture, Tohoku University

Juan Uriagereka Department of Linguistics, University of Maryland

Myriam Uribe-Etxebarria Department of Basque Philology, University of the Basque Country, and Basque Center for Language Research

Ewa Willim Institute of English Philology, Jagiellonian University

Publications by
Howard Lasnik

Working Papers

(1971) On the notion "stronger derivational constraint." In *Quarterly Progress Report 100*, 178 182. Research Laboratory of Electronics, MIT, Cambridge, Mass.

(1971) A general constraint: Some evidence from negation. In *Quarterly Progress Report 101*, 215 217. Research Laboratory of Electronics, MIT, Cambridge, Mass.

(1981) On a lexical parameter in the Government-Binding Theory. In James Pustejovsky and Victoria Burke, eds., *Markedness and learnability*, 97 106. (University of Massachusetts Occasional Papers in Linguistics 6.) GLSA, University of Massachusetts, Amherst. [Reprinted, with minor revisions, in Lasnik 1990; *Essays on restrictiveness and learnability*.]

(1987) A note on indirect negative evidence. In Yasuo Ishii, Rosalind Thornton, Sung-Ho Ahn, and Eva Bar-Shalom, eds., *UConn working papers in linguistics 1*, 19 26. UConn Working Papers in Linguistics, University of Connecticut, Storrs.

(1990) Metrics and morphophonemics in early English verse. In Javier Ormazabal and Carol Tenny, eds., *UConn working papers in linguistics 3*, 29 40. UConn Working Papers in Linguistics, University of Connecticut, Storrs.

(1993) Lectures on minimalist syntax. (UConn Working Papers in Linguistics Occasional Paper 1.) UConn Working Papers in Linguistics, University of Connecticut, Storrs. [Reprinted, with minor revisions, in Lasnik 1999, *Minimalist analysis*.]

(1994) Operators and obviation. In Jun Abe, Lisa Ferro, Laurel LaPorte-Grimes, Daiko Takahashi, and Miyuki Yamashina, eds., *Papers*

in honor of David Michaels, 99 110. (UConn Working Papers in Linguistics 4.) UConn Working Papers in Linguistics, University of Connecticut, Storrs.

(1995) A note on pseudogapping. In Rob Pensalfini and Hiroyuki Ura, eds., *Papers on minimalist syntax*, 143 163. (MIT Working Papers in Linguistics 27.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass. [Reprinted, with minor revisions, in Lasnik 1999, *Minimalist analysis*.]

(1998) Some reconstruction riddles. In Alexis Dimitriadis, Hikyoung Lee, Christine Moisset, and Alexander Williams, eds., *Proceedings of the 22nd Annual Penn Linguistics Colloquium*, 83 98. (University of Pennsylvania Working Papers in Linguistics 5.1.) Penn Linguistics Club, University of Pennsylvania, Philadelphia.

Journal Articles

(1970) [with Michael Brame] A derived nominal requiring a non-sentential source. *Linguistic Inquiry* 1, 547 549.

(1971) [with Thomas Wasow] On the descriptive adequacy of interpretive theories. *Foundations of Language* 7, 429 430.

(1972) [with Robert Fiengo] On nonrecoverable deletion in syntax. *Linguistic Inquiry* 3, 528.

(1973) [with Robert Fiengo] The logical structure of reciprocal sentences in English. *Foundations of Language* 9, 447 468. [Reprinted, with minor revisions, in Lasnik 1989, *Essays on anaphora*.]

(1974) [with Robert Fiengo] Complement object deletion. *Linguistic Inquiry* 5, 535 571. [Reprinted in Minoru Yasui, ed., *Selected theses on linguistics*. Tokyo: Eichosha, 1976.] [Reprinted, with minor revisions, in Lasnik 1989, *Essays on anaphora*.]

(1976) Remarks on coreference. *Linguistic Analysis* 2, 1 22. [Reprinted, with minor revisions, in Lasnik 1989, *Essays on anaphora*.]

(1976) [with Robert Fiengo] Some issues in the theory of transformations. *Linguistic Inquiry* 7, 182 191. [Reprinted, with minor revisions, in Lasnik 1990, *Essays on restrictiveness and learnability*.]

(1977) [with Noam Chomsky] Filters and control. *Linguistic Inquiry* 8, 425 504. [Reprinted, with minor revisions, in Lasnik 1990, *Essays on restrictiveness and learnability*.]

(1977) [with Joseph Kupin] A restrictive theory of transformational grammar. *Theoretical Linguistics* 4, 173 196. [Reprinted, with minor revisions, in Lasnik 1990, *Essays on restrictiveness and learnability*.]

(1978) [with Noam Chomsky] A remark on contraction. *Linguistic Inquiry* 9, 268 274.

(1981) On two recent treatments of disjoint reference. *Journal of Linguistic Research* 1.4, 48 58. [Reprinted, with minor revisions, in Lasnik 1989, *Essays on anaphora*.]

(1981) [with Robert Freidin] Disjoint reference and *wh*-trace. *Linguistic Inquiry* 12, 39 53. [Reprinted, with minor revisions, in Lasnik 1989, *Essays on anaphora*.]

(1984) [with Mamoru Saito] On the nature of proper government. *Linguistic Inquiry* 15, 235 289. [Reprinted, with minor revisions, in Lasnik 1990, *Essays on restrictiveness and learnability*.]

(1985) Illicit NP movement: Locality conditions on chains? *Linguistic Inquiry* 16, 481 490. [Reprinted, with minor revisions, in Lasnik 1989, *Essays on anaphora*.]

(1985) [with Stephen Crain] On the acquisition of pronominal reference. *Lingua* 65, 135 154.

(1986) On accessibility. *Linguistic Inquiry* 17, 126 129.

(1986) [with Andrew Barss] A note on anaphora and double objects. *Linguistic Inquiry* 17, 347 354. [Reprinted, with minor revisions, in Lasnik 1989, *Essays on anaphora*.]

(1988) Subjects and the q-Criterion. *Natural Language & Linguistic Theory* 6, 1 17.

(1989) The nature of triggering data. *Behavioral and Brain Sciences* 12, 349 350.

(1991) Language acquisition and two types of constraints. *Behavioral and Brain Sciences* 14, 624 625.

(1991) [with Irene Heim and Robert May] Reciprocity and plurality. *Linguistic Inquiry* 22, 63 101.

(1991) [with Irene Heim and Robert May] On "Reciprocal scope." *Linguistic Inquiry* 22, 173 192.

(1991) [with Tim Stowell] Weakest crossover. *Linguistic Inquiry* 22, 687 720.

(1992) Case and expletives: Notes toward a parametric account. *Linguistic Inquiry* 23, 381 405. [Also published, in slightly different form, in Robert Freidin, ed., *Current issues in comparative grammar*, 162 189. Dordrecht: Kluwer.]

(1995) Case and expletives revisited: On Greed and other human failings. *Linguistic Inquiry* 26, 615 633. [Reprinted, with minor revisions, in Lasnik 1999, *Minimalist analysis*.]

(1997) A gap in an ellipsis paradigm: Some theoretical implications. *Linguistic Analysis* 27, 166 185.

(1999) On feature strength: Three minimalist approaches to overt movement. *Linguistic Inquiry* 30, 197 217.

(1999) [with Zeljko Boskovic *] How strict is the cycle? *Linguistic Inquiry* 30, 691 703.

(In press) [with Nicholas Sobin] The *who/whom* puzzle: On the preservation of an archaic feature. *Natural Language & Linguistic Theory*.

Articles in Books

(1975) On the semantics of negation. In Donald J. Hockney, ed., *Contemporary research in philosophical logic and linguistic semantics*, 279 311. Dordrecht: Reidel.

(1978) [with David Michaels] A reanalysis of English vowel alternations. In Wolfgang U. Dressler and Wolfgang Meid, eds., *Proceedings of the 12th International Congress of Linguists*, 808 811. Innsbrucker Beiträge für Sprachwissenschaft.

(1981) Learnability, restrictiveness, and the evaluation metric. In C.L. Baker and John McCarthy, eds., *The logical problem of language acquisition*, 1 21. Cambridge, Mass.: MIT Press. [Reprinted, with minor revisions, in Lasnik 1990, *Essays on restrictiveness and learnability*.]

(1981) Restricting the theory of transformations: A case study. In Norbert Hornstein and David Lightfoot, eds., *Explanation in linguistics*, 152 173. London: Longmans. [Reprinted, with minor revisions, in Lasnik 1990, *Essays on restrictiveness and learnability*.]

(1981) [with Robert Freidin] Core grammar, Case theory, and markedness. In Adriana Belletti, Luciana Brandi, and Luigi Rizzi, eds., *Theory of markedness in generative grammar*, 407 421. Scuola Normale Superiore,

Pisa. [Reprinted, with minor revisions, in Lasnik 1990, *Essays on restrictiveness and learnability*.]

(1988) [with Robert Fiengo, C.-T. James Huang, and Tanya Reinhart] The syntax of *wh*-in-situ. In Hagit Borer, ed., *Proceedings of the Seventh West Coast Conference on Formal Linguistics*, 81 98. Stanford, Calif.: CSLI Publications.

(1989) On certain substitutes for negative data. In Robert J. Matthews and William Demopoulos, eds., *Learnability and linguistic theory*, 89 105. Dordrecht: Reidel. [Reprinted, with minor revisions, in Lasnik 1990, *Essays on restrictiveness and learnability*.]

(1990) Syntax. In Daniel Osherson and Howard Lasnik, eds., *An invitation to cognitive science*. Vol. 1, *Language*, 5 21. Cambridge, Mass.: MIT Press.

(1991) On the necessity of binding conditions. In Robert Freidin, ed., *Principles and parameters in comparative grammar*, 7 28. Cambridge, Mass.: MIT Press. [Reprinted, with minor revisions, in Lasnik 1989, *Essays on anaphora*.]

(1991) [with Mamoru Saito] On the subject of infinitives. In Lise M. Dobrin, Lynn Nichols, and Rosa M. Rodriguez, eds., *CLS 27: Papers from the 27th Regional Meeting of the Chicago Linguistic Society*, 324 343. Chicago Linguistic Society, University of Chicago, Chicago, Ill. [Reprinted, with minor revisions, in Lasnik 1999, *Minimalist analysis*.]

(1992) Subjacency. In William Bright, ed., *International encyclopedia of linguistics*. Vol. 4, 92 93. Oxford: Oxford University Press.

(1992) Traces. In William Bright, ed., *International encyclopedia of linguistics*. Vol. 4, 170 171. Oxford: Oxford University Press.

(1992) Two notes on control and binding. In James Higginbotham, Sabine Iatridou, Utpal Lahiri, and Richard Larson, eds., *Control and grammar*. 235 251. Dordrecht: Kluwer.

(1993) [with Noam Chomsky] The theory of principles and parameters. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, eds., *Syntax: An international handbook of contemporary research*. Vol. 1, 506 569. Berlin: Walter de Gruyter. [Reprinted, with minor revisions, in Noam Chomsky, *The Minimalist Program*, 13 127. Cambridge, Mass.: MIT Press, 1995.]

(1994) Noam Chomsky on anaphora. In Carlos Otero, ed., *Noam Chomsky: Critical assessments*. Vol. 1, 574 606. London: Routledge.

(1994) Operators and obviation. In Young-Suk Kim, Byung-Choon Lee, Kyoung-Choon Lee, Hyun-Kwon Yang, and Jong-Yurl Yoon, eds., *Explorations in generative grammar: A festschrift for Dong-Whee Yang*, 294 309. Seoul: Hankuk.

(1994) Syntax. In V.S. Ramchandran, ed., *Encyclopedia of human behavior*. Vol. 4, 369 382. San Diego, Calif.: Academic Press.

(1995) The forms of sentences. In Daniel Osherson, Lila Gleitman, and Mark Liberman, eds., *An invitation to cognitive science*. Vol. 1, 283 310. 2nd ed. Cambridge, Mass.: MIT Press.

(1995) Last Resort. In Shosuke Haraguchi and Michio Funaki, eds., *Minimalism and linguistic theory*, 1 32. Tokyo: Hituzi Syobo. [Reprinted, with minor revisions, in Lasnik 1999, *Minimalist analysis*.]

(1995) Last Resort and Attract F. In Leslie Gabriele, Debra Hardison, and Robert Westmoreland, eds., *Papers from the 6th Annual Meeting of the Formal Linguistics Society of Mid-America*. Vol. 1, 62 81. Indiana University Linguistics Club, Bloomington.

(1995) Verbal morphology: *Syntactic structures* meets the Minimalist Program. In Héctor Campos and Paula Kempchinsky, eds., *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero*, 251 275. Washington, D.C.: Georgetown University Press. [Reprinted, with minor revisions, in Lasnik 1999, *Minimalist analysis*.]

(1997) Levels of representation and the elements of anaphora. In Hans Bennis, Pierre Pica, and Johan Rooryck, eds., *Atomism and binding*, 251 268. Dordrecht: Foris.

(1998) Exceptional Case marking: Perspectives old and new. In Zeljko Boskovic, * Steven Franks, and William Snyder, eds., *Formal Approaches to Slavic Linguistics: The Connecticut Meeting 1997*, 187 211. Michigan Slavic Publications, Department of Slavic Languages and Literatures, University of Michigan, Ann Arbor.

(1999) Minimalism. In Robert A. Wilson and Frank C. Keil, eds., *MIT encyclopedia of the cognitive sciences*, 548 550. Cambridge, Mass.: MIT Press.

(1999) On the locality of movement: Formalist syntax position paper. In Michael Darnell, Edith Moravcsik, Frederick Newmeyer, Michael Noonan, and Kathleen Wheatley, eds., *Functionalism and formalism in linguistics*. Vol. 1, 33 54. Amsterdam: John Benjamins.

(1999) Pseudogapping puzzles. In Shalom Lappin and Elabbas Benmamoun, eds., *Fragments: Studies in ellipsis and gapping*, 141 174. Oxford: Oxford University Press.

(1999) Chains of arguments. In Samuel David Epstein and Norbert Hornstein, eds., *Working minimalism*, 189 215. Cambridge, Mass.: MIT Press.

(In press) Derivation and representation in modern transformational syntax. In Mark Baltin and Chris Collins, eds., *Handbook of syntactic theory*. Oxford: Blackwell.

(In press) Patterns of verb raising with auxiliary 'be'. In Kimberley Broderick and Bart Hollebrandse, eds., *Proceedings of the 1995 UMass Conference on African American English*. GLSA, University of Massachusetts, Amherst.

(In press) [with Cédric Boeckx] Transformational grammar. In Robert Clark, editorial director, and Karin Kessels and Jeanette Baxter, publication eds., *Annotated bibliography of English studies*. Swets and Zeitlinger.

(In press) [with Carlos Otero] Chomsky. In John Joseph and Pierre Swiggers, eds., *Biographical dictionary of Western linguistics since 1450*. London: Routledge.

Books

(1988) [with Juan Uriagereka] *A course in GB syntax: Lectures on binding and empty categories*. Cambridge, Mass.: MIT Press. [This book has been translated into and published in Korean and Basque.]

(1989) *Essays on anaphora*. Dordrecht: Kluwer.

(1990) *Essays on restrictiveness and learnability*. Dordrecht: Kluwer.

(1992) [with Mamoru Saito] *Move* a*: Conditions on its application and output*. Cambridge, Mass.: MIT Press.

(1999) *Minimalist analysis*. Oxford: Blackwell.

(2000) [with Marcela Depiante and Arthur Stepanov] Syntactic structures *revisited: Contemporary lectures on classic transformational theory*. Cambridge, Mass.: MIT Press.

Books Edited

(1990) [with Daniel Osherson] *An invitation to cognitive science*. Vol. 1, *Language*. Cambridge, Mass.: MIT Press.

Ph.D. Dissertations
Supervised by
Howard Lasnik

Phoebe Huang (1977) WH Fronting and Related Processes

Margaret Allen (1978) Morphological Investigations

Mona Anderson (1979) Noun Phrase Structure

Craig Hoffman (1980) Phrase Structure, Subcategorization, and Transformations in the English Verb Phrase

Lori Davis (1984) Arguments and Expletives: Thematic and Non-thematic NPs

Sungshim Hong (1985) A and A ' Binding in Korean and English: Government-Binding Parameters

Samuel David Epstein (1987) Empty Categories and Their Antecedents

Elaine McNulty (1988) The Syntax of Adjunct Predicates

Juan Uriagereka (1988) On Government

Ana Varela (1988) Binding in Spanish: A Theoretical and Experimental Study

Robyne Tiedeman (1989) Government and Locality Conditions on Syntactic Relations

Aliaa Abd El-Moneim (1989) Role of INFL

Sung-Ho Ahn (1990) Korean Quantification and Universal Grammar

Rosalind Thornton (1990) Adventures in Long-Distance Moving: The Acquisition of Complex Wh-Questions

Yasuo Ishii (1991) Operators and Empty Categories in Japanese

Shu-Ying Yang (1991) Dative Alternation in Chinese and English

Eun-Ji Lee (1992) On the Extended Projection Principle

Jeong-Shik Lee (1992) Case Alternation in Korean: Case Minimality

Michiya Kawai (1992) Missing Object Constructions and Null Operator Predication

Naoko Nemoto (1993) Chains and Case Positions: A Study from Scrambling in Japanese

Jun Abe (1993) Binding Conditions and Scrambling without A/A ' Distinction

Myriam Uribe-Etxebarria (1993) On the Typology of Negative Polarity Licensing

Daiko Takahashi (1994) Minimality of Movement

Rhanghyeyun Lee (1994) Economy of Representation

Myung-Kwan Park (1994) A Morpho-Syntactic Study of Korean Verbal Inflection

Jai-Hyoung Cho (1994) Scrambling: Crossover, Reconstruction and Binding Theory

Zeljko Boskovic * (1995) Principles of Economy in Nonfinite Complementation

Keun-Won Sohn (1995) Negative Polarity Items, Scope, and Economy

Javier Ormazabal (1995) The Syntax of Complementation: On the Connection between Syntactic Structure and Selection

Roger Martin (1996) A Minimalist Theory of PRO and Control

Jeong-Seok Kim (1997) Syntactic Focus Movement and Ellipsis: A Minimalist Approach

Satoshi Oku (1998) A Theory of Selection and Reconstruction in the Minimalist Perspective

Ayumi Matsuo (1998) A Comparative Study of Tense and Ellipsis

Kazuko Yatsushiro (1999) Case Licensing and VP Structure

Sandra Stjepanovic * (1999) What do Second Position Cliticization, Scrambling and Multiple Wh-Fronting Have in Common?

Masao Ochi (1999) Constraints on Feature Checking

Preface

This book is dedicated to Howard Lasnik. It began as part of the celebration of his twenty-fifth anniversary at the University of Connecticut in 1997. For a good part of those twenty-five years Howard was the only syntactician in the Department of Linguistics at UConn, and he was serving as major advisor to as many as twenty doctoral students in any given year. His day at the office began at eight in the morning and ended at ten or eleven at night. His door was always open, and across from him at his desk there was usually a student, and another pacing the hall outside. The list of Howard's doctoral students that appears in this book shows his extraordinary productivity as a teacher and advisor. The list of his publications shows his extraordinary productivity as a scholar. Howard is probably as close as one can get to the ideal teacher/scholar.

In September 1997 a special colloquium was organized to mark Howard's twenty-five years at UConn. Howard and two of his former studentsZeljko Boskovic * and Juan Uriagerekagave talks, after which Howard was presented with a collection of essays by current and former students entitled *"Is the Logic Clear?" Papers in Honor of Howard Lasnik* (Department of Linguistics, UConn, 1997) and with the projected table of contents for this book. Both titles, "*Is the Logic Clear?*" and *Step by Step*, are intended to evoke the character of Howard's teaching and writing, but we would need many more titles and the volumes that go with them to do full justice to his work and what it has meant to his students and colleagues.

Of the contributors to this volume Boskovic,* Hoshi, Martin, Murasugi, Ormazabal, Takahashi, Uriagereka, and Uribe-Etxebarria were Howard's doctoral students at UConn. Willim also studied there, but she took her degree at her home university, the Jagiellonian University in Cracow, to which Howard traveled to serve on her dissertation examining committee.

Howard also worked regularly with graduate students at MIT, where he was an associate dissertation advisor for Barss, Demirdache, Johnson, and Saito. Barss, Johnson, and Saito have been Howard's colleagues at UConn in the past. Boskovic, * is a colleague now. Chomsky was Howard's dissertation advisor at MIT.

Of the chapters in this book, some focus on derivational issues in transformational grammar: they deal with the nature of movement or feature attraction with respect to *wh*-movement (Boskovic, Barss), A-movement (Ormazabal, Takahashi), q-role assignment (Saito and Hoshi), and the operation determining quantifier scope relations (Johnson). Others focus on more representational issues: they deal with the nature of phrase structure in tense and aspect (Demirdache and Uribe-Etxebarria), relative clauses in Japanese (Murasugi), and nominal phrases in Polish (Willim). Chomsky's chapter deals with both representational and derivational approaches. We summarize each of the chapters below.

Using some of the ideas and technology of Chomsky's *The Minimalist Program* (MIT Press, 1995), Boskovic, shows that a number of previously unexplained empirical phenomena, such as the striking differences between French and English with respect to *wh*-movement, can be predicted. Boskovic's * major claim is that lexical insertion (merger) of phonologically null elements (such as a C containing a strong [+wh] feature) after Spell-Out is possible in principle. This, combined with the assumption that Merge always expands the phrase marker (the Extension Condition), explains the well-known fact that French allows *wh*-phrases to remain in situ only in matrix questions. Differences between French and languages such as English are described in terms of a simple morphophonological parameter.

Barss also takes Chomsky 1995 as his starting point. He argues that certain interpretive and distributional facts about *wh*-questions raise empirical challenges for several of its key assumptions. Chomsky (1995) suggests that there is no covert *wh*-movement, speculating that some mechanism serves to interpret *wh*-in-situ at LF. Barss, on the other hand, argues that *wh*-phrases interpreted in situ correlate with, among other things, the lack of pair-list readings, the absence of Superiority effects, and the impossibility of inverse linking. Thus, he concludes that some *wh*-phrases (those that lack these properties) must be moving to [Spec, CP] covertly.

Ormazabal argues that minimalist assumptions regarding the locality of movement make several interesting empirical predictions. A consequence

of Chomsky's (1995) Minimal Link Condition is that at most two arguments (the external and at most one internal) of the verb can successfully raise out of VP. This predicts that in ditransitive constructions with external arguments, one of the internal arguments must remain internal to VP. Rather than complicating the theory in order to allow such movements, hence concluding that FL is less elegant in some significant sense, Ormazabal argues that this is exactly the right result, as it provides a principled syntactic explanation for a number of rather peculiar constraints on clitic clusters that were previously thought to be morphological in nature.

Takahashi investigates the nature of feature movement. Chomsky (1995) speculates that, given no other considerations, displacement operations should in general involve purely movement of formal features. That whole categories ever move overtly (generalized pied-piping) is merely a consequence of PF demands (the idea being that PF cannot interpret a category out of which feature movement has occurred). Takahashi suggests that if PF demands are the only thing that forces pied-piping, phonologically empty elements should allow pure feature movement in the overt syntax. From his analysis of control into adjunct clauses in Japanese, Takahashi concludes that such movement is in fact attested; hence, there is no need to block it arbitrarily.

Saito and Hoshi focus on the light verb construction in Japanese. One interesting property of this construction is that not all of the arguments of the q-assigning noun appear within its projection; rather, one or more appear as "complements" to the non-q-assigning light verb. Saito and Hoshi argue that the q-assigning noun incorporates into the light verb in LF, at which point it can assign its q-roles to any arguments appearing outside its original projection. They further argue that this incorporation obeys the Last Resort condition on movement. As a consequence, they assume that q-roles are formal features of the q-assigner that can enter checking relations, an assumption that has significant consequences for the theory of conceptual relations. They argue that this supports Chomsky's (1995) arguments for eliminating D-Structure and the Projection Principle.

Johnson considers the structures that yield quantifier scope interactions of various sorts. He rejects the classical QR analysis for empirical reasons. He then considers several Case-theoretic analyses of quantifier scope, concluding that they too are empirically inadequate. Rather, he argues that the movement operation responsible for yielding quantifier scope

ambiguities is of the same kind as scrambling found in Dutch and German. His analysis is very much in the minimalist spirit, in the sense that he attempts to reduce QR to an independently needed movement operation. However, the theoretical status of scrambling is less clear than that of movement for Case reasons, especially with respect to the minimalist conjecture that movement is driven solely by the need to check morphological features, such as Case and agreement.

Demirdache and Uribe-Etxebarria explore temporal relations from a new perspective, drawing on the surprising crosslinguistic manifestation of these relations in a locative guise. This connection poses two sets of important questions about the nature of functional categories. One is how and why certain substantive categories come to be "grammaticalized" into the functional system. The other is how and why cross-categorial relations, such as the one they explore between tense and aspect, emerge in the computational system. The chapter reaches into new descriptive areas in presenting a theory of temporal relations that draws parallels between the tense and aspect systems.

Murasugi investigates relative clauses in Japanese. On the basis of Kayne's theory developed in *The Antisymmetry of Syntax* (MIT Press, 1995), Murasugi argues that Japanese lacks relative clauses altogether, employing complex DPs instead. In other words, so-called relative clauses in Japanese have the structure [IP [D NP]], where no operator movement is involved. Murasugi then shows that the absence of true relative clauses in Japanese follows from a well-known constraint on antecedent-trace relations coupled with Kayne's theory of word order. She draws on a variety of evidence, such as dialectal variation and acquisition data, to make her arguments. Furthermore, her analysis, much like Willim's, touches on important questions regarding the nature of crosslinguistic variation.

Willim examines the structure of nominal phrases in Polish. Beyond its empirical contributions, much of the significance of the chapter relates to the treatment of parametric variation. Willim argues that Polish lacks the category D, thus immediately raising the serious issue of whether parametric variation is just a matter of the morphological realization of categories or whether entirely unlexicalized parameters may exist. If the latter is in fact the case, a second fundamental question is how a language without D can express (semantic) notions that presuppose D (assuming the universality of such notions). Willim offers several speculations about how this question might be addressed, including the possibility that the

mapping between the syntax and the semantics is nontrivial, or alternatively that Polish does have Ds after all, but they are inserted in LF.

Chomsky's chapter is relevant to all the others, and directly or indirectly touches on all the issues discussed in the book. The discussion develops and radicalizes many of the ideas discussed in chapter 4 of Chomsky 1995, the paper on which most of the chapters in this volume are based. For one, derivations gain more importance (much in the spirit of early work by Howard in the 1960s), in particular by way of the notion "phase." This allows Chomsky to solve some puzzles that the previous version of the Minimalist Program faced, especially with regard to the evaluation set for economy purposes. A central topic of the chapter is radical locality, which the notion "phase" contributes to. Trying to avoid computational complexity, Chomsky argues that derivational decisions are made at successive computational stages, with no look-ahead. He replaces the notion Attract with the more principled Agree and shows that this approach allows recalcitrant data to be explained and new surprising predictions to be made. Case is no longer the cause of movement, simply what activates it. And, interestingly, the model involves a single syntactic component, thus in effect eliminating LF computations proper.

We believe that all of these chapters fit very well within the scope of minimalism. Some provide direct support for minimalist ideas or analyses. Others, including Chomsky's own contribution, present internal challenges, which is both to be expected and appreciated given that the Minimalist Program is just that: a program. All of the chapters combine creativity of ideas with attention to details of linguistic analysis, which should stimulate further research on these specific areas and, more generally, the role of minimalist principles in language.

All of the chapters in this volume owe much to Howard Lasnik, though perhaps not as much as their authors themselves do. In each chapter Howard's presence can be witnessed in the form of his many specific, often brilliant ideas and suggestions. In addition to our inexpressible debt of gratitude to Howard, as editors we would like to extend our thanks to the following individuals. To Robert Freidin and Akira Watanabe, who served generously as commentators and advisors on this project, thus participating fully in this celebration. Their mark can be found in every chapter. To the editors of *"Is the Logic Clear?"*, Jeong-Seok Kim, Satoshi Oku, and Sandra Stjepanovic, * who coordinated their work with ours with considerable patience given the pressure of their earlier deadline, never losing sight of the goal we shared. To Amy Brand, who had the vision to

bend the practices of the MIT Press when it comes to festschrifts and who went along with this homage to Howard precisely because Howard himself is an exception to the usual practices: To all those who wanted to participate in this volume and we had to beg not to, or it would be immense, including Howard's many students outside of UConn and the co-authors of many of his papers. To Anne Mark, whose contribution can be found on every page of this book. And finally, to Cédric Boeckx, who compiled the index and updated references in the page proofs out of friendship and appreciation for Howard.

< previous page page_xxiv next page >

Introduction
Some Possible Foundations of the Minimalist Program

Roger Martin and Juan Uriagereka

The goal of this introduction is to clarify the conceptual bases and potential ramifications of the minimalist approach to the language faculty, at least as we understand them. In so doing, we briefly mention the chapters in this volume, which provide an excellent and up-to-date sample of research in the program, as well as a fitting tribute to Howard Lasnik, one of its main pillars and its sharpest, most positive critic. Our own intellectual debt to him is so great that not even the most elaborate sentence could seriously express it.

Two Minimalist Theses

What one might call a "weak minimalist thesis" is nothing new. The drive for simple and nonredundant theories of the world (or Occam's razor) is taken for granted in the core sciences. Even within the more specialized science of linguistics, this working methodology has brought undeniable success. From such a perspective, minimalism is just a new way to refer to what many people have been doing for a long time: seeking the best way to theorize about a particular domain of inquiry. We think of this thesis as *methodological minimalism* (or, in plain English, business as usual).

We have no qualms about such methodology, surely a necessary aspect of scientific inquiry. However, there is a more subtle take on the program, a "strong minimalist thesis," which for lack of a better term we may refer to as *ontological minimalism*. In his contribution to this book, Noam Chomsky poses the question, "[H]ow well is FL [the language faculty] designed?" (p. 92). Importantly, he is not asking, "How good is our theory of FL?" but "How good is FL itself?" Since this is an admittedly peculiar question to ask, we should consider it in some detail.

Let us first continue with Chomsky's statement of the issue:

> Suppose that a super-engineer were given design specifications for language: "Here are the conditions that FL must satisfy; your task is to design a device that satisfies these conditions in some optimal manner (the solution might not be unique)." The question is, how close does language come to such optimal design? (p. 92)

We return in the next section to what Chomsky takes to be the goal and metric of this optimal design (meeting the demands of interface specifications). For now we just want to emphasize that how well this super-engineered language works, in principle, may have very little to do with how well we as scientists manage to explain the fact. Our explanation may be good, mediocre, or terrible, and it would still be true that FL works as well as it does, perhaps optimally as just suggested. Chomsky himself explicitly makes such a point: that "the question of optimal design has nothing to do with the issue of 'best theory' for FL (however intricate and 'imperfect' the design of the system)" (p. 141, n. 12).

It is another question, though, whether the strong minimalist thesis is sensible. Chomsky is rhetorically ambivalent about this. For example, he asserts that "[t]here are minimalist questions, but no minimalist answers, apart from those found in pursuing the program: perhaps that it makes no sense, or that it makes sense but is premature" (p. 92). At the same time he devotes his entire chapter to addressing the matter, and he claims that "[t]he substantive thesis . . . that language design may really be . . . approaching a 'perfect solution' to minimal design specifications . . . would be surprising, hence interesting if true" (p. 92). He adds that

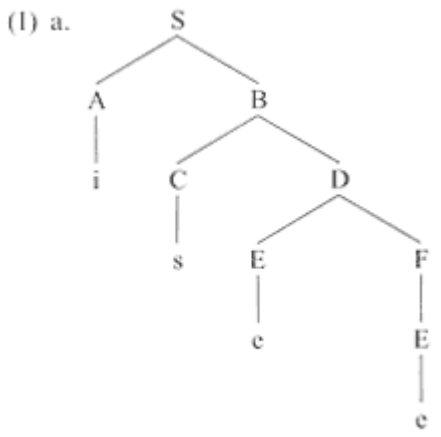
> [i]f the question is real, and subject to inquiry, then the [strong thesis] might turn out to be an even more radical break from the tradition than [the principles-and-parameters model] seemed to be. Not only does it abandon traditional conceptions of "rule of grammar" and "grammatical construction" that were carried over in some form into generative grammar, but it may also set the stage for asking novel questions that have no real counterpart in the earlier study of language. (p. 92)

Although it is in principle possible to separate the strong and the weak minimalist theses, in practice it is not so easy, and often the minimalist methodology has immediate ontological consequences of the right sort. An example that nicely illustrates this fact comes from the work of Lasnik and Kupin (1977).

Traditionally, a phrase marker (PM) is taken to be a set of strings arranged in familiar ways, whereby the crucial *is a* relation (which allows

us to distinguish chunks of the terminal string as particular phrases) can be adequately defined. For instance, consider the object informally represented in (1a) along with its more precise mathematical representation in (1b).

(1) a.



b. PM = {S, AB, ACD, ACEF, ACeF, ACeE, ACEE, ACEe,
ACee, AsD, AsEF, AsEE, AsEe, AseF, AseE, Asee, iB, iCD,
iCEF, iCEE, iCEe, iCeF, iCeE, iCee, isD, isEF, isEE, isEe, Asee,
iseF, iseE, isee}

An object like (1) is generated by applying the rewrite rules in (2a) in any order, one at a timefor instance, as in (2b).

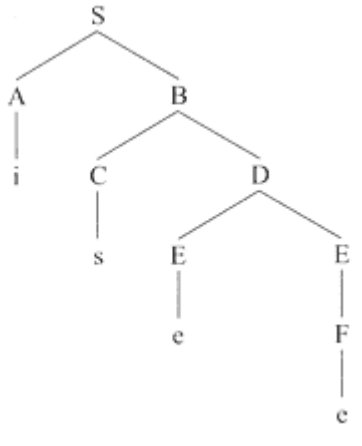
(2) a. Starting axiom: S ® A B
b. Rewrite rules: [1] B ® C D
[2] D ® E F
[3] F ® E
[4] A ® i
[5] C ® s
[6] E ® e

b. S
AB (by starting axiom)
iB (by [4])
iCD (by [1])
isD (by [5])
isEF (by [2])
iseF (by [6])
iseE (by [3])
isee (by [6])

The phrase marker in (1b) is nothing but a collection of all possible ways in which the grammar in (2a) allows us to equivalently generate *isee*, including (2b) but others as well; that is, it is a set of strings for the equivalent derivations of a given terminal string like *i^s^e^e*.

Lasnik and Kupin noted both a limitation and an inelegance stemming from Chomsky's (1955) formulation of phrase markers. To see the limitation, compare (1a) with (3a), and (1b) with (3b).

(3) a.



b. PM = {S, AB, ACD, ACEF, ACeF, ACeE, ACEE, ACEe,
ACee, AsD, AsEF, AsEE, AsEe, AseF, AseE, Asee, iB, iCD,
iCEF, iCEE, iCEe, iCeF, iCeE, iCee, isD, isEF, isEE, isEe, Asee,
iseF, iseE, isee}

The object in (1a) is clearly different from the one in (3a). In the latter F immediately dominates *e*, whereas in the former this is not the case; that is, different rules are involved. However, the phrase markers in (1b) and (3b) that describe these two different graphs are identical; put differently, there is no way of encoding the visual distinction witnessed in the graphs merely in terms of the set-theoretic object.

The inelegance in the phrase marker relates to how the *is a* relations are determined. We want to say, for instance, that the sequence *ee* in our phrase marker is a D, or that *see* is a B. Consider which particular strings in the set-theoretic object are actually used to determine this. For instance, one way we can know that *ee* is a D is by comparing the terminal string (4a) with the string (4b).

(4) a. isee
b. isD

As long as both (4a) and (4b) are members of the phrase marker, we know that this combination of strings is possible only if D stands in a domination relation with regard to *ee*. However, we could have reached the same conclusion by comparing the strings in (5), which are also members of the phrase marker.

(5) a. ACee
b. ACD

But surely we do not need *both* (4) and (5) in order to reach this conclusion. In fact, it turns out that we only need to consider the terminal string and those strings that contain at most one nonterminal element to completely determine the *is a* relationsnamely, we need the strings underlined in (6).

(6) PM = {S, AB, ACD, ACEF, ACeF, ACeE, ACEE, ACEe, ACee,
AsD, AsEF, AsEE, AsEe, AseF, AseE, Asee, iB, iCD, iCEF, iCEE.
iCEe, iCeF, iCeE, iCee, isD, isEF, isEE, isEe, Asee, iseF, iseE, isee}

Lasnik and Kupin (1977) provided a simple way to explicitly define such an object, which they called a *reduced phrase marker*. The reduced phrase marker (RPM) for the case in question is given in (7).

(7) RPM = {S, iB, iCee, isD, isEe, Asee, iseF, iseE, isee}

(7) is, needless to say, a much simpler object than the nonreduced version in (6). But (7) is also a more accurate representation of phrases in natural language. Recall the limitation that Lasnik and Kupin found. Although (7) cannot capture the difference between (1a) and (3a), there are ways to modify things so that the relevant relation is encoded. For example, we could complicate the definition of phrase marker in such a way that the fact that in one of the graphs F is rewritten as E, whereas in the other graph E is rewritten as F, is represented. But Lasnik and Kupin did just the opposite. Rather than complicating the phrase marker, they reduced it to the bare minimum in (7). In so doing, they were predicting that the kinds of representations witnessed in (1a) and (3a) are impossible linguistic representations. In other words, they correlated their purely methodological move with an empirical result, which they showed to be true.

Lasnik and Kupin's theory had another important consequence. Whereas Chomsky's (1955) phrase markers were directly generated in terms of rewrite rules, reduced phrase markers were not. Assuming reduced phrase markers are actual linguistic representations (i.e., mental representations

of some sort), this immediately entails that rewrite rules are orthogonal to the human linguistic makeup. A theory without rewrite rules turns out to be not only much simpler, but also more restrictive in terms of the logical problem of language acquisitionan argument that was emphatically made by Stowell (1981).

In sum, this is an instance where Occam's razor leads, first of all, to a more adequate model of human linguistic competence. More importantly, the ensuing modelone that at the same time contains simple objects and is consistent with a principles-and-parameters approach to grammarappears to be, in nontrivial ways, more elegantly designed than its theoretical predecessor.

But now consider a different example, where the results of methodological minimalism are not (at least obviously) ontologically minimalist as well. In the 1980s many linguists were concerned with the distribution of traces. In a nutshell, a trace was taken to require an antecedent very near by. Another way to say this is that if a barrier of some sort separates the trace from the antecedent, an ungrammatical result ensues. This is illustrated by the contrast in (8).

(8) a. [Whoi did [you think [[*t*i left]]]]?
b. *[Whoi did [you think [that [*t*i left]]]]?

BARRIER

Interestingly, adjunct traces were shown to differ from argument traces in this respect.

(9) [Whyi do [you think [that [Mary left John *t*i]]]]?

(9) should be compared with the ungrammatical (8b); descriptively, the deletable complementizer *that* disallows the argument-antecedent association in (8b) but not the adjunct-antecedent association in (9). In turn, compare the examples in (10), supposing the existence of an intermediate trace of the *wh*-phrase in the lowest C.

(10) a. [Who$_i$ do [you wonder [whether [to admit [t_i [t_i loves Mary]]]]]]?

BARRIER

b. *[Why$_i$ do [you wonder [whether [to admit [t_i [John loves Mary t_i]]]]]]?

BARRIER

Whereas the argument trace in (10a) only needs to relate appropriately to the intermediate trace (thus does not care about any subsequent barriers), the adjunct trace in (10b) seems to be sensitive to a subsequent barrier. These differences between argument traces and adjunct traces are summed up in (11).

(11) a. Only argument traces are sensitive to deletable complementizers.
b. An adjunct trace requires a completely barrier-free path between itself and its antecedent.

Lasnik and Saito (1984) came up with an ingenious way of accounting for all these facts, which we present here, somewhat anachronistically, in more or less the terms used by Epstein (1991).

(12) a. A trace is marked +g if bound by an antecedent across no barriers.
b. Once a trace is marked +g, it cannot be marked g.
c. Once a trace is indexed, it must be marked +g or g.
d. A g trace can be deleted if no information is lost.
e. A g trace left by the end of the derivation is ruled out.
f. An argument index must exist in D-Structure.
g. Free indexation takes place at LF.

None of these conditions are unreasonable. (12a c) are just an explicit statement of the fact that traces must locally relate to an antecedent. (12d) turns out to be crucial for the analysis, and is normal under the standard assumption that the grammar keeps only recoverable information. (12e) simply states that if a trace has no local antecedent, the resulting representation is ruled out. And (12f) and (12g) amount to saying that D-Structure only cares about arguments, not adjuncts. As a result, when an argument moves, it can leave behind a coindexed trace, whereas when an adjunct moves overtly, it crucially need not leave a coindexed trace.

Let us now reconsider (8), repeated as (13), assuming the machinery in (12).

(13) a. [Whoi did [you think [*t*i [*t*i left]]]]?
+g +g
b. *[Whoi did [you think [that [*t*i left]]]]?
BARRIER

In (13a) the initial trace is locally bound by the intermediate trace, which is locally bound by the *wh*-phrase. As a result, all traces are +g. In con-

trast, presence of the complementizer in (13b) blocks the possibility of local binding of the initial trace, which is thus marked g. Furthermore, we cannot delete this trace to salvage the derivation since that would result in the loss of argument information.

Consider next (11a), repeated as (14).

(14) [Whoi do [you wonder [whether [to admit [*t*i [*t*i loves Mary]]]]]]?

BARRIER

Whereas the lower trace is marked +g, the upper one is not, since its would-be antecedent is not local enough. Yet the sentence is good because the intermediate trace can be deleted without any loss of information.

Let's now look at (9), repeated as (15).

(15) [Why do [you think [that [Mary left John]]]]?

Assuming adjuncts are not (or at least need not be) indexed at D-Structure, they do not leave a coindexed trace when they undergo overt movement. This amounts to saying they do not need to leave a trace at all, or at least the trace need not be g-marked. Of course, (15) can be a question not only about your reasons for thinking something but also about Mary's reasons for leaving John. However, if there are effectively no adjunct traces at S-Structure, the question immediately arises how *why*, which appears overtly in the higher clause, can modify the lower clause.

Lasnik and Saito proposed that we should free ourselves from preconceptions concerning movement, such as its proceeding upward or its involving command. After all, independent principles, such as those encoded by the assumptions in (12), ensure that the *result* of movement will be of a certain sort. Why, on top of that, should we complicate the theory in such a way that directionality and other structural properties are stated as part of the movement rule itself? Suppose we adopt the simplest theory, which is the one with fewest assumptions, and let movement proceed anywhere, anytime, in any way. This allows for the derivation in (16), among others.

(16) a. [Why do [you think [that [Mary left John]]]]?
b. Covertly move *why* down without leaving a trace:
[do [you think [that [Mary left John why]]]]
c. Freely assign indices in the LF component:
[do [you think [that [Mary left John whyi]]]]

d. Covertly delete *that*:
[do [you think [*e* [Mary left John whyi]]]]
e. Covertly move *why* up, leaving coindexed traces:
[whyi do [you think [*t*i [Mary left John *t*i]]]]
f. Assign +g or -g:
[whyi do [you think [*t*i [Mary left John *t*i]]]]
+g +g

The operations that are of most interest take place when the derivation is on its way from S-Structure to LF. (16b) and (16e) in effect recreate the original movement except that now movement can leave an indexed trace of *why* in the embedded clause (presumably necessary to get the embedded reading) after free indexing in (16c). Also, this time we can delete the complementizer in the covert component, as in (16d). With the complementizer out of the way, assignment of +g to the indexed traces of *why* proceeds as in (16f).

Finally, we return to (11b), repeated as (17).

(17) a. [Why do [you wonder [whether [to admit [[John loves Mary]]]]]]?
b. Covertly move *why* down without leaving a trace:
[do [you wonder [whether [to admit [[John loves Mary why]]]]]]
c. Freely assign indices in the LF component:
[do [you wonder [whether [to admit [[John loves Mary whyi]]]]]]
d. Covertly move *why* up, leaving coindexed traces:
[whyi do [you wonder [whether [to admit [*t*i [John loves Mary *t*i]]]]]]
e. Assign +g or g:
[whyi do [you wonder [whether [to admit [*t*i [John loves Mary *t*i]]]]]]
BARRIER

In this case we cannot get rid of the complementizer *whether* since that would be a nonrecoverable deletion (intuitively, *whether* carries meaning). Thus, the covert movement in (17d) results in leaving an indexed trace that is too far away from its antecedent (i.e., is marked g as in (17e)). Why can we not just delete the offending g trace in (17e), exactly as we did in (14)? The difference is that the intermediate trace in (14) was created in the overt component, whereas the intermediate trace in (17e) was left in the covert component. We cannot, as it were, come back and delete the trace in (17e) once the outer cycle of the covert component has been reached, as it has in (17d).

Lasnik and Saito's theory allowed not only completely unconstrained rule applications, but in fact unconstrained rule types as wellin other words, not "Move a" but, as they called it, "Affect a," which aside from movement included deletion, insertion, or whatever was necessary to accord with the principles. Is that a simpler, more elegant theory? On methodological grounds it certainly is. The alternative theory, in addition to some version of (12), has an axiom demanding certain defining properties of transformations, such as their being limited to movement, involving upward paths, and so forth. Call those defining properties D. Lasnik and Saito's theory did not have to assume D, hence had one fewer axiom or set of axioms; that, on anybody's count, is a simpler theory. Of course, the resulting system was less constrained in that it allowed more possible derivations, including the kinds of derivational gymnastics we have just outlined. But the point of all of that was empirical as well as conceptual: those derivations were needed to account for the facts.

The two classic analyses described above have fared rather differently. Whereas Lasnik and Kupin's theory of phrase markers was assumed essentially without discussion (and is the basis, for instance, for the system in Chomsky 1981), Lasnik and Saito's theory of movement generated a variety of reactions. For one, Chomsky has consistently rejected a grammar that allows back-and-forth movements. The methodological cost of this is clear: Chomsky's theory needs a new set of axioms, usually referred to as "last resort," so that movements are always motivated according to some, as yet unclear, representational demands. The obvious question is, why go in this theoretically more costly direction?

One might think that Chomsky's alternative is actually empirically more adequate than Lasnik and Saito's, which could justify a move away from sheer theoretical elegance. But so far Chomsky's conjectures with regard to an overall theory of movement have in fact been considerably less adequate, at least on standard descriptive grounds; for example, there is no longer any systematic account of the adjunct/argument asymmetries that Lasnik and Saito's theory explained.

One could also try to argue that Chomsky's theory is in some sense more natural than Lasnik and Saito's, but that too is a difficult argument to make. We may have metrics of elegance (few, but some: symbol or axiom count, symmetry, harmony among subcomponents, and perhaps others), but just what are our metrics of naturalness? Or what exactly was unnatural about Lasnik and Saito's model? That there was no rationale

for movement? That there were no constraints on possible transformations? That the model relied on covert operations? None of those questions have obvious answers.

We believe the reason Chomsky went in a direction not supported by methodological elegance is that he relies on the strong minimalist thesis:

> The [Minimalist] [P]rogram presupposes the common goal of all inquiry into languageto discover the right theoryand asks further *why language is that way*. More narrowly, it seeks to discover to what extent minimal conditions of adequacy suffice to determine the nature of the right theory. (p. 92, our emphasis)

In other words, what the Minimalist Program is concerned with is not so much how few axioms the theory has, but whether the theory is consistent with a language design that satisfies certain minimal conditionsin the case just discussed, for instance, a motivation for movement. When the best theory happens not to match the best design, so much the worse for the best theory.

Although subservient to the strong thesis, the weak minimalist thesis is itself uncontroversial, and furthermore of much methodological value. As Chomsky puts it:

> It brings to light what might be fundamental problems, where empirical evidence and minimalist expectations conflict . . . . [W]e can ask whether [certain standard solutions to descriptive problems] are of roughly the order of complexity of the original problem . . . [so] whether they are genuine solutions . . . . The Minimalist Program helps to focus attention on such issues, and perhaps to address them by showing that elimination of descriptive technology yields empirical results that are as good as, possibly even better than, before. (p. 93)

But beyond that, the toughest question is what the strong thesis really implies, and why we should take it seriously. Chomsky is extremely cautious about the matter:

> Questions of this kind are not often studied and might not be appropriate at the current level of understanding, which is, after all, still quite thin in a young and rapidly changing approach to the study of a central component of the human brain, perhaps the most complex object in the world, and not well understood beyond its most elementary properties. (p. 93)

At the same time Chomsky's work in the last decade has forcefully embraced the underlying idea, and many theorists have followed suit. In the next section we discuss some specifics of the strong minimalist thesis, together with what we think may support it, thus motivating it beyond the evident fact that it would be interesting if it holds true.

A Mind Plan

There is a significant passage in Chomsky's contribution to this book, of the sort which, as far as we are aware, is unique in his writings to date. For good reasons (e.g., "[l]ittle is known about evolution of higher mental faculties" (p. 141, n. 15)), he calls it an "evolutionary fable."

> Imagine some primate with the human mental architecture and sensorimotor apparatus in place, but no language organ. It has our modes of perceptual organization, our propositional attitudes . . . but no way to express its thoughts by means of linguistic expressions, so that they remain largely inaccessible to it, and to others. Suppose some event reorganizes the brain in such a way as, in effect, to insert FL. To be usable, the new organ has to meet certain "legibility conditions." Other systems of the mind/brain have to be able to access expressions generated by states of FL . . ., to "read" them and use them as "instructions" for thought and action. We can try to formulate clearlyand if possible answerthe question of how good a solution FL is to the legibility conditions, and these alone. (p. 94)

Compare Chomsky's fable with a very different, more familiar one found in Pinker 1994:

> Selection could have ratcheted up language abilities by favoring the speakers in each generation that the hearers could best decode, and the hearers who could best decode the speakers . . . . Grammars of intermediate complexity . . . could have symbols with a narrower range, rules that are less reliably applied, modules with fewer rules . . . (pp. 365 366)

Whereas Chomsky's tale describes a sudden event, Pinker's is gradualistic. The latter view fits very well with the standard neo-Darwinian evolutionary line: gradual changes in genotypes, whose corresponding phenotype survives the changing environment (or does it better than competing ones), result in a leaner and meaner individual that outsmarts, outreproduces, and ultimately outlives some duller version of itself. The problem is that language does not seem to have appeared gradually at all. If we are to judge by what looks like the abrupt and combined emergence of ritual art, specialized tools for dozens of diversified and extremely specific tasks, and a complex social life, including the organized hunting of major game, then it appears that language suddenly *was*for those activities are unthinkable without the whole of language.

Similarly, Pinker's view is constructive whereas Chomsky's is modular. It is one thing to say that phonetics/phonology, syntax, and semantics evolved hand in hand, in terms of communicative or expressive success. But to say that (essentially) phonetics/phonology and semantics were

entirely in place, and then syntax suddenly connected them, is a very different story. Once again the facts seem to support the modular view. It is well known, for instance, that the anatomical structure needed for speech was in place well before (one hundred thousand years prior to) the emergence of FL as we presently understand itan event that very likely occurred some forty thousand years ago (Holden 1998). Similarly, if we are to judge from the abilities involved in tasks such as making tools for making other tools, fire control, and care for the sick and elderly, it is not unreasonable to claim that our prelinguistic predecessors had many, if not all, of our fundamental perceptual and even propositional modes. The admittedly partial picture that the fossil record shows is one of dramatic transformations, not gradual adaptations.

More seriously, the logic behind Pinker's line of reasoning falls apart as soon as one considers the sorts of linguistic properties that Chomsky is alluding to. The driving force behind the adaptive approach to the evolution of linguistic forms is their effect on language use; some genotypic change results in a phenotypic improvement, in the sense that the streamlined linguistic form is associated with some environmental advantage. But let's consider this matter more earnestly, not just with pretheoretical notions such as "symbols with a narrower range" or "rules that are less reliably applied." (Here, and in what follows, we are essentially adapting ideas from Lasnik's (1999) critique of functionalism.)

Take for example the notion of feature attraction that Takahashi discusses in this volume, and its limitations in terms of the already mentioned notion of movement as a last resort or in terms of locality conditions of various sorts. Obviously, as a result of those conditions the speaker's capacity to communicate a thought is reduced, as any of the stars in the examples show. What communicative advantage could there be in *not* being able to express the propositions and attitudes that, in the absence of those narrow conditions, the system would readily allow? Or take the parameters discussed by Boskovic * in his chapter, as a result of which languages differ massively in the surface shapes that questions can take. What is the advantage of open dimensions of this sort, whose direct consequence is systems that are mutually incomprehensible?

It's easy to give a vague justification for general properties of symbols or all-purpose rules, particularly because no other known biological system exists to serve as a point of comparisonhow would evolution have proceeded if we articulated through our noses, or if logic resided in a stomachal brain? The real concern is accounting for the fundamental

properties of FL, and in this sense there does not seem to be any obvious correlation between linguistic form (knowledge of language) and the uses to which it is eventually put. Surely there is some such correlation, or we would not be using language right now. But the adaptationist view needs something stronger: not just a correlation, but a correlation whose existence justifies the endurance of a randomly evolved form, in terms of survival of the fittest.

Here, more sophisticated analyses speculate that FL exists (with all its conditions) in order to disambiguate otherwise hard-to-understand expressions (Pinker and Bloom 1990). But again it is hard to see how such a claim can stand up to closer scrutiny. Ambiguity resolution is surely not a necessary condition for the existence of a grammatical process; there are no ambiguities in a sentence like (18) (setting aside the obvious yet irrelevant fact that *it* could be referential).

(18) *John seems that it is likely *t* to leave.

(18) violates a condition on locality and is thus simply impossible. Likewise, ambiguity resolution is not a sufficient condition for the existence of a grammatical rule. For example, the sentence in (19) has two well-known readings but to the best of anybody's knowledge only one syntactic structure.

(19) Smith's murderer is insane.

If ambiguity resolution is neither necessary nor sufficient for a grammatical condition to exist, one wonders what is meant by the claim that the former implies the latter.

Perhaps more perversely, grammatical processes often entail the *emergence* of an ambiguity that did not exist before. For example:

(20) a. John loves Sue and John loves Mary too.
b. John loves Sue and Mary too.

The sentence in (20a) is completely unambiguous; yet the one in (20b), where a syntactic process of ellipsis has applied, is ambiguous (a reading where Mary loves Sue is also possible). We are not saying, incidentally, that (20b) has a single syntactic representation with two different interpretations; that may or may not be the case, depending on how ellipsis works. Our point is simply this: the mere fact that ellipsis is a possible phenomenon in language allows for something with, let us say, the phonetic representation corresponding to (20b), obviously an ambiguous rep-

resentation. If ellipsis did not exist as a syntactic process, this ambiguity wouldn't even be possible.

In sum, Chomsky's evolutionary fable is, at worst, better than its competitors and, at best, even somewhat plausible. However, it poses a serious, immediate question (which is why we're spending so much time discussing it): why should it be that FL is a *good* solution to legibility conditions? This is the crux of a conceptual justification for the Minimalist Programin particular, the strong thesis that concerns us here.

The picture that emerges is one where, in fact, FL is not a particularly good solution to standard usability conditions, or perhaps is just as good as it is, with no particular mention of optimality traits. In fact, Chomsky speculates that FL is not such a good solution to learnability conditions either, or for that matter other kinds of conditions one might care to relate to why FL evolved adaptively; indeed, he insists on the idea that legibility conditions alone are what matters for FL (p. 96).

If FL is optimal in this intended, purely structural sense, the fact is surprising and in fact follows from nothing in the neo-Darwinian thesis on evolution. Chomsky puts it this way:

> Some basic properties of language are unusual among biological systems, notably the property of *discrete infinity*, . . . [the fact] that the language faculty is *non-redundant*, in that particular phenomena are not "overdetermined" by principles of language, . . . [and] the role of "principles of *economy*." (1995, 168; emphasis added)

But the emergence of the FL "mind plan" is not the only difficulty for neo-Darwinian stories in the theory of evolution. Also to be explained are "body plan" convergences in organ structures without shared functions (e.g., Fibonacci patterns in chordate skins, mollusc shells, jellyfish tentacle arrangements, plant phyllotaxis, microtubules within cytoskeletons in eukaryotic cells), or the very concept of speciation (how to go from a mutant to a group of mutants that stay close enough to matter for reproduction) and individual (what makes a prokaryotic cell turn into a eukaryotic one, in the process subsuming nucleus, mitochondria, organelles, through symbiosis with other microorganisms, and then start aggregating to what we now see?). This is without going past surface problems, readily understood to the layperson.

A particularly interesting example of a "body plan" that does not seem to square with the standard neo-Darwinian logic of evolution (and that, surprisingly, turns out to be interesting to the linguist) is provided by West, Brown, and Enquist (1997). These authors analyze the vertebrate

cardiovascular system as a fractal-space-filling network of branching tubes, under the assumptions that the energy dissipated by this transportation system is minimized and that the size of terminal tubes (reaching subtissue levels) does not vary across species. In so doing, they are able to deduce scaling laws (among distantly related vertebrates) that have been known to exist for some time, but have not previously been accounted for.

Biological diversity (from metabolism to population dynamics) correlates with body size. Allometric scaling laws relate some biological variable to body mass M, by elevating M to an exponent b and multiplying that by a constant characteristic of a given organism. This would lead one to think that b should be a multiple of 1/3, so that the cube root of an organism's mass relates to some of its internal functions. Instead, what researchers have found is that b involves not cube roots but *quarter* roots; this is unexpected if one is dealing with standard geometric constraints on volume. For example, the embryonic growth of an organism scales as M 1/4, or the quarter root of its mass.

The details of how a fractal network succeeds in involving quarter powers as the scaling factor are complex, but the gist of what happens is this. Imagine a function that details the way holes are evenly distributed in a sponge, assuming they become larger as the sponge grows. The first, outermost layer that harbors holes relates to the entire sponge in terms of a cubic power, since we are just speaking of standard volumes so far. But now consider the holes themselves. Imagine the next layer of the sponge, also full of holes (smaller than those in the first layer), and the next layer, and so on. Call *f* the function distributing the holes in a given layer. Obviously, *f* will depend on hole size and the total volume of the layer. But to express *f* for the entire sponge we need, apart from the three dimensions we already have for volume, a fourth dimension. This is because *f* is not just a constant quantity that describes the hole distribution; instead, it involves crucial variables tied, in part, to general volume. It is this fourth dimension that carries us from cubic proportions, characteristic of volumes, to fractal proportions in terms of fourth powers.

The fractal model predicts anatomical and physiological facts with an incredible degree of accuracy, as (21) shows (where the predicted (P) and observed (O) numbers express the scaling exponent, a multiple of 1/4).

(21)	aorta radius	P = 3/8 = .375, O = .36
	circulation time	P = 1/4 = .25, O = .25
	cardiac frequency	P = 1/4 = .25, O = .25
	metabolic rate	P = 3/4 = .75, O = .75

The list is actually much longer. West, Brown, and Enquist (1997) observe that "the predicted scaling properties do not depend on most details of system design, including the exact branching pattern, provided it has a fractal structure" (p. 126). This is very important, considering that we may be dealing with species that have relatively few genes in common. In other words, the regularities and convergences are not the obvious result of standard adaptations triggered by some useful function in the phenotypes of the relevant individualsthere are no such common useful functions. There is nothing obvious that directly relates, say, the metabolic rate of an amphibian and the pregnancy time of a mammal, or the cardiac frequency of a bird and the respiratory rate of a fish (all properties of the relevant sort).

Of course, it may always be the case that some ancestor common to all the relevant species showing accurate convergences had some property P in its phenotype, which (given some friendly environment) adaptively selected for the relevant genetic change. However, it would have to be a rather ancient ancestor since, although West, Brown, and Enquist still have not predicted this in their studies, quarter-power scaling is known to extend throughout the plant and animal kingdoms, including micro-organisms. Thus, to insist on the adaptive line would seem to be mere insistence, with no evidence to back it up and no way to falsify its claims.

One might, however, reasonably ask what the alternative is. As it turns out, the logic of an alternative to neo-Darwinian adaptation exists, and was articulated classically (in modern terms) by D' Arcy Thompson in his study *On Growth and Form* (1917). Regarding this precursor of contemporary complexity studies, Goodwin (1994) says:

> [H]e single-handedly defines the problem of biological form in mathematical terms and re-establishes the organism as the dynamic vehicle of biological emergence. Once this is included in an extended view of the living process, the focus shifts from inheritance and natural selection to creative emergence as the central quality of the evolutionary process. And, since organisms are primary loci of this distinctive quality of life, they become again the fundamental units of life, as they were for Darwin. Inheritance and natural selection . . . become parts of a more comprehensive dynamical theory of life which is focussed on the dynamics of emergent processes. (p. xiii)

The "ontological bite" of West, Brown, and Enquist's study can be easily conceived in this light, and further recalls the following statement of goals of a new (or in some ways very old) biology by Kauffman (1995):

> [M]uch of the order seen in development arises almost without regard for how the networks of interacting genes are strung together. Such order is robust and emergent, a kind of collective crystallization of spontaneous structure . . . . Here is spontaneous order that selection then goes on to mould . . . . Examples that we shall explore include the origin of life as a collective emergent property of complex systems of chemicals, the development of the fertilized egg into the adult as an emergent property of complex networks of genes controlling one another's activities, and the behavior of coevolving species in ecosystems that generate small and large avalanches of extinction and speciation . . . . [T]he order that emerges depends on robust and typical properties of the systems, not on the details of structure and function. (p. 18)

Kauffman aptly sums up this view: "Under a vast range of different conditions, the order can barely help but express itself" (p. 19).

Needless to say, one can ask about West, Brown, and Enquist's model the kind of question Pinker (1994) asks of Chomsky's general contention that "some event reorganizes the brain in such a way as, in effect, to insert FL" (p. 94). Pinker writes:

> [T]he possibility that there is an undiscovered corollary of the laws of physics that causes brains of human size and shape to develop the circuitry for [FL] seems unlikely for many reasons . . . . What set of physical laws could cause a surface molecule guiding an axon . . . to cooperate with millions of other such molecules to solder together just the kinds of circuits that would compute . . . grammatical language? (p. 363)

The presuppositions underlying this question are as curious in the case of brains as they would be for fractal networks. One could deny West, Brown, and Enquist's results by posing the rhetorical question, what set of physical laws could cause a molecule guiding the relevant units in the networks to cooperate with millions of other such molecules to solder together just the kinds of circuits that would result in those particular networks? Obviously, no one knows, just as no one knew at the turn of the twentieth century what sets of physical laws could predict the property of chemical valence. What scientists did was to approach the problem from the vantage point of both chemistry and physics, eventually unifying the two perspectives by dramatically extending the more basic science to account for the results of the specialized onethrowing out the chemical results was apparently not suggested.

Pinker's other reason to doubt Chomsky's general claim about the accidental, nonadaptively justified emergence of the language faculty comes from the claim that large brains are, per se, maladaptive, hence (according to Pinker) could have emerged only as a result of some good associated

function. One may grant the premise (large brains are metabolically costly, they entail difficult births, and perhaps others) and yet not agree with the conclusion. First, the logic of evolution dictates that only if a maladaptive trait results in genes' not being transmitted would the individuals bearing it not survive. A few (or even many) women's deaths in childbirth, or the need for abundant food to keep brains functioning, do not in and of themselves entail the extinction of hominids any more than they guarantee their survival. A careful study of their environmental circumstances, competitors, predators, prey, and overall habits and habitats is needed before concluding that a maladaptive trait entails anything whatsoever in evolutionary terms.

Moreover, even if one were to come up with a reasonable scenario along those lines, just about any advantage of large brains could offset the disadvantagesit doesn't have to be language. For instance, Gould (1991) refers to circulatory benefits of large brains for slender australopithecines living in hot savannah conditions. That, or something like that, may have been the only advantage of large brains.

Finally, and most importantly, the crucial change in the human brain (from not having to having FL) took place long after it became large. The large brain developed hundreds of thousands of years before the human language faculty emerged. Indeed, Neanderthals are known to have had larger brains than ours, yet they lacked all of the crucial (serious) properties one associates with human language and became extinct soon after present-day humans spread across the planet.

The bottom line is this: nobody knows why brain size increased during the course of evolution, or what that resulted in. Fodor (1998) sums the situation up in his review of Pinker 1998:

> [W]hat matters with regard to . . . whether the mind is an adaptation is not how complex our behaviour is, but how much change you would have to make in an ape's brain to produce the cognitive structure of a human mind. And about this, exactly nothing is known. That's because nothing is known about the way the structure of our minds depends on the structure of our brains . . . . Unlike our minds, our brains are, by any gross measure, very like those of apes. So it looks as though relatively small alterations of brain structure must have produced very large behavioural discontinuities in the transition from the ancestral apes to us. If that's right, then you don't have to assume that cognitive complexity is shaped by the gradual action of Darwinian selection on pre-human behavioural phenotypes . . . . [M]ake an ape's brain just a little bigger (or denser, or more folded, or, who knows, greyer) and it's anybody's guess what happens to the creature's behavioural repertoire. (p. 12, col. 2)

Why should linguists care about any of this? What should be important to linguists is not how the physical support of FL came to be what it is, but what FL's properties are and perhaps even why they happen to be those, instead of others. In the last respect, it is very instructive, from a minimalist perspective at least, that fractal structures of the sort in West, Brown, and Enquist's (1997) study have all of the properties that Chomsky ascribes to FL and furthermore happen to be those that he insists (1995, 168) "are unusual among biological systems." The structural elegance of these fractal structures should be obvious from the sponge scenario discussed above, including their discrete infinity (the network could, in principle, "branch" forever) and their structural economy (the network represents an optimal packaging of features that depend on a central axis). As for the systems' being nonredundant, with principles not over-determining the form of the relevant structures, one could correlate the mysterious linguistic variation mentioned above with the fact that the overall structure studied by West, Brown, and Enquist is implemented differently in different creatures, depending on a variety of parameters external to the system, such as the flexibility of the tubing walls or the viscosity of the circulating fluid. In sum, Chomsky's three basic properties obtain in these systems.

We believe this is very important because, to the extent that it is true, it puts the strong minimalist thesis in a new light, without having to face any special crisis concerning the nature of the reality it studies or needing to demonstrate its "ontological clout" any more than any deep biological system has to. Chomsky (1995) observes that fundamental linguistic properties are "more like what one expects to find (for unexplained reasons) in the study of the inorganic world" (p. 168). It may turn out that the inorganic world is more central to the organic one than is customarily assumed, if those terms even make sense at this level of abstraction. Why any of this should be true, no one knows. Regardless, the key is that these properties (structural economy, plastic nonredundancy, and discrete infinity) hold of both a core biological system and the linguistic onegood news for those who believe the linguistic system is, after all, biological. The bottom line is that "mind" and "body" are ultimately be-having pretty much alike.

Ontological Minimalism

Chomsky, in class lectures at least, has often appealed to the same metaphor that Kauffman uses in the earlier quotation: crystallization. He

seems to think that grammar could have emerged in roughly the way a crystal does, only at a more complex and arcane level of reality, whatever that might be. Quite plausibly, this general idea is what prompts him to suggest the sudden brain reorganization that gave our ancestors FL, as opposed to a gradual set of changes, each motivated by some adaptive benefit. But what difference does any of this make to the working linguist?

Chomsky concerns himself with just that question when he asks, "What is 'good design'?" (p. 94). And, although he suggests we may put aside this important matter for now, he adds, "The appropriate place to seek answers is in the hard sciences, where understanding is far deeper and intuitions are far more firmly grounded." At the same time, he warns, "Caution is in order in appealing to such considerations. Given some empirically supported conclusion, it is often possible to construct plausible conceptual grounds for it, and for alternatives" (p. 94).

There is nothing much one can do about that, other than being cautious. Nonetheless, in this section we pursue more seriously the idea that a good research strategy in our early explorations of the strong minimalist thesis is indeed, following Chomsky's advice, to seek answers and to ground "economy" explanations in the hard sciences. Reasonably, if the linguistic system has the elegance one expects of a chemical reaction or a standard physical system, it should plausibly be obeying principles concerned with maximums and minimumssuch as those involved in entropic equilibrium or least action/distance, as Fukui (1996) observesand in general the sorts of checks and balances that find their way into existence through conservation laws of various sorts.

If one is generous with interpretation, conservation laws seem to hold in various places in FL. Though there are obviously no quantities to conserve in linguistics, *qualitative patterns* of certain sorts are systematically conserved. For example, lexical information does not speak to phrasal information, which does not speak to discourse information.

(22) a. Every student didn't observe the holiday.
b. every student's nonobservance of the holiday

(23) a. Every student thought he shouldn't observe the holiday.
b. Every student attended the class. He didn't observe the holiday.

(22a) has two readings, one where *not* takes scope outside *every* and one where it takes scope inside *every*. In contrast, in (22b) *non* cannot take scope outside *every*, and thus the noun phrase denotes the lack of observance of the holiday by the entire group of students. This is explained if

the negative cannot interact with *every* in (22b) although it can in (22a). Why? Because words and phrases constitute different levels of representation. Similarly, in (23a) the quantifier *every* is able to bind the variable *he*, with the result that the pronoun covaries with the variable instantiating *every student*. However, covariance is impossible in (23b). Why? Because phrases and discourses constitute different levels of representation.

It looks at if some informational qualities are best captured at word level (see Saito and Hoshi's chapter in this volume, which explores light verb formation), others at phrase level (see Demirdache and Uribe-Etxebarria's chapter, which studies temporal restrictions in terms of postulating a locative phrase to encode tense), and still others at sentence level (as in Barss's chapter, where certain long-distance relations among *wh*-phrases are studied). This recalls conservation laws demanding, say, the conservation of matter/energy, or particle spin, or the like. Such laws basically establish a cut in the known universe where certain properties hold constant. In fact, this is why they are so important: not because they refer to quantities, but because they provide as clear a domain of "naturalness" as exists in the natural sciences. Our point is that, inasmuch as domains of naturalness are known to exist within FL, we could perhaps think of them as being akin to the domains of conservation defined elsewhere in the universe, thus justifying their presence within the system.

That decision would have good consequences with regard to other tacitly assumed properties of the linguistic system. For instance, consider the very fact that derivations are not allowed to lose information or, conversely, that they are not allowed to gain information either. Syntactic derivations, as all the chapters in this volume show, are nothing but manipulations of given arrays of lexical choices, governed by specific principles that constrain relevant procedures. In older models this followed from the fact that LF was seen as being mapped from D-Structure (the Projection Principle). However, once D-Structure is dispensed with for good empirical and arguable conceptual reasons (it is not an interface level), just what does it mean to say that information is neither created nor destroyed in the derivation of LF? Created or destroyed with respect to what? One of the main reasons behind having a set of lexical tokens from which LF is mapped (i.e., a lexical array or numeration) is precisely to answer this question. The answer, however, directly encodes a domain of conservation: the tokens in the numeration, and only those, are chains at LF. This is not a condition on the mapping between levels, since the lexical array is not a level; but it is a perfectly legitimate conservation law (corre-

sponding to the word "cut" mentioned above). It is not very different from the oldest conservation law in physics, the one demanding that nothing be created or destroyed in a chemical reaction: the first law of thermodynamics.

We should note that nothing in a communication code demands that it should obey a conservation law; most codes do not. Encryption is based on the idea that a message is garbled with all sorts of junk, which in the absence of the key will lead its would-be interpreters astray. Likewise, most formal systems contemplate the possibility of vacuous statements, which are not eliminated simply because they do not harm the system, and because eliminating them would complicate the general statement of the formal language. Not so with human language, which does not seem to tolerate vacuity and seems to encode just the information that is strictly necessary to make the system work. Why? An ontological take on the Minimalist Program suggests that answering that question is akin to answering the question of why matter/energy is not destroyed, or why quantum mechanical processes conserve certain quantities. In general, no one knows why language does not tolerate vacuity and has other conservative properties, but here linguists share their uncertainty with all other scientists, who also do not know why conservation laws obtain everywhere in the universe.

Of course, the existence of a domain of information encapsulation does not mean that information cannot flow between levels; it does, as Ormazabal's chapter shows for relations between word and phrase levels through agreement processes, Johnson's does for relations between phrase and sentence levels by studying how quantifier phrases locally reach their scope-taking site within the sentence, and Barss's does for sentence-level interpretation of moved *wh*-phrases. Nonetheless, these interlevel dependencies are not as free as relations within a given level. And they never skip intermediate "cuts"; no known relation goes from the word directly to the sentence, for instance, skipping the phrase level altogether. Both the limitations on possible cross-level relations and the existence of layers of implication among the various "cuts" indirectly argue for the pre-supposed levels themselves. Many of these ideas are more or less tacitly assumed, but they come to the fore once one subscribes to the strong minimalist thesis, and we are trying to suggest they make good sense as design specifications within a natural systemor at the very least, they make as much sense as similar conditions do in other natural systems.

Seen in this light, other instances where the linguistic system exhibits structural elegance, symmetry, economy, and the like, are no more or less

troublesome than similar instances elsewhere in the inorganic world. Take, for example, locality principles of the kind explored in this volume by Takahashi, Ormazabal, Boskovic, * and Chomsky. Certainly, locality is very common in the standard physical universe. For example, a field transmits information in a wavelike fashion; it does not make sense to speak of a classical magnetic field that skips a portion, so that a magnet attracts a nail across an intervening paper clip. One might conjecture that locality in language is, mutatis mutandis, of the same sort. Chomsky posits a "probe" element that relates to a "goal" in much the same terms. If, in its search down a phrase marker, the probe hits a matching goal, the search blindly ends. If that particular "agreement" relation (between "probe" and "goal") happens to be what the system requires for convergence, everything is fine; otherwise, the derivation crashes. Crucially, derivations do not have the ability to "smartly" search for ways of bypassing the blind, fieldlike procedure just described.

Similarly, recall the last resort strategy we alluded to earlier, which is central to several chapters in the volume, including Chomsky's, where it takes the form of Suicidal Greed, a development of Lasnik's (1995a,b) Enlightened Self-Interest. Dynamic systems behave, in some respects at least, in terms very reminiscent of those implied by Suicidal Greed. It's obvious, for instance, that whirlpools in fluid currents don't arise randomly and that they emerge in regions where the fluid is unstable (e.g., after it flows past a rock). This can be connected with a third, less obvious behavior: as fluid velocity increases, these eddies become more unstable, to the point where their partial symmetry breaks into turbulence. Points of turbulence can therefore be thought of as "enslaving" the eddies, which in turn can be thought of as "enslaving" the current at large (see, e.g., Mainzer 1994). When we say transformations are greedy, what we might ultimately mean is that we expect those dynamic points in the system to occur at crucial and, as it were, unstable regionsthose where some (historically randomized) morphological checking takes place. Could it be that morphological checking is to syntactic derivations as rocks are to streams: attracting points for the formation of chains, in linguistics, or eddies, in currents? That would offer a rationale for the movement of otherwise stable bare phrases.

Chomsky relates many of these curious properties of the language faculty to one particular design specification: "Reduce computational complexity." However, are computational concerns the reason for the existence of these properties or is lack of computational complexity

merely a welcome consequence? As Chomsky notes, if it turns out that computational complexity is relevant to FL, that would be a "puzzling fact about the nature of the mind/brain" (p. 99). In other words, it is by no means obvious why computational concerns should arise in a system that characterizes the structure of a body of knowledge. The same situation arises for the sorts of biological systems discussed in the previous section, many of which (plant phyllotaxis, embryonic development, protein-folding specifications, and various others) are nowadays standardly modeled in computational terms, and where, just as in the linguistic system, computational complexity appears to be a fundamental factor in deciding on the right models.

Why computational concerns should matter to systems characterizing the general form of organisms is unclear. Scientists' only conjecture is that form, in this universe, is the result of some process that can be characterized computationally. Of course, that process is ultimately evolutionary (in the broad sense discussed in the previous section, beyond neo-Darwinian considerations), but it somehow permeates down to the developmental unfolding of an organism, and apparently also to systemic dynamics that hold of adult organisms and that make crucial use of organic systems based on the deployment of some genetic array, given by evolution, into a more elaborate structure induced by development. The case of FL here is obvious, its initial state being genetic and its adult state depending on input environmental information, but many other such systems exist (e.g., all known instances of limb growth, sex determination in many species of fish, morphological class specifications in insect drones, and innumerable core behavioral patterns in various creatures, such as those involved in playing; see Goodwin 1994).

If for whatever reason or reasons (perhaps unrelated ones ultimately rooted in the physical properties of the system in question) economy notions happen to have arisen in these systems, then they will result in computational advantages, including, in the case of the linguistic system, those that Chomsky mentions: eliminating "(a) superfluous elements in representations [and] (b) superfluous steps in derivations," "reduc[ing] 'search space' for computation: 'Shortest Movement/Attract,' [etc.]," providing "'local determinability' conditions barring 'look-ahead'" (p. 99), and surely many others.

It may seem strange to view as "computational" not just the workings of FL, but also other systems in the organic world. The issue here is whether, aside from incorporating the standard logical workings of a

computer, a computational system must in some sense be "representational." The folds of a protein or the steps leading from egg to individual are obviously not "representational" (of anything else). Yet there is a whole respected field that calls itself "computational biology." Rather than challenging this legitimate discipline, which appears to be teeming with results, it seems more appropriate to question whether the "representational" view is necessary for a computational account.

Chomsky comments on this in passing. He explicitly notes that in generative grammar "the term *representation* is a technical one, with no '*representation*' relation in the sense of representational theories of ideas, for example" (p. 140, n. 8). However, as Rey (forthcoming) observes, there is still an issue about the symbols in the system, the phrases, the chains, notions such as "command" and "agree," and so forth. Are they, in any real sense, representing some kind of "idea"? The question is not entirely meaningless, especially given Chomsky's suggestion that

> FL can be regarded as a "language organ," in the informal sense in which the visual system, the immune system, and the circulatory system are commonly described as organs of the body: not objects that can be removed leaving the rest intact, but subsystems of a more complex structure that we hope to understand by investigating parts that have distinctive characteristics, and their interactions. (p. 90)

Pretheoretically, the immune and circulatory systems are no more representational than the system responsible for protein folding; however, the visual system may be representational in some nontrivial ways (see Marr 1982). Though all of these organic systems seem to have logico-computational aspects to them, is FL like the visual system in possibly being representational, or like all others in clearly not being so? (Or does the question ultimately even make sense, at some deeper level of abstraction?)

We do not know the answer to that question (or to the related, and obviously much harder, question of what it might mean for a natural system to be representational). In our opinion, the contributors to this volume would address it differently. For the chapters by Boskovic, * Murasugi, Ormazabal, Saito and Hoshi, and Takahashi, it would not seem to matter whether symbols correspond to anything, since what these authors are interested in are the details of symbols' computational destinies: when they are allowed to be displaced, how far, in what guise, across which domains, and so forth. However, the rest of the studies appear to have

more of a stake in something that goes beyond derivational dynamics: *wh*-phrase absorption and how it affects LF structures (Barss); the structures that enter into determining quantifier scope, as well as what other structures they do or do not resemble (Johnson); a Tense element that is substantively claimed to be locative (Demirdache and Uribe-Etxebarria); and the possibility of determiner interpretations in nominals without a D category (Willim). In all of these instances, substantive interpretations for various representations seem not just appropriate but indeed necessary if the logic of the chapters is to hold. Chomsky's chapter needs both: it involves obvious computational considerations, but also substantive claims about, for instance, interpretable features.

To conclude, we want to admit that our position regarding the ultimate nature of these minimalist devices, be they derivational or representational, is purposely rather weak. Whatever their origin might be, they have nice computational consequences, in effect making FL possible. Furthermore, it should be clear that we are not concerned with answering the question of what gave rise to form in nature, or even why it can be modeled computationally. Our position is much more modest: FL is a very good, and well-understood, instance of natural form. Therefore, we feel (standard) linguistic studies should be taken quite seriously when these general questions are being evaluated within the sciences of complex systems, since they provide a unique analysis of one of the most intriguing natural systems there is. Moreover, when linguists consider the nature of their inquiry and, most importantly, how the elusive notion of "good design" should be characterized, adopting the perspective of the core sciences will not hurt.

More bluntly, the ontological claim is simply that there is no reason to isolate linguistics from everything else that is going on out there, be it in physics, chemistry, dynamics, or whatever, including yet to be explored regions of some scientific cut of reality. Importantly, we are not suggesting limiting our view to psychology or biology, or at least not those disciplines as they are customarily understood. Chomsky goes as far as suggesting that

> we might replace [proposals about psychological or neurological reality] by a substantive (but extraordinarily strong) empirical hypothesis, namely, the [strong minimalist] thesis . . . : an optimal solution to legibility conditions satisfies all other empirical tests as well. The reformulated thesis replaces the obscure notion of "linguistic evidence" by the meaningful notion: satisfaction of interface conditions. (p. 97)

Provocative as that may seem, it is entirely in line with everything suggested in this chapter. To the extent that those interface conditions are part of reality, our point is that although it may be surprising that meeting those conditions should be achieved optimally, it is just as surprising that organisms meet other structural conditions for their very existence on this planet, or perhaps in this universe. The resulting structures can be modeled as "efficiently computed" by (in terms of) a mind or (in fact, equally efficiently) a developing embryo or other systems; but in our view that is more plausibly the consequence of a very intriguing, general, somehow biologically rooted state of affairs, rather than the raison d'être for structural elegance within a given organism.

Note

We would once again like to express our sincere gratitude to Howard Lasnik, this time not so much for his monumental contributions to the field of linguistics, but on a more personal note for being a unique teacher, advisor, and friend. We also owe a debt we could never possibly repay to Noam Chomsky, who not only inspired many of the ideas contained in this introduction but also commented extensively on an earlier version. Neither of them should be blamed for our errors or misinterpretations. In addition, we would like to thank the following people, whose comments, suggestions, discussion, and in some cases disagreement have helped us sharpen our ideas and improve the exposition: Jun Abe, Cédric Boeckx, Jorge Campos, Juan Carlos Castillo, Stephen Crain, John Drury, Robert Freidin, Elena Herburger, Norbert Hornstein, Hisatsugu Kitahara, David Lightfoot, David Michaels, Jairo Nunes, Wayne O'Neil, Carlos Otero. Paul Pietroski, David Poeppel, Philip Resnik, Georges Rey, Robyne Tiedeman, and Akira Watanabe. This research has been partially funded by a grant from the National Science Foundation (SBR 9601559; Uriagereka) and by the Japan Society for the Promotion of Science (Martin).

References

Chomsky, Noam. 1955. The logical structure of linguistic theory. Ms., Harvard University and MIT, Cambridge, Mass. [Revised 1956 version published in large part in 1975 (New York: Plenum) and 1985 (Chicago: University of Chicago Press).]

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Epstein, Samuel David. 1991. *Traces and their antecedents*. Oxford: Oxford University Press.

Fodor, Jerry. 1998. The trouble with psychological Darwinism. *London Review of Books*, January 22, 11 13.

Fukui, Naoki. 1996. On the nature of economy in language. *Cognitive Studies* 3, 51 71.

Goodwin, Brian. 1994. *How the leopard changed its spots*. London: Weidenfeld & Nicolson.

Gould, Stephen Jay. 1991. Exaptation: A crucial tool for evolutionary psychology. *Journal of Social Issues* 47, 43 65.

Holden, Constance. 1998. No last word on language origins. *Science* 282, 1455 1458.

Kauffman, Stuart. 1995. *At home in the universe: The search for laws of self-organization and complexity*. New York: Oxford University Press.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Lasnik, Howard. 1995a. Case and expletives revisited: On Greed and other human failings. *Linguistic Inquiry* 26, 615 633.

Lasnik, Howard. 1995b. Last Resort. In Shosuke Haraguchi and Michio Funaki, eds., *Minimalism and linguistic theory*, 1 32. Tokyo: Hituzi Syobo.

Lasnik, Howard. 1999. On the locality of movement: Formalist syntax position paper. In Michael Darnell, Edith Moravcsik, Frederick Newmeyer, Michael Noonan, and Kathleen Wheatley, eds., *Functionalism and formalism in linguistics*, vol. 1, 33 54. Amsterdam: John Benjamins.

Lasnik, Howard, and Joseph J. Kupin. 1977. A restrictive theory of transformational grammar. *Theoretical Linguistics* 4, 173 196.

Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15, 235 289.

Mainzer, Klaus. 1994. *Thinking in complexity*. Berlin: Springer-Verlag.

Marr, David. 1982. *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman.

Pinker, Steven. 1994. *The language instinct*. New York: Morrow.

Pinker, Steven. 1998. *How the mind works*. New York: W.W. Norton.

Pinker, Steven, and Paul Bloom. 1990. Natural language and natural selection. *Brain and Behavioral Studies* 13, 707 784.

Rey, Georges. Forthcoming. Chomsky, intentionality, and the CRTT. In Louise Antony and Norbert Hornstein, eds., *Chomsky and his critics*. Oxford: Blackwell.

Schoemaker, Paul. 1991. The quest for optimality: A positive heuristic of science? *Brain and Behavioral Studies* 14, 205 245.

Stowell, Tim. 1981. Origins of phrase structure. Doctoral dissertation, MIT, Cambridge, Mass.

Thompson, D'Arcy. 1917. *On growth and form*. Cambridge: Cambridge University Press.

West, Geoffrey, James Brown, and Brian Enquist. 1997. A general model for the origin of allometric scaling laws in biology. *Science* 280, 122 125.

Chapter 1
Minimalism and Asymmetric *Wh*-Interpretation

Andrew Barss

On the classical Government-Binding (GB) model of *wh*-movement (initiated by the foundational proposals made in Chomsky 1976, 1977, 1981, Higginbotham and May 1981, Huang 1982, Lasnik and Saito 1984, 1992, and May 1985) all *wh*-expressions have been moved to Ā (operator) positions before LF. The *wh*-expression is interpreted as a restricted interrogative quantifier; its logical scope position is identified with its LF c-command domain, and the lowest trace it binds is interpreted as a logical variable whose value is to be specified in the answer to the question.

In the past several years the unanimity with which this view was generally held has begun to dissipate, and several intriguing proposals in the literature now suggest that some *wh*-expressions are interpreted in situ. Some of the motivation for this view is found in the representational and derivational economy conditions that together constitute the core of the Minimalist Program (Chomsky 1993, 1995). Other motivation stems from detailed consideration of the link between island effects and subtle, but consistent, semantic and pragmatic factors in the interpretation of *wh*-in-situ, a line of inquiry beginning with Pesetsky 1987 and continuing with Cinque 1990, Reinhart 1993, 1995, and much other work.

In this chapter I will juxtapose these two positions on *wh*-questions and suggest (a) that both mechanisms*wh*-scope assignment via movement, and *wh*-scope assignment and interpretation without movementmust be available in the grammar of LF and (b) that they yield similar but distinct interpretational effects. In the course of advancing this view, I will have cause to closely evaluate and examine the minimalist motivations for massive reduction of LF movement, and I will suggest that the most parsimonious view of *wh*-interpretation leads us to partially reject this stance. As the repertoire of formal devices for interpretation of *wh*-operators increases, so does the need to delineate their differences. My concern

throughout will be to carefully examine the differences between the two interpretational schemas with the goal of partially identifying the best aspects of each.

1.1
Two Interpretation Mechanisms for *Wh*-in-Situ

1.1.1
Movement and In-Situ Interpretation Systems

On the classical view of the LF form of interrogatives, all *wh*-phrases are moved to the specifier position of a [+wh] C position (see, e.g., May 1985, Rizzi 1990). The *Wh*-Criterion forces each overtly in-situ *wh*-expression to raise covertly. 1 Semantically, the LF position of the *wh*-phrase corresponds to its scope position. The concept of "scope" for *wh*-phrases has two facets. First, the particular CP projection with which the *wh*-phrase associates constitutes its absolute scope position. All and only matrix-level *wh*-phrases are interpreted as contributing to the direct, to-be-answered question; all embedded interrogatives constitute nonanswered portions of complement questions. By inference, then, the LF position of a *wh*-phrase in a particular sentence can be adduced from whether or not it is answered in the response to the question. Thus, in (2) the LF position of *what* is matrix level, whereas in (1) it is embedded, as indicated by the answer pattern.

(1) Who asked where to buy what?
Answer: John did.

(2) Who asked where to buy what?
Answer: John asked where to buy books.

The *Wh*-Criterion is satisfied on either scope assignment for *what*.

Independent of this distinction in absolute scope is the *relative* scope of the *wh*-operator with respect to other operators with the same absolute scope, as discussed in May 1985 and subsequent work. Thus, (3) has two interpretations, one in which *wh* has wide scope with respect to the quantifier (3a), and one in which it has narrow scope (3b).

(3) What did everyone see yesterday?
a. [wh!x: thing(x)] ["y: person(y)] [y saw x yesterday]
b. ["y: person(y)] [wh!x: thing(x)] [y saw x yesterday]

On the analysis proposed by Higginbotham and May (1981), the scope-assigned *wh* is interpreted as a restricted interrogative operator, and the question as a whole as a statement of maximal ignorance about the true

values of the variables bound to the *wh*. The denotation of the question is a set of *theories* of the structure of the world, differing from one another in the values assigned to the variables. A complete answer to the question is information inconsistent with all but one of the theories; a partial answer is information inconsistent with at least one theory.

Singular *wh*-phrases have the same sort of presupposition as the (strong) existential quantifier *some N*¢. (4) presupposes that exactly one woman left the party at 10, and (5) that exactly one woman was seen in the garden.

(4) Which woman left the party at 10?

(5) Which woman did you see in the garden?

Central to Higginbotham and May's discussion is the rule *Absorption*, which they propose in part to solve a puzzle about the presuppositions of multiple questions in which each *wh* is singular, as in (6).

(6) Which man helped which woman?

As Higginbotham and May observe, (6) should presuppose that exactly one man helped a woman and that exactly one woman was helped by a man (and ask for the identity of each). But they note that the question has a dominant reading on which this presupposition of singularity is relaxed, and the question instead presupposes the existence of several distinct man-woman pairs that stand in the helping relation. To answer this puzzle, they propose a formal logical rule that maps two (or more) *structurally adjacent* unary quantifiers (i.e., quantifiers that bind one variable each) into binary (*n*-ary, in the general case) quantifiers. Since the logical type of the derived quantifier is different from the inputs, the singular presuppositions are relaxed, giving rise to the multiple-pair interpretation. In updated theoretical terms, the syntactic condition of structural adjacency is essentially the requirement that the two operators stand in a mutual government relation (as in May's (1985, 1989) framework), or simply that they are specifiers of the same C0 head, in minimalist terms.

The logical properties of absorbed operators are quite complex; for details, see the original discussion, as well as May 1990 and Higginbotham 1993. What is pertinent here is the following set of ideas:

(7) Absorption has a syntactic precondition of structural adjacency, which I take to be as defined in May 1985, 1989 (mutual m-command of the two operators).

(8) Absorption creates a complex *n*-ary quantifier, within which the original restrictions on the unary quantifiers are conjoined as a complex restriction and to which all the variables bound by the input quantifiers are bound.

(9) Absorption is a sufficient, and necessary, precondition to the multiple-pair interpretation of multiple questions with singular *wh*-phrases.

One other aspect of Absorption is germane to my discussion. Higginbotham and May offer Absorption as a solution to the "crossed anaphora" puzzle of Bach-Peters sentences like (10).

(10) Which pilot who shot at it hit which MIG that had chased him?

As they observe, the semantic problem with such sentences is that the question has two variable-binding operators, each of which contains a pronoun anaphoric on the other; whichever *wh*-phrase is interpreted with wider scope will thus contain a variable of the other operator, but this variable will lie outside the scope of the intended binder. Consequently, anaphora should fail for one of the pronouns. Essentially, the two distinct operators must have *some* relative scope; and this creates the problem.

Absorption offers an elegant solution. The output of Absorption is a single operator, and each pronoun is within the restriction on this operator 2 and thus may be bound in just the same way as the pronoun in simple cases like *which man who lost his pen*. Thus, the final piece of this system that we will need to keep in mind is the following theorem:

(11) Absorption is a necessary precondition for crossed anaphora.

1.1.2
Economy and Covert Wh-*Movement*

Like much other classical GB work, Higginbotham and May's theory of question semantics relies on the hypothesis that all *wh*-expressions have been moved to scope position prior to LF. This hypothesis is questioned under minimalist approaches to grammar, especially under Chomsky's (1993, 1995) feature-checking theory of *wh*-movement. Basically, in a language like English with one occurrence of obligatory overt *wh*-movement for each [+wh] C, the simplest account of a multiple question like (12) is that the [+wh] feature on C is checked by the overt raising of one *wh*-phrase, and covert raising of the in-situ *wh* is motivated (at best) by its own morphological feature specification. The semantic interpretation of the in-situ interrogative phrase alone does not constitutein strict minimalist termsmotivation for raising it covertly.

(12) [C' C[+wh] [[who] saw [which movie]]]?

This point is elaborated by Chomsky (1995), who proposes that covert movement is purely featuralonly the [+wh] feature of *which movie* can raise, leaving the lexical content in situ. 3 On the one hand, this proposal is inconsistent with the principal syntactic assumption of Higginbotham and May's system (which requires full phrasal movement, not just featural movement): on the other hand, it requires that a proper semantics be offered to interpret *wh*-in-situ. This semantics is offered by Reinhart (1993, 1995) and outlined below.

1.1.3
Choice Functions and Wh-*in-Situ*

A quite different approach to *wh*-in-situ is advanced by Reinhart (1993, 1995), who develops a different logical-semantic system for interpreting it without LF movement. Reinhart suggests that certain *wh*-phrases, specifically those corresponding to Pesetsky's (1987) D-linked operators, can remain in situ at LF, their interpretation being secured via interrogative quantification over *existential choice functions*. A choice function maps a set onto one member of that set, so that, for example, there is a choice function that maps the set of ex-presidents of the United States onto Abraham Lincoln, another that maps it onto Richard Nixon, and so forth. In Reinhart's system the *wh*-phrase in object position in (13) is interpreted via the formula (14).

(13) Who admires which former US president?

(14) [which *x*: human(*x*)] [which *f*] [*x* admires *f* (former US president)]

The N' *former US president* denotes the relevant set, and the determiner *which* (on this interpretation) is interpreted as a variable over choice functions, taking the N' denotation as its argument. Thus, informally, the question asks for specification of the identity of two things, a person and a choice function, such that the former admires the output of applying the latter to the set of former US presidents. An answer of the form *John admires Abraham Lincoln* appropriately answers the question, since identification of the output of the function is sufficient to distinguish that function from all others (from the one whose value is George Washington, from the one whose value is Thomas Jefferson, etc.).

On my reading of Reinhart's proposal, the interrogative operator *which f* is introduced in the LF component at the CP level via a sort of existential closure, so that no movement at all is involved. As a result, *wh*-phrases with the internal structure *which N*' can be interpreted without

moving them from their overtly pronounced position. As Reinhart is careful to establish, this interpretation mechanism is limited solely to this class of *wh*-expressions, explaining Pesetsky's initial generalization that island effectsparticularly Superiority effectsare ameliorated for instances of *wh*-in-situ with exactly this internal structure.

(15) I wonder what which clerk thinks we should buy.

(16) *I wonder what John bought why.

(17) *I wonder what who thinks we should buy.

As a result of this restriction, the *wh*-in-situ *which clerk* in (15) can be interpreted in situ, whereas *who* of (17)lacking the proper internal compositionmust be interpreted via LF movement, creating a Superiority effect, akin to (16).

1.1.4 *Overt Movement of D-Linked* Wh

A fact not immediately explained by Reinhart's no-movement-interpretation system is why overtly moved D-linked *wh* shows much weaker Subjacency effects than non-D-linked *wh*, as discussed in detail by Cinque (1990), among others.

(18) ?*What did you wonder who was going to buy *t*?

(19) Which book did you wonder who was going to buy *t*?

This is of course contingent on whether we take Subjacency to be a constraint on derivations or on representation (see Lasnik and Saito 1984, Freidin 1978, and Browning 1991 for discussion of each view). Suppose we take it as a constraint on (LF) representations only. Thus, a Subjacency violation could be "repaired" by structural reconstruction, that is, by undoing the overt movement of *which book* (itself driven by the strong [+wh] feature of matrix C). This results in the Subjacency-satisfying intermediate LF representation (20), which, after introduction of Reinhart's function operator, will be the interpretable (21).

(20) you wonder who was going to buy [which book]

(21) [wh! *f*] you wonder who was going to buy *f* (book)

Such repair in (18) will be possible only insofar as *what* is interpretable as D-linked, which seems to accord well with the facts. I thus suggest that this reconstructive option be included in the pre-LF operations, fitting well with the idea that Reinhart's semantics is the appropriate interpreta-

tion system for Chomsky's minimalist conception of LF. However, in discussion below I will abstract away from this point, assuming that in the unmarked case an overtly moved interrogative NP is interpreted in its surface position, without reconstruction. My main concern below will be with the LF status (syntactic and semantic) of interrogative NPs left in situ at S-Structure. Are they all covertly raised, as on the classical view? Are they all left in situ, as the strong version of Reinhart's system would suggest? Or are they in principle available for either interpretive option?

There are thus two syntactically and semantically distinct operations for interpreting *wh*-in-situ: raise it to [Spec, C] at LF and interpret it via quantification over objects satisfying the non-*wh* part of the phrase (e.g., the N'); or leave it in situ and interpret it via quantification over choice functions whose argument is the set denoted by the non-*wh* part. The first mechanism is in general available, but subject to island constraints; the second is immune from island effects, but constrained in the form of the *wh*-expressions it can apply to.

As noted above, although the two interpretive mechanisms are formally distinct in their semantics, they have considerable overlap. Either mechanism suffices to associate a *wh*-phrase with scope arbitrarily higher than the initial position of the *wh*, and either is answered by identification of an object.

The hypothesis I would like to explore is that, although the initial similarities between the two schemas make it difficult to distinguish them (particularly for simple questions with single *wh*-phrases), semantic differences do exist. Taken together, these differences will suggest that both interpretive schemas are needed in a minimalist conception of LF. I will show that LF-displaced *wh* is different from LF-in-situ *wh* in at least two crucial respects. First, only displaced *wh* can undergo logical Absorption or participate in any dependency contingent on Absorption. Second, displaced *wh* has wider relative scope than an LF-in-situ interrogative NP with which it shares the same absolute scope.

1.2 The First Puzzle: The Multipair Reading and Superiority

1.2.1 The Loss of the Multipair Reading When Superiority Is Violated

Consider the following pair, read with neutral intonation (i.e., avoiding strong focal stress on the *wh*-phrases):

(22) Which man do you think helped which woman yesterday?

(23) Which woman do you think which man helped yesterday?

The multipair reading is easily available for (22), but not for (23). 4 (23) is a question about one pair only and is thus essentially similar to the conjunction of singular questions (24).

(24) Which woman do you think a man helped yesterday, and who was he?

Recall that, on Higginbotham and May's (1981) theory, the multipair interpretation of a multiple singular question is contingent on application of the rule Absorption. In turn, Absorption is contingent on LF raising of the in-situ expression to the same [Spec, C] as the overtly moved one with which it absorbs. In structural terms, the multipair reading of (22) requires an LF representation akin to (25), in which each *wh*-phrase is in an LF-raised position in which it m-commands the other.

(25) [which woman]2 [which man]1 do you think *e*1 helped *e*2 yesterday

In the same terms, the unavailability of the multipair reading for (23) would result from the ungrammaticality of (25) as an LF representation of that sentence. Since we know that (25) is a grammatical LF representation per se, why is it not a grammatical LF representation for the S-Structure representation (23)? What is the difference between the licit derivation mapping (22) onto (25), and the illicit one mapping (23) onto (25)?

Our answer to this puzzle appeals to the major derivational economy principle proposed in Chomsky 1995, the Shortest Movement Condition.

> a can raise to target K only if there is no legitimate operation Move b targeting K, where b is closer to K [than a is]. (Chomsky 1995, 296)

(A "legitimate operation" is defined as one that results in a feature-checking configuration (Chomsky 1995, 280).) By this principle, the overt extraction of the object *wh*-phrase *which woman* to check the [+wh] feature of the matrix C0 in (23) is permitted only if there exists no shorter movement that would result in checking the same feature. Problematically, a shorter movement seemingly does exist with just this effect: movement of the subject *wh*-phrase *which man*. Before we can fully answer the question of why (23) is interpretively limited to the unipair reading, we must answer a more basic question: why is it grammatical at all?

1.2.2
Constraining the Distribution of [+wh]

One way in which the "classical" GB theory of *wh*-movement has been reencoded in the Minimalist Program is in the characterization of the distribution of the feature [+wh]: it is borne by all *wh*-words (*which, how, who*, etc.), by all maximal projections with a [+wh] specifier (*which N', how many N', how A'*, etc.), and by all C0s that head a selected question. Since all features require checking, this effectively derives the *Wh*-Criterion, since all occurrences of [+wh] on IP-internal constituents will have to raise to [Spec, CP] for feature checking.

The derivational paradox raised by the grammaticality of both (22) and (23) disappears if we permit the occurrence of [+wh] to be optional on those *wh*-expressions that need not raise to [Spec, C] for interpretation. That is, suppose the premovement representation for (23) is (26).

(26) [$C^0_{[+wh]}$ [you think [which man helped which woman$_{[+wh]}$ yesterday]]]

Here, since only one of the interrogative phrases bears the feature [+wh], the violation of the Shortest Movement Condition is only apparent. There is only one legitimate operation for (26), since "legitimate" is defined quite narrowly in terms of feature checking. As a result, the only LF representation permitted for (26) will be structurally isomorphic to its S-Structure representation, namely, (23). (23) is grammatical both as an S-Structure representation and as an LF representation, since it is derived by the shortest movement possible that checks the [+wh] feature on C0.

Let us consider the full paradigm. (22) and (23) contain two interrogative phrases. Under the proposal that [+wh] feature assignment is optional, we have four input representations, each of which maps to a unique S-Structure representation, and then to a unique LF representation. The four derivations are as follows.

(27) *Case 1:* No [+wh] feature is assigned to either interrogative NP.
Associated derivation
No legitimate movement operation will exist, and the derivation will crash because the [+wh] feature on C0 cannot be checked.

(28) *Case 2:* [+wh] is assigned to both interrogative NPs.
a. *Premovement representation*
[$C^0_{[+wh]}$ [you think [NP[+wh] which man]1 helped [NP[+wh] which woman]2 yesterday]]

b. *At S-Structure*

[[NP[+wh] which man]1 $[C^0_{[+wh]}$ [you think [*e*1 helped [NP[+wh] which woman]2 yesterday]]]

c. *At LF*

[[NP[+wh] which woman]2 [NP[+wh] which man]1 $[C^0_{[+wh]}$ [you think [*e*1 helped *e*2 yesterday]]]

d. *Associated interpretation*

unipair (if Absorption does not apply)
multipair (if Absorption applies)

(29) *Case 3:* [+wh] is assigned to *which man*, and not to *which woman*.

a. *Premovement representation*

$[C^0_{[+wh]}$ [you think [NP[+wh] which man]1 helped [NP which woman]2 yesterday]]

b. *At S-Structure*

[[NP[+wh] which man]1 $[C^0_{[+wh]}$ [you think [*e*1 helped [NP[+wh] which woman]2 yesterday]]]

c. *At LF*

[[NP[+wh] which man]1 $[C^0_{[+wh]}$ [you think [*e*1 helped [NP[+wh] which woman]2 yesterday]]]

d. *Associated interpretation*

unipair only (structural preconditions for Absorption unmet)

(30) *Case 4:* [+wh] is assigned to *which woman*, and not to *which man*.

a. *Premovement representation*

$[C^0_{[+wh]}$ [you think [NP which man]1 helped [NP[+wh] which woman]2 yesterday]]

b. *At S-Structure*

[[NP[+wh] which woman]2 $[C^0_{[+wh]}$ [you think [[which man]1 helped *e*2 yesterday]]]

c. *At LF*

[[NP[+wh] which woman]2 $[C^0_{[+wh]}$ [you think [[which man]1 helped *e*2 yesterday]]]

d. *Associated interpretation*

unipair only (structural preconditions for Absorption unmet)

The in-situ subject in (23) is blocked from raising to [Spec, C] by the fact that it was "passed over" in the operation of Move that raised the object interrogative. This in turn requires that this in-situ interrogative remain in situ throughout the derivation, including LF. As a result, the

structural conditions for Absorption are not met in any derivation whose well-formed S-Structure representation is (23). Since Absorption is the formal mechanism underlying the multipair interpretation of multiple interrogatives, such "anti-Superiority" examples are limited to the unipair interpretation.

One formal point remains to be discussed here. Since (e.g.) *which man* in (23) cannot move to [Spec, CP] at LF, it is interpreted in situ. This means that the LF representation for (23) is basically as follows, filling in the choice function operator and its variable:

(31) [[NP[+wh] which woman]2 (wh*f*) $[C^0_{[+wh]}$ [you think [[*f*(man)]1 helped *e*2 yesterday]]]

Once we include this interrogative operator over choice functions, it is not immediately clear why Absorption should not occurtwo interrogative operators are present in the same scopal position, seemingly m-commanding one another. Why can't they absorb, giving rise to the multipair reading?

The answer to this question will follow from a constraint that May (1990, 125) imposes on Absorption, called S-Invariance. This constraint requires type parallelism between the two input operators and the output operator. In (31) the two interrogatives are vastly different: one is a quantifier over individuals, restricted by a set, and the other is an unrestricted quantifier over choice functions. This type difference in the two operators is sufficient to block Absorption, as desired. 5

We now have achieved a welcome result: the asymmetry in interpretation between (22) and (23) follows from economy constraints on LF derivations. This particular way of achieving this result depends on the view that the [+wh] feature is genuinely morphological: like abstract Case, its presence or absence on an NP is formally distinct from its semantic interpretation. An interrogative expression may, or may not, bear this feature. Interrogatives bearing [+wh] must raise to [Spec, CP], must participate in Shortest Movement Condition effects, and are interpreted via Higginbotham and May's system of restricted quantification (over individuals, in the cases discussed here). Interrogatives not bearing this feature cannot raise to [Spec, CP] (or at least cannot remain there at LF), do not necessarily participate in Shortest Movement Condition effects, and must be interpreted in situ, via the choice function option.

In the next sections I will consider three additional puzzles, each having to do with asymmetries in interpretation with multiple interrogatives. I

will show that each puzzling asymmetry is accounted for directly under the mixed LF system advocated above.

1.3
The Second Puzzle:
A Crossed Binding Asymmetry

Recall that Absorption plays a key role in resolving the seeming paradox of crossed anaphora, in cases like (10), repeated here.

(10) [Which pilot who shot at it1]2 hit [which MIG that had chased him2]1?

After Absorption each pronoun is within the restriction of the single, binary operator and thus constitutes a properly bound variable. This is the one principled exception to a general fact about operator-bound pronominals, namely, that a pronoun inside the subject of a clause cannot be operator-bound to a quantified or interrogative object, as shown in (32) (33).

(32) *Which pilot who shot at it2 hit [which MIG]2?

(33) *[His1 mother] loves [every boy]1.

(See May 1990 for additional discussion; see also note 6.)
Now consider the following examples:

(34) [Which student who took it2]1 do you think [*t*1 flunked [which test that had frightened him1]2] rather badly?

(35) ?*[Which test that had frightened him1]2 do you think [which student who took it2]1 flunked *t*2 rather badly?

In (34) *wh*-movement respects Superiority; thus, in the framework advanced here the in-situ interrogative can raise to [Spec, C] and undergo Absorption (licensing crossed binding). In (35), however, Superiority is violated, "freezing" the subject *wh* in situ and forcing it to be interpreted in that position (via the choice function operator). The Absorption-contingent analysis of crossed binding predicts, correctly, that crossed binding should fail for (35).

(34) is unproblematic: the in-situ *wh* raises at LF and absorbs with the overtly moved operator, and each pronoun is properly bound within the derived binary operator. The puzzling case is the ungrammatical (35). The puzzle is explained directly under present assumptions. Since the lower interrogative NP has been overtly raised to [Spec, CP], the embedded

subject interrogative must not bear the feature [+wh]. As a result, it cannot raise in the LF component and thus cannot absorb with the overtly moved interrogative. Crossed binding occurs only as a result of Absorption; thus, it cannot occur here, and anaphora fails.

Note the import of both puzzles outlined above. In each case an asymmetry in interpretive options arises in pairwise comparisons of "normal" and anti-Superiority patterns of *wh*-extraction. Any attempt to interpret the in-situ occurrences of *wh* in the example pairs via the *same* logical interpretation mechanisms encounters difficulty. It is only in a system that incorporates both a method for interpreting raised *wh* and a method for interpreting *wh*-in-situ that asymmetries of this sort can be accommodated. Hence, we conclude that two mechanisms are available for in-situ interpretation: covert *wh*-raising of the entire *wh*-phrase to [Spec, C], and in-situ interpretation via choice functions. Fundamentally, the empirical asymmetries justify formal interpretive asymmetries. 6

1.4
The Third Puzzle: Asymmetry in *Wh*-QNP Interpretations

Consider the following pair:

(36) [Which professor]2 did everyone tell *e*2 to grade [which student]1 strictly?

(37) [Which student]1 did everyone tell [which professor]2 to grade *e*1 strictly?

(36) is multiply ambiguous.7 Of particular interest are two available interpretations on which the *wh*-phrases have absorbed (yielding the multipair reading). On one of these, the universal QNP has wider scope than the complex *wh*-operator; on the other, the scope orders are reversed. Informally, these two readings are as follows:

(38) a. i. which pairs <*x*, *y*>, *x* a student & *y* a professor, are such that for every person *z*: *z* told *y* to grade *x* strictly
ii. "Everyone told Prof. Smith to grade Jones strictly, everyone told Prof. Thompson to grade O'Reilly strictly, . . ."
b. i. for every person *z*: which pairs <*x*, *y*>, *x* a student & *y* a professor, are such that *z* told *y* to grade *x* strictly
ii. "As for Sally, she told Prof. Smith to grade O'Reilly strictly and Prof. Thompson to grade Bush strictly. David told Prof. Jennings to grade Chung strictly, and told . . ."

What is important here is that the *wh*-operator that permits the multipair reading can have wide or narrow scope with respect to the universal quantifier. The distribution across multiple pairings of students and professors is thus independent of the relative scope of the operators. Example (36) exhibits the normal pattern of *wh*-extraction, respecting Superiority. Within the current framework this means that the in-situ expression can raise at LF, and the two interrogatives can absorb, forming a binary *wh*-operator. The two readings follow.

In example (37) neither of these readings exists. On the system advanced here, this is precisely as expected: both of these readings are contingent on Absorption, Absorption is contingent on LF raising of the *wh*-phrase, and LF raising of the *wh*-phrase is blocked by the interplay of the optional assignment of the morphological [+wh] feature and the Shortest Movement Condition.

On my judgment, (37) has a four-way ambiguity, resting on the relative scope of the universal quantifier and the two unabsorbed *wh*-operators. None of these four readings permits multipair answers. 8

(39) a. i. for which *x*, *x* a professor, for which *y*, *y* a student, for every person *z*: *z* told *x* to grade *y* strictly
ii. "Everyone told Prof. Smith to grade Johnson strictly."
b. i. for every person *z*, for which *x*, *x* a professor, for which *y*, *y* a student: *z* told *x* to grade *y* strictly
ii. "Sally told Prof. Smith to grade Johnson strictly, David told Prof. Jennings to grade Chung strictly, Ellen . . ."
c. i. for which *y*, *y* a student, for every person *z*, for which *x*, *x* a professor: *z* told *x* to grade *y* strictly
ii. "Sally told Prof. Smith to grade Johnson strictly, David told Prof. Jennings to grade Johnson strictly; Ellen told Prof. Thompson to grade Johnson strictly, . . ."
d. i. for which *x*, *x* a professor, for every person *z*, for which *y*, *y* a student: *z* told *x* to grade *y* strictly
ii. "Sally told Prof. Smith to grade Jennings strictly, David told Prof. Smith to grade Jones strictly, Ellen told Prof. Smith to grade O'Reilly strictly."

1.5 The Fourth Puzzle: The Comparative Superlative

The final interpretive asymmetry that emerges under the anti-Superiority extraction pattern is one affecting the construal of the comparative

superlative. Szabolcsi (1986) (to whom the term *comparative superlative* is due) draws attention to the comparative use of the superlative construction, as in (40).

(40) a. John climbed the highest mountain.
b. It was John who called Mary the earliest.

As Szabolcsi observes, (40) has two distinct interpretations. On one reading, dubbed the *absolute superlative* reading, John climbed the highest among all mountains, namely, Everest. On the other reading, he simply climbed the highest mountain among those climbed by the other climbers in his comparison set. This is the *comparative superlative* (CS) interpretation, the type of reading I will focus on here. 9

Szabolcsi observes that the CS reading is contingent on the presence of a free variable within the scope of the CS-related operator. More precisely, the CS encodes a comparison across elements satisfying a variable *x*, only if *x* is within the scope of the CS, and the CS is within the scope of the operator that binds *x*. For example, in (40a) the subject *John* is analyzed as the default focus, so a partial interpretation of (40a) is as in (41).

(41) for *x* = John, *x* climbed the highest mountain

Presence of the CS operator (glossed as EST, following Szabolcsi's notation) triggers a conversion of the open sentence *x climbed . . .* into a comparison across members of a (discourse-specified) set containing John. The full logical representation of (41) is essentially (42).

(42) for *x* = John, *x* climbed a higher mountain than any other *y*

In general, Szabolcsi argues, a sentence in which the schema (43) is instantiated will be interpretable essentially as (44).

(43) . . . OP*x* [. . . EST [*P*(*x*) . . .]]

(44) *P*(*x*) & "*y* ¹ *x* [more (*P*(*x*)) than (*P*(*y*))]

Szabolcsi further observes that the CS reading is maximally available when the variable ranges over individuals, and is degraded when the variable ranges over different types. Thus, there is an obvious contrast between (45), with a clear CS interpretation (comparing across various values for the *wh*-variable), and (46), which disfavors the same sort of interpretation.10

(45) [Which man]1 did you introduce *e*1 to the most women?
(i.e., "Which man did you introduce to more women than you introduced any other man to?")

(46) [Why] did you introduce John to the fewest women?
(#"What is the reason *r* such that you introduced John to fewer women for reason *r* than you introduced him to for any other reason?")

This limitation will play a crucial role below.

With that much in place (the crucial scope relation between the CS operator and the other operator, and the apparent restriction of the CS to variables over individuals), consider the following pairs:

(47) a. [Which man]1 do you think [*e*1 gave the most flowers to [which woman]2 yesterday]?
b. [Which woman]2 do you think [which man]1 gave the most flowers to *e*2 yesterday?

(48) a. [Which newspaper]1 does Karl believe *e*1 published [which story]2 the earliest?
b. [Which story]2 does Karl believe [which newspaper]1 published *e*2 the earliest?

In the (a) examples the dominant reading is one in which comparison is across pairs (of men and women, or newspapers and stories). The proper answer to (48a), for example, will be a pair <a, b>, such that *a* is a newspaper, *b* is a story, and *a* published *b* earlier than any other newspaper published its story. Similarly, the proper answer to (47a) will be specification of a male-female couple such that more flowers were given by the man to the woman than were given within any other couple. This follows straightforwardly, 11 since in the LF representation for the (a) examples (see (49), for (47a)) both interrogative NPs will be moved to matrix [Spec, CP], binding variables over individuals in the lower clause. The CS operator can thus determine a comparison across sets of pairs (of men and women, or newspapers and stories).

(49) [which woman]2 [which man]1 do you think *e*1 gave the most flowers to *e*2 yesterday

The (b) sentences have a markedly different dominant reading. In each case the comparison is only over alternative values of the variable bound by the overtly moved interrogative. In (47b), for example, the comparison is not across man-woman pairs, but across individual women (e.g., "Which woman do you think a man gave more flowers to than any other woman, and who was he?"). Why should this be?

Again, following the argument that the overtly in-situ interrogative in the (b)-type examples is trapped in situ in the LF component, the only option for interpreting *which man* is via Reinhart's (1993, 1995) existential choice function mechanism. Thus (ignoring the CS operator for the moment), the semantic representation for (47b) will be (50).

(50) [which woman]2 (wh*f*) do you think *f*(man) gave the most flowers
to *e*2 yesterday

Here, there is one variable over individuals, and one variable over functions. Recall that Szabolcsi (1986) independently argued (see discussion of (46)) that the CS interpretation degrades, or disappears altogether, when the variable in its scope is anything other than a variable over individuals. I suggest that this is exactly the explanation for the (a) (b) contrasts above. In (50) there is only one variable over individuals (because of the interplay between the Shortest Movement Condition and the two interpretive options for interrogatives), corresponding exactly to the locus of comparison in the interpretation of (47b). Only a theory of interrogative interpretation that can divide the operators' variables across precisely these lines is capable of capturing these sharp distinctions.

1.6
Conclusion

Current literature on interrogative interpretation contains two formally different mechanisms for interrogative NP construal: interpretation as a moved restricted quantifier and interpretation in situ, via existential closure over choice functions. In this chapter I have advocated a hybrid theory of *wh*-interpretation, arguing that both of these leading approaches to interrogative interpretation are necessary elements in a proper theory of LF. Rather than viewing these two theoretical devices as rivals, one should, I have suggested, take them as complementary devices. Each has approximately the same overall function, assigning logical scope to an interrogative operator. They differ in subtle but important ways, including their capacity to license comparative readings of the superlative, their capacity to participate in multipair interpretations, their ability to commute in scope with noninterrogative quantifiers, and their interaction with pronominal anaphora.

Notes

A note on the origins of the work presented in this chapter. A very early version of the chapter, drawing rather different conclusions, was presented as a NELS talk in

1990 and (in emended form) at colloquia at the University of Arizona, the University of Connecticut, and the University of Washington in 1991 and 1992. The material in its current form was developed in 1994, 1995, and 1997 seminars at the University of Arizona. I am grateful to audiences at all those presentations for their feedback. I am particularly indebted to several people whose detailed comments on earlier versions of this chapter, and careful consideration of the data it presents, have added greatly to it: Howard Lasnik, Richard Larson, Jim Higginbotham, David Pesetsky, Janet Fodor, Janet Nicol, Mario Montalbetti, and Joe Emonds.

It is indeed a pleasure to be able to contribute this chapter to a volume honoring Howard Lasnik for his many fundamental contributions to linguistics during the first twenty-five years of his astounding career. Here's to you, Howard.

1. The *Wh*-Criterion is stated as follows (May 1985, Rizzi 1990):

(i) At LF every [+wh] C0 must have a [+wh] specifier.

(ii) At LF every [+wh] projection must be in specifier position of a [+wh] C0.

2. Specifically, the output binary *wh*-operator will be of the form in (i), via Higginbotham and May's theory.

(i) [wh *x*, wh *y*: *x* a pilot who shot at *y* & *y* a MIG that chased *x*]

The second occurrences of the variables *x* and *y* correspond to the logical translations of the pronouns in (10). Note that each such variable is within the scope of its binding *wh*-operator, hence the possibility of grammatical anaphora.

3. To elaborate on this point, under Chomsky's (1995) system it is the feature [+wh] itself that must raise at LF. The feature is classified as interpretable and is thus subject to LF raising. Since LF movement is separated from the Spell-Out operation, only the feature will raise, leaving all associated lexical content in situ (i.e., there is no pied-piping in the LF component under Chomsky's system, including no raising of the *wh*-determiner or its associated nominal restriction). The semantics of Higginbotham and May's Absorption operation requires full phrasal raising in LF, since the raised operators are combined in part by conjoining their restrictions external to the scope of the binary operator thus formed.

4. This empirical point was initially made in Barss 1990; it was made independently in unpublished work by Phil Branigan at about the same time. As the judgments on such examples are complicated, I am grateful to the many speakers who provided judgments, which confirm the empirical points made here.

5. To make a similar point in a slightly different way, suppose that the choice function operator is explicitly identified as such in the sentence's logical representation. We would then replace (31) by something like (i).

(i) [wh!*f*: "*Z*"*z*: *f*(*Z*) = *z* ® *z* Î *Z*][wh!*y*:woman(*y*)] you think [*f*(man) helped *y*]

As before, the operator corresponding to the overtly moved *wh*-expression is a quantifier over individuals, restricted by a set. The other operator is a quantifier over functions and is restricted in the manner indicated. Either way, the type difference between the two operators is obvious.

6. An anonymous reviewer observes that the analysis of crossed binding I assume here (the Absorption analysis of Higginbotham and May (1981) and May (1985)) has been criticized in certain details by Clark and Keenan (1986) and by Epstein (1989). The problem for Absorption raised by Clark and Keenan is addressed explicitly by May (1990), who presents, as a solution, the principle of S-Invariance. I believe the core of the objection is not germane to the flow of argument in this chapter, and in any event May's solution is invoked here (see section 1.2.2) for independent reasons (principally to prevent absorption of type-distinct *wh*-operators).

Epstein's objection to Higginbotham and May's system is more intricate and is directly relevant to the present discussion, as it is aimed squarely at the analysis of crossed binding. The two central examples in Epstein's article relevant to the assumptions I make in the text are (ia) and (iia) (see Epstein's (29)); these are the LF representations of (ib) and (iib).

(i) a. [S[NP his2 mother] [VP[everyone]2 [VP saw *e*2]]]
b. *His2 mother saw everyone2.

(ii) a. *[S'/CP[which pilot who shot at it3]2 [S *e*2 [VP[every MIG that chased him2]3 [VP hit *e*3]]]]
b. [Which pilot who shot at it3]2 hit [every MIG that chased him2]3?

Epstein observes that the theory of adjunction employed by May (1985), together with May's invocation of the Path Containment Condition as the principle ruling out weak crossover violations, would seem to illicitly permit bound anaphora in (i), a classic example of weak crossover. In May's system the VP-adjoined quantifier takes, as its maximal scope, all the material it m-commands, namely, everything within S. Consequently, the pronoun in (i) is within the absolute scope domain of *everyone*. Since the paths from the pronoun to QNP, and from the trace to QNP, do not illicitly interact, nothing blocks this type of example, as May analyzes all cases of weak crossover as violations of the Path Containment Condition. Epstein concludes that the Path Containment Condition is too weak to be the correct principle governing weak crossover and instead adopts Higginbotham's (1980) theory, where a pronoun may be dependent on a QNP only if it occurs to the right of the QNP (at LF). Via this linear-order-dependent principle, the pronoun in (i) cannot function as a bound variable. The first problem is thus solved.

The relatedness of this argument to my discussion emerges when we consider how Epstein treats crossed binding. Epstein argues that crossed binding, as exemplified in (ii), is not permitted, since no matter how the operators are arranged at LF, one will be to the right of the corresponding pronoun and thus incapable of binding that pronoun as a variable. Essentially, Epstein argues that *all* cases of crossed binding in the literature should be viewed as ungrammatical. As further support for this conclusion, he suggests that (on his judgments) they are, in fact, all unacceptable.

With regard to Epstein's empirical claim, my judgments differ from his, instead according with those cited in Higginbotham and May 1981, May 1985, 1989, and the text of this chapter (which reflect my judgments and those of the colleagues cited in the acknowledgments).

With regard to Epstein's theoretical claim (that principles of pronominal-variable binding such as Higginbotham's (1980) block crossed binding), he is surely correct, *with respect to pre-Absorption operators*. Indeed, one of the central initial motivations for Absorption was the phenomenon of crossed binding. However, in the post-Absorption representations for cases like (ii), there is no weak crossover violation. Absorption is a principled (and independently motivated) means of avoiding weak crossover violations. The lack of weak crossover effects in examples like (ii) is a consequence of the operation that converts the two unary operators into one binary operator. This correlation between Absorption and the suppression of weak crossover is central to the explication I give in the text of examples like (34) and (35).

Thus, although Clark and Keenan's and Epstein's arguments have advanced our understanding of the interplay between the syntax of LF, the semantics of quantification, and pronominal variable binding, they do not adversely affect the core components of the theory of Absorption crucial to my conclusions in this chapter.

7. In addition to the two readings delineated in (38a,b), this example has the four interpretations given in (39a d). I focus in the text on (38a,b) because these are the two readings *not* available for example (37).

8. In the sense used here. When one or both *wh*-operators have scope narrower than that of *everyone*, the wide-scope quantifier will distribute answers to the question across members of the domain of quantification. This distribution is fundamentally different from the multipair reading produced via Absorption, since in this case there will be one student/professor for each element in the domain of *everyone*, rather than a series of pairs, which is what the true multipair reading would provide.

9. The absolute superlative reading is suppressed in many cases, as in (i).

(i) John called Mary the earliest.

The CS reading is that on which John called Mary earlier than anyone else in his comparison set did. The absolute reading is the absurd proposition that John called Mary at the time *t* that is prior to all other distinct times *t'*.

10. If *John* is read as focus, there is a CS reading available for (46), basically as in (i).

(i) Why did you introduce John to fewer women than you introduced anyone else to?

The CS reading in question, however, is the one that is comparative across reasons, as discussed in the text.

11. In her original discussion, Szabolcsi (1986) states that the CS reading is strictly local and cannot be licensed across a tensed clause boundary. I disagree with this judgment, as do the other speakers I have consulted, for cases of the following sort:

(i) a. Who said that Mary told John the most jokes?
b. for which *x*: *x* said that Mary told John more jokes than any other *y* did

Indeed, there is apparently a subtle de dicto de re contrast when the variable is within the lower clause and the *wh*-phrase is given wide scope.

(ii) a. Who does Mary think John told the most jokes to?
b. De dicto: for which *x*: Mary thinks [John told more jokes to *x* than he told to anyone else]
c. De re: for which *x*: Mary thinks John told more jokes to *x* than she thinks he told to anyone else

On the de dicto construal the numerical comparison is within Mary's belief-world; on the de re construal the numerical comparison is attributed by the speaker. This distinction is discussed in more detail in Barss, in preparation.

References

Barss, Andrew. 1990. Optional movement, absorption, and the interpretation of *wh*-in-situ. Paper presented at NELS 21, Université du Québec à Montréal; ms., University of Arizona, Tucson.

Barss, Andrew. In preparation. Derivations and dependencies. Ms., University of Arizona, Tucson.

Browning, M.A. 1991. Bounding conditions on representation. *Linguistic Inquiry* 22, 541 562.

Chomsky, Noam. 1976. Conditions on rules of grammar. In *Essays on form and interpretation*. New York: North-Holland.

Chomsky, Noam. 1977. On *wh*-movement. In Peter Culicover, Thomas Wasow, and Adrian Akmajian, eds., *Formal syntax*. New York: Academic Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale and Samuel Jay Keyser, eds., *The view from Building 20*. Cambridge, Mass.: MIT Press. [Also in Noam Chomsky, *The Minimalist Program*. Cambridge, Mass.: MIT Press (1995).]

Chomsky, Noam. 1995. Categories and transformations. In *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Cinque, Guglielmo. 1990. *Types of Ā-dependencies*. Cambridge, Mass.: MIT Press.

Clark, Robin, and Edward Keenan. 1986. The Absorption operator and Universal Grammar. *The Linguistic Review* 5, 113 136.

Epstein, Samuel David. 1989. On adjunction and pronominal variable binding. *Linguistic Inquiry* 20, 307 319.

Freidin, Robert. 1978. Cyclicity and the theory of grammar. *Linguistic Inquiry* 9, 519 549.

Higginbotham, James. 1980. Pronouns and bound variables. *Linguistic Inquiry* 11, 679 708.

Higginbotham, James. 1993. Interrogatives. In Kenneth Hale and Samuel Jay Keyser, eds., *The view from Building 20* . Cambridge, Mass.: MIT Press.

Higginbotham, James, and Robert May. 1981. Questions, quantifiers, and crossing. *The Linguistic Review* 1, 41 80.

Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT, Cambridge, Mass.

Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15, 235 289.

Lasnik, Howard, and Mamoru Saito. 1992. *Move x*. Cambridge, Mass.: MIT Press.

May, Robert. 1985. *Logical Form: Its structure and derivation*. Cambridge, Mass.: MIT Press.

May, Robert. 1989. Interpreting Logical Form. *Linguistics and Philosophy* 12, 387 435.

May, Robert. 1990. A note on quantifier absorption. *The Linguistic Review* 7, 121 127.

Pesetsky, David. 1987. *Wh*-in-situ: Movement and unselective binding. In Eric Reuland and Alice ter Meulen, eds., *The representation of (in)definiteness*. Cambridge, Mass.: MIT Press.

Reinhart, Tanya. 1993. Interpreting *wh*-in-situ. Ms., Tel Aviv University.

Reinhart, Tanya. 1995. Interface strategies. (OTS Working Papers of Theoretical Linguistics 95-002). Utrecht Institute of Linguistics OTS, Utrecht University.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.

Szabolcsi, Anna. 1986. Comparative superlatives. In Naoki Fukui, Tova R. Rapoport, and Elizabeth Sagey, eds., *Papers in theoretical linguistics*. (MIT Working Papers in Linguistics 8.) MITWPL. Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Chapter 2
Sometimes in [Spec, CP], Sometimes in Situ

Zeljko Boskovic *

Languages differ systematically with respect to how they treat *wh*-phrases in multiple questions. In English one and only one *wh*-phrase undergoes fronting in multiple questions.

(1) Who did John give what?

In Slavic languages, on the other hand, all *wh*-phrases undergo fronting, though they do not all land in [Spec, CP] in all Slavic languages. (See Rudin 1988. In Bulgarian, from which (2) is drawn, all fronted *wh*-phrases are located in [Spec, CP].)

(2) Koj kakvo e kupil?
who what is bought
'Who bought what?'

Languages such as Japanese show the opposite pattern: they allow all *wh*-phrases to remain in situ in multiple questions.

(3) John-wa dare-ni nani-o ageta ka?
John-TOP who-DAT what-ACC gave Q
'Who did John give what?'

Finally, there are "mixed" languages such as French, which displays both the English and the Japanese patterns. Thus, both (4a), where one *wh*-phrase moves to [Spec, CP], and (4b), where all *wh*-phrases remain in situ, are good as true nonecho multiple questions. The Bulgarian pattern, on the other hand, is not attested in French, as illustrated by (4c).

(4) a. Qu' a-t-il donné à qui?
what has-he given to whom
'What did he give to whom?'
b. Il a donné quoi à qui?
he has given what to whom
c. *Qu'à qui a-t-il donné?

In this chapter I will examine *wh*-constructions in French. It turns out that the English and the Japanese patterns are not always both available in French. In fact, whereas the English pattern is always available, the distribution of the Japanese pattern is rather limited, which suggests that the two languages employ different strategies in in-situ questions. The goal of the chapter is to determine when the Japanese pattern is allowed in French and to account for the limited distribution of the in-situ strategy in this language. I will show that the analysis I propose has interesting theoretical consequences. It provides evidence that lexical insertion can take place in LF. In fact, I will show that even elements with strong features are capable of entering structure in LF under certain well-defined conditions. The analysis will also provide evidence that feature movement is subject to locality conditions on movement. I will also examine consequences of the proposed analysis of French for the principles of economy, as well as pure syntactic *wh*-movement languages such as English and pure *wh*-in-situ languages such as Japanese.

2.1
Embedded Questions in French

As noted above, in simple matrix questions French allows both the English and the Japanese patterns regarding *wh*-phrases in multiple questions. Thus, both (5a) and (5b) are grammatical. It is well known, however, that the Japanese pattern is disallowed in embedded questions. Whereas (6a) is acceptable, (6b) is not; the latter sharply contrasts with (5b) and Japanese (7).

(5) a. Qui as-tu vu?
whom have-you seen
'Who did you see?'
b. Tu as vu qui?

(6) a. Pierre a demandé qui tu as vu.
Pierre has asked whom you have seen
'Pierre asked who you saw.'
b. *Pierre a demandé tu as vu qui.

(7) Peter-wa [anata-ga dare-o mita ka] tazuneta.
Peter-TOP you-NOM who-ACC saw Q asked
'Peter asked who you saw.'

The contrast between (5b) and (6b), which has always resisted a satisfactory account (see note 4), follows given certain minimalist assumptions in

Chomsky 1995. In particular, the contrast follows from the way lexical insertion is performed in the minimalist system and from Chomsky's definition of strong features. In the minimalist system, which dispenses with D-Structure, lexical insertion takes place through the operation Merge, which is defined in such a way that it always must expand structure (i.e., make the existing tree bigger). As a result, Merge can take place only at the root of a tree; it cannot take place in embedded positions. In non-minimalist terms, it is subject to the cycle. Merger generally takes place in overt syntax. As noted in Chomsky 1995, this follows without stipulation. Thus, if an NP such as *John* is inserted in LF, the derivation crashes because LF cannot interpret phonological features of *John*. If, on the other hand, *John* is inserted in PF, PF will not know how to interpret the semantic features of *John*. The only way to derive a legitimate PF representation and a legitimate LF representation is for *John* to be inserted before the level of S-Structure is reached. PF then strips off the phonological features of *John*, and the semantic features of *John* proceed into LF. This line of reasoning allows lexical insertion to take place in PF and LF under certain conditions. To be more precise, it allows PF insertion of semantically null lexical elements and it allows LF insertion of phonologically null elements. Here we are interested in the second possibility. (Chomsky actually suggests that PF insertion should be altogether banned, though without giving empirical justification. The *do of do*-support, however, appears to be a plausible candidate for PF insertion.)

Returning now to (5b), notice that the complementizer in (5b) is phonologically null. Since *wh*-movement must take place overtly in at least some constructions in French, I assume that the [+wh] feature is strong in French. Suppose now that we insert the complementizer with strong [+wh] feature in the LF representation of (5b). In Chomsky's (1994) system this would not be possible. Chomsky (1994) considers strong features illegitimate LF objects. As a result, a derivation that enters LF with a strong feature necessarily crashes. This is, however, not the case under Chomsky's (1995) approach to strong features, which I will adopt here. 1 Chomsky (1995) defines strong features derivationally. In particular, he defines strong features as features that are not tolerated by the derivation and therefore must be eliminated via checking as soon as they are inserted into the tree. This conception of strong features allows strong features to be inserted in LF as long as they are checked immediately upon insertion. Nothing, then, seems to prevent the complementizer with a strong [+wh] feature from being inserted in the LF representation of (5b).2 Once the

complementizer is inserted, the *wh*-phrase moves into the checking domain of the complementizer, checking its strong [+wh] feature. Note also that since the lexical insertion of the complementizer takes place at the root of the tree, it is a legitimate instance of Merge. Under this analysis, (5b) then has the S-Structure and LF representations in (8a) and (8b), respectively. The reason why no overt *wh*-movement takes place in (5b) is trivial. The CP projection, containing the [+wh] feature, is not present in the overt syntax.

(8) a. S-Structure: [IP tu as vu qui]
b. LF: [CP qui C [IP tu as vu]]

Notice that in-situ questions in French are not restricted to arguments. Thus, (9) is also a well-formed nonecho question. The construction in (9) can be derived in the same way as (5b), with the LF insertion of the complementizer with a strong [+wh] feature. 3

(9) Jean a réparé la voiture comment?
Jean has fixed the car how
'How did Jean fix the car?'

Now consider embedded questions such as (6b). In order to account for the ungrammaticality of (6b), we need to rule out the derivation in which the interrogative complementizer is inserted into the embedded clause in LF. The derivation is in fact straightforwardly ruled out because it involves an instance of Merge that does not expand the tree (i.e., applies in an embedded position). To satisfy the requirement that Merge expand the structure, the complementizer with strong [+wh] feature must be introduced into the structure in (6b) in overt syntax before the higher structure is built. The *wh*-phrase then must move overtly to [Spec, CP] to check the strong feature of the interrogative complementizer as soon as the complementizer is merged into the structure. Overt *wh*-movement is clearly the only option with embedded questions. The contrast between (5b) and (6b) is thus straightforwardly accounted for.4

The crucial fact that the interrogative complementizer is phonologically null in the constructions under consideration is what enabled us to delay the insertion of the complementizer until LF in (5b), which eventually made it possible to account for the contrast between (5b) and (6b). That the fact that the complementizer is phonologically null in the constructions in question plays a crucial role in the phenomena under consideration and is not just a coincidence is confirmed by data concerning *wh*-constructions in Serbo-Croatian (SC).

SC is a multiple *wh*-fronting language that, as shown in Rudin 1988, cannot place more than one fronted *wh*-phrase in [Spec, CP]. In Boskovic * 1997b I show that SC exhibits an interesting pattern with respect to the Superiority Condition. In particular, SC exhibits Superiority effects exactly in those contexts where *wh*-movement must take place in French. (This holds for all the contexts discussed below in which French displays obligatory *wh*-movement.) Where French does not have to have syntactic *wh*-movement, SC does not exhibit Superiority effects. This is illustrated in (10) (11), which show that fronted *wh*-phrases in SC are freely ordered in matrix questions but not in embedded questions.5

(10) a. Ko sta* kupuje?
who what buys
'Who buys what?'
b. Sta ko kupuje?*

(11) a. Zavisi od toga [ko sta kupuje].
depends on it who what buys
'It depends on who buys what.'
b. ?*Zavisi od toga [sta ko kupuje].
c. [Ko koga voli], taj o njemu i govori.
who whom loves that-one about him even talks
'Everyone talks about the person they love.'
d. ?*[Koga ko voli], taj o njemu/o njemu taj i govori.
e. Ima [ko sta da kupi].
has who what PARTICLE buys
'Someone has something to buy.'
f. ?*Ima [sta ko da kupi].

To account for the parallelism between the contexts in which SC exhibits Superiority effects and the contexts in which *wh*-movement must take place in French, I propose in Bošković 1997b that SC is a French-type language with respect to when *wh*-movement must take place. Given that, as in French, *wh*-movement does not have to take place in SC matrix questions, the Superiority Condition is trivially satisfied in (10): Since, as in French, *wh*-movement must take place in SC embedded questions, the Superiority Condition is operative in (11). (11b,d,f) are then ruled out by the Superiority Condition. Given this analysis, the only difference between French and SC is that even *wh*-phrases that do not move overtly to [Spec, CP] in SC must undergo fronting. Thus, no *wh*-phrase is allowed

to stay in situ in (10) (11). For example, **Ko kupuje* sta * 'Who buys what?' is ungrammatical. This fronting of *wh*-phrases is clearly independent of *wh*-movement, as indicated by the fact that the construction is ungrammatical even on the echo question reading of *sta*. The same holds for **Jovan kupuje sta* 'Jovan buys what?'. For the purposes of this chapter we may assume that fronting of SC *wh*-phrases that does not land in [Spec, CP] takes place in PF, as suggested in Boskovic* 1997b, and thus follows *wh*-movement. Stjepanovic* (1995) shows that the fronting takes place for focusing reasons. In what follows I will ignore this focus fronting of *wh*-phrases and concentrate on *wh*-movement (i.e., movement to [Spec, CP]), which, as suggested above, takes place in (11) but not in (10), SC being a French-type language with respect to when *wh*-movement must take place. As noted above, this enables us to account for the contrast between (10) and (11) with respect to Superiority.

Note that the interrogative complementizer in SC (10) (11) is phonologically null. As a result, the analysis of the contrast between French (5b) and (6b) with respect to the obligatoriness of *wh*-movement can be readily extended to the SC constructions in question. SC, however, also has a phonologically realized interrogative complementizer *li*.6 If the LF C-insertion analysis is on the right track, we would expect that, in contrast to (10), matrix *li* questions will exhibit Superiority effects, a phenomenon that reveals when *wh*-movement must take place in SC. Since *li* is not phonologically null, it must be introduced into the structure overtly, in contrast to phonologically null complementizers. Overt movement to [Spec, CP] then must take place in *li* questions to eliminate the strong [+wh] feature. (Since SC does not differ from French with respect to when overt *wh*-movement must take place, I assume that the [+wh] feature in SC is strong.) As illustrated in (12), the prediction is borne out. Example (12b) clearly contrasts with (10b) on the true question reading of the second *wh*-phrase.

(12) a. Ko li sta kupuje?
who C what bought
'Who on earth buys what?'
b. ?*Sta li ko kupuje?*

The contrast between (10b) and (12b) indicates that in order for the *wh*-movement option not to be forced in matrix questions in SC, the interrogative complementizer must be phonologically null, as expected under the current analysis.

It is well known that even in some dialects of French C may be phonologically realized in questions. More precisely, some French dialects allow the complementizer *que* to appear in questions, as illustrated in (13).

(13) Qui que tu as vu?
whom C you have seen
'Who did you see?'

Under the LF C-insertion analysis we predict that in-situ questions will not be possible with the overt complementizer *que*. Since, owing to its phonological content, the complementizer must be introduced in the overt syntax, we would expect it to always trigger syntactic *wh*-movement. As (14) shows, the prediction is borne out. (14) is unacceptable for all speakers, including those who accept (13).

(14) *Que tu as vu qui?

The contrast between (13) and (14), as well as the fact that matrix overt-complementizer questions differ from matrix null-complementizer questions with respect to the availability of the in-situ strategy, strongly confirms the LF C-insertion analysis.

Two questions arise at this point: what is the status of (5b) in English, and why are both the in-situ and *wh*-movement options allowed with matrix questions in French? Consider first the counterpart of (5b) in English.

(15) (*)You saw who?

The status of this construction is not quite clear. (15) does not seem to be as good as French (5b) on the true question (i.e., nonecho) interpretation. However, there still seems to be a contrast between (15) and embedded questions such as (16).

(16) *John wonders you saw who.

If (15) is acceptable on the true question reading, its grammaticality, as well as the contrast between (15) and (16), can be accounted for in the same way as (5b) and the contrast between (5b) and (6b). On the other hand, if (16) is unacceptable on the true question reading, we can account for its ungrammaticality by assuming that the interrogative complementizer in English constructions such as (15) is lexically specified as a phonological affix that must be affixed to a verbal head at PF, a rather natural move given *do*-support and subject-aux(iliary) inversion in English matrix questions. 7 *Do*-support and subject-aux inversion can be

assumed to take place in order to provide an appropriate host for the affixal complementizer. The derivation in which the interrogative complementizer is introduced in the LF representation of (15) is then straight-forwardly ruled out because the requirement that the complementizer be affixed to a verbal head in PF cannot be satisfied. (The derivation crashes because LF cannot interpret the phonological requirement on the interrogative complementizer.)

Let us now turn to embedded questions. Notice first that, for some reason that remains unclear, the interrogative complementizer in embedded questions in English does not seem to be a PF affix, given that neither *do*-support nor subject-aux inversion takes place in English embedded questions. (Compare *What did John buy?* and *What can John buy?* with *I wonder what John bought and I wonder what John can buy*.) 8 The derivation in which the interrogative complementizer in (16) is introduced into the structure in LF, which would delay movement of *what* to the embedded CP projection until LF, then cannot be ruled out by appealing to PF requirements on the complementizer. However, as discussed above with respect to French (6b), the derivation in question is still ruled out because it involves an instance of Merge (i.e., lexical insertion) that does not expand the tree. Since Merge can take place only at the root of a tree, a complementizer with strong [+wh] feature (I assume that the [+wh] feature is strong in English) must be introduced in the overt syntax before the higher structure is built. *Wh*-movement then must take place overtly in the construction in question to check the strong [+wh] feature of the complementizer upon insertion. We thus derive *John wonders who you saw* instead of (16). Notice that we can still account for the fact that (16) is worse than (15) on this analysis because the constructions are ruled out in different ways: the operation of Merge in an embedded position and the failure to check a strong feature immediately upon insertion should be taken to result in stronger unacceptability than the failure of a PF affix to be affixed to an appropriate host at PF because it was inserted at LF.

Now let us consider why French allows (5a) in addition to (5b) in spite of Procrastinate, which requires that all operations take place as late as possible. We can account for the existence of both (5a) and (5b) in French if French has both a PF affix and a nonaffix interrogative complementizer, the former being used in (5a) and the latter being used in (5b). The fact that, in contrast to what happens in (5a), inversion does not and cannot take place in (5b) (see (17)) may be relevant here. (Note that the ungrammaticality of (17) is fully consistent with the IP analysis of French

in-situ questions.) However, the grammaticality of (18), where nothing seems to "support" the interrogative complementizer, casts doubt on this analysis.

(17) *As-tu vu qui?

(18) Qui tu as vu?
whom you have seen
'Who did you see?'

The most promising way of accounting for the grammaticality of both the in-situ and the *wh*-movement strategy in French matrix questions seems to be to exempt Merge from Procrastinate. The phonologically null complementizer could then be merged either overtly or covertly in French, giving either (5a) or (5b). Having Merge apply freely either overtly or covertly in spite of Procrastinate is actually a rather natural move in Chomsky's (1995) system. Procrastinate was originally posited in order to delay movement until LF whenever this is possible (see Chomsky and Lasnik 1993, 535). In Chomsky's (1995) system, however, this can be achieved without Procrastinate. Chomsky observes that a natural consequence of the standard minimalist assumption that movement is driven by feature checking is that, all else being equal, the operation Move should apply to features and not to syntactic categories. Overt movement, which feeds PF, still has to apply to whole categories, given the natural assumption that lexical items with scattered features cannot be interpreted/pronounced at PF. Since the considerations of PF interpretability are not relevant to LF, in LF the operation Move applies only to features. As observed in Chomsky 1995 and attributed to Hisatsugu Kitahara and Howard Lasnik, in this system LF movement is in a sense always cheaper than overt movement since it carries less material: whereas overt movement affects whole categories, covert movement affects only features. Since covert movement is more economical than overt movement, it should be preferred to overt movement. Given feature movement, there is no longer any need for an independent principle such as Procrastinate to ensure that movement is delayed until LF whenever possible. Since this was the main reason for positing Procrastinate in the first place, it seems that Procrastinate can be eliminated. And if Procrastinate is eliminated, Merge should be free to apply either overtly or covertly. Introduction of feature movement into the theory does not affect Merge; that is, it does not appear to make LF Merge in any sense cheaper than overt Merge. There is only one scenario under which it actually does. Suppose

that a merged element triggers movement. The movement itself could be cheaper if the element in question is merged in LF rather than in the overt syntax, since it would be a case of feature movement. Notice, however, that we need global economy of derivation to make LF Merge cheaper in the case in question. This cannot be achieved with local economy of derivation, which is clearly conceptually more appealing than global economy of derivation (see Chomsky 1995 and Collins 1997 for much relevant discussion), because local economy of derivation does not allow look-ahead. If the insertion itself is not cheaper at point A than at point B, by local economy of derivation it should be allowed to apply at either point A or point B. Merger of the interrogative complementizer in the constructions under consideration then should be able to apply either covertly or overtly. As a result, we account for the fact that both (5a) and (5b) are acceptable. In (5a) the interrogative complementizer is merged into the structure overtly, triggering overt *wh*-movement, and in (5b) it is merged covertly. As discussed above, covert applications of Merge are still restricted, since only phonologically null elements can be introduced into the tree in LF. And both covert and overt Merge are constrained by the requirement that Merge must expand the structure (i.e., apply only at the root of the tree). 9

To summarize, we have seen that for the "no overt *wh*-movement" option to be available in French, the interrogative complementizer must be phonologically null and located at the root of the tree. The analysis whereby the complementizer is inserted at LF provides a natural way of capturing these two conditions. As discussed above, only phonologically null elements can be inserted covertly, and lexical insertion can take place only at the root of the tree. The behavior of French in *wh*-in-situ questions thus follows straightforwardly given the way lexical insertion takes place in the minimalist system and given Chomsky's (1995) conception of strong features. The fact that under the current analysis the strength of the [+wh] feature in French can be kept constant (i.e., there is no need to assume that the [+wh] feature of the French complementizer is sometimes weak and sometimes strong to account for the availability of more than one option with respect to *wh*-movement in French) should be considered an advantage of the analysis. Assuming that the [+wh] feature can be either strong or weak in French would amount to sneaking in optional movement, which is by its very nature incompatible with the minimalist system. The optionally strong/weak [+wh] feature analysis is also empirically inferior to the current analysis since it fails to account for the fact

that the in-situ strategy is not always allowed in French. Chomsky (1995) assumes that the [+wh] feature is interpretable and therefore does not have to be checked when it is weak (see the discussion in section 2.3). It is not clear, then, how the ungrammaticality of (6b) can be accounted for if the [+wh] feature can be weak in French. The same holds for other contexts discussed below in which *wh*-movement is obligatory in French (see (19a), (20a), and (21a)). As shown below, the data in question receive a straightforward account under the analysis involving a strong [+wh] feature plus LF C-insertion. To the extent that it is successful, the analysis provides evidence that lexical insertion, in fact even the insertion of elements with strong features, can take place in LF in certain well-defined configurations. 10 This in turn provides evidence for Chomsky's (1995) approach to strong features, which leaves room for covert lexical insertion of elements with strong features. We have also seen that if we adopt the local rather than the global view of economy of derivation, we can account for the fact that both the in-situ and the *wh*-movement strategies are available in French root questions, which in turn provides evidence that the local view of economy of derivation is superior to the global view on empirical grounds. Turning to English, we have seen that, if real, the differing behavior of French and English with respect to the availability of the in-situ strategy in matrix questions can be accounted for given the plausible assumption that the interrogative complementizer is a phonological affix in English. In this respect, it is worth noting that the current analysis provides a uniform account of the different behavior of English and French with respect to the obligatoriness of subject-aux inversion and the possibility of *wh*-in-situ in matrix questions.

I should note here that Chomsky (1995) explicitly bans covert insertion of elements with strong features, an operation that plays a crucial role in the current analysis. However, Chomsky's empirical reasons for imposing this ban are very weak, and the way he implements it is conceptually very problematic. Chomsky appears to have two empirical reasons for banning LF insertion of elements with strong features: to account for the ungrammaticality of English (15), and to account for what he calls the general unacceptability of in-situ questions with adjuncts even in French-type languages. However, we have seen that the ungrammaticality of (15), as well as the contrast between (15) and (5b), can be accounted for even if we allow covert insertion of elements with strong features. As for in-situ adjunct questions, they are clearly allowed in French, as illustrated in (9). Note also that Chomsky does not discuss the fact that the in-situ strategy

is not always available in French even with arguments, as illustrated by the contrasts between (5b) and (6b), and between (13) and (14) (see also section 2.2). It is difficult to see how these contrasts can be accounted for in a principled way without invoking LF insertion of phonologically null elements with strong features, which is certainly a possibility in the theory (i.e., all the mechanisms needed in the LF C-insertion analysis are expected to be available given the basic minimalist assumptions in Chomsky 1995). 11

2.2
Long-Distance and Negated Questions in French

In the previous section I have discussed the well-known difference between matrix and embedded null-complementizer questions in French with respect to the availability of *wh*-in-situ. The *wh*-in-situ strategy, however, is not always available in French matrix null-complementizer questions. Thus, my informants find long-distance questions in which a finite-clause boundary intervenes between a [+wh] [Spec, CP] and a *wh*-in-situ fully acceptable only on the echo question reading. *Wh*-movement is obligatory in the contexts in question.12

(19) a. ?*Jean et Pierre croient que Marie a vu qui?
Jean and Pierre believe that Marie has seen whom
'Whom do Jean and Pierre believe that Marie saw?'
b. Qui Jean et Pierre croient-ils que Marie a vu?

(20) a. *Jean et Marie se demandent si Pierre aime qui?
Jean and Marie wonder if Pierre loves whom
'Whom do Jean and Marie wonder if Pierre loves?'
b. ??Qui Jean et Marie se demandent-ils si Pierre aime?

A question arises now why, unlike in short-distance questions, the in-situ strategy is not available in French long-distance questions.13 The question again receives a principled answer in the minimalist system. Furthermore, the answer turns out to have interesting theoretical consequences.

Consider (19a). ((20a) can be accounted for in the same way.) We need to rule out the derivation in which the complementizer with a strong [+wh] feature is inserted at the root of the tree in the LF representation of (19a), with the *wh*-phrase undergoing movement to check the strong [+wh] feature of the complementizer immediately upon the insertion. If this derivation were to yield a legitimate output, we would not be able

to account for the ungrammaticality of (19a) on the true (nonecho) question reading of the *wh*-phrase. If the movement of *qui*, motivated by checking the strong [+wh] feature of the complementizer, were overt, it would involve substitution into [Spec, CP]. Note, however, that since the complementizer is inserted in the covert syntax in the derivation under consideration, the movement in question will also have to be covert. Recall now that all LF movement necessarily involves pure feature movement. Chomsky (1995) argues that in LF formal features moveor, to be more precise, adjointo the heads bearing matching features. LF movement thus necessarily involves head movement, that is, adjunction to X0 elements. Instead of moving to [Spec, CP], *qui* will then have to undergo head movement by adjoining to the matrix complementizer in the LF representation of (19a) in the derivation under consideration. It is well known, however, that movement to X0 positions is subject to very strict locality restrictions. As far as I know, no grammatical instance of head movement of the "necessary" length has ever been reported. I therefore suggest that the derivation in question is ruled out because it involves movement to an X0 position that violates locality restrictions on head movement.

Notice that the movement of *qui* to the matrix complementizer in (19a) violates locality restrictions on head movement even if we adopt Roberts's (1992) (see also Rivero 1991) relaxed version of the Head Movement Constraint, which subjects head movement to Relativized Minimality and by doing so allows movement to an X0 position of type a to cross X0 positions of type b but not type a. To reach the matrix complementizer, $\bar{A}$-head position, in the derivation in question, *qui* must skip another $\bar{A}$-head, namely, the embedded-clause complementizer, which violates even Roberts's relaxed Relativized Minimality version of the Head Movement Constraint. On the other hand, in short-distance questions such as (5b), LF movement of the *wh*-phrase to the interrogative complementizer does not cross any $\bar{A}$-heads, so that Relativized Minimality is respected. In other words, it represents a grammatical instance of movement to X0 positions under Roberts's approach to locality restrictions on such movement. This approach thus enables us to account for the contrast between (5b) and (19a).

It also enables us to account for the degraded status of negative *wh*-in-situ constructions such as (21a) on the true (nonecho) question reading of the *wh*-phrase (cf. the grammatical (21b)).

(21) a. ?*Jean ne mange pas quoi?
Jean NEG eats NEG what
'What doesn't Jean eat?'
b. Que ne mange-t-il pas?

In order to derive a legitimate output for (21a) on the true question, nonecho reading of the *wh*-phrase in situ, the complementizer must be inserted in LF and the *wh*-phrase must move to the complementizer to check its strong [+wh] feature immediately upon insertion of the complementizer. However, the movement of the *wh*-phrase violates Roberts's version of the Head Movement Constraint. Like the movement of the *wh*-phrase to the matrix complementizer in (19a), movement of *quoi* to the complementizer, an Ā-head, crosses another Ā-head, this time negation, and is expected to yield an ungrammatical output under this approach to movement to X0 positions.

Adopting this approach to locality conditions on movement to X0 positions thus enables us to account not only for the ungrammaticality of (19a), but also for the ungrammaticality of (21a) and the contrast between (5b) and (19a)/(21a). Since this analysis is crucially based on Chomsky's feature movement hypothesis, which forces all LF movement to be adjunction to X0 positions, to the extent that it is successful, the analysis provides evidence for the feature movement hypothesis. 14 The analysis also provides evidence that feature movement is subject to locality restrictions on movement.15 In fact, feature movement seems to be subject to essentially the same locality restrictions as overt movement to X0 positions, which confirms Chomsky's intuition that feature movement in some sense involves head movement. It remains to be seen, however, how the Relativized Minimality version of the Head Movement Constraint can be integrated into Chomsky's (1995) system and what its consequences for the system will be. The needed version of the Head Movement Constraint does not readily fit into a system based on Attract instead of Move. However, the same holds for certain other locality constraints on movement, for example, the traditional Condition on Extraction Domain (CED) phenomena of Huang (1982)in fact, almost all of them except the *Wh*-Island Constraint. There thus still seems to be a need for a conception of the Shortest Movement Condition that considers movement from the point of view of the moved element. In contrast to Attract, such a conception of the Shortest Movement Condition can readily accommodate the Relativized Minimality version of the Head

Movement Constraint. As shown in Takahashi 1994, it can also accommodate the traditional CED phenomena and the Coordinate Structure Constraint, which remain mysterious under Attract. 16

2.3
Wh-Phrases That Do Not Move in LF

Now consider multiple questions in which a clausal boundary or negation intervenes between a *wh*-in-situ and a [+wh] C whose specifier is overtly filled by another *wh*-phrase. (22a) and (22b) are better than (19a) and (21a) on the true question, nonecho reading of the *wh*-phrases in situ.

(22) a. Qui croit que Marie a vu qui?
who believes that Marie has seen whom
'Who believes that Marie saw whom?'
b. Qui ne mange pas quoi?
who NEG eats neg what
'Who does not eat what?'

The grammaticality of these constructions on this reading indicates that the *wh*-phrases in situ are not undergoing LF movement to C. If they were, we would expect (22a,b) to have the same status as (19a) and (21a), which is not the case. This is expected in the minimalist system. In this system, movement to a [+wh] [Spec, CP], or adjunction to a [+wh] C, takes place to check the strong [+wh] feature of C. In (22a,b) this is done by the *wh*-phrases that are located in [Spec, CP] at S-Structure. There is therefore no need for the *wh*-phrases that are located in situ in the S-Structure representation of (22a,b) to undergo LF movement to C, in contrast to the *wh*-phrases in (19a) and (21a). (Note that in Chomsky's (1995) system, the [+wh] feature of the *wh*-phrases is considered to be interpretable and weak and therefore does not have to be checked.) Last Resort, which bans superfluous steps, then prevents them from undergoing LF movement to C.17

It seems plausible, however, that a *wh*-phrase in situ must be somehow associated with a [+wh] C, possibly as a way of scope taking, in order to ensure proper interpretation (see Baker 1970, Chomsky 1973, Reinhart 1995, and Williams 1986, among many others). In (22a,b) this can be done through the mechanism of unselective binding, with the interrogative complementizer binding the *wh*-phrases in situ. Notice that this cannot be done in (19a) and (21a), since the strong [+wh] feature of the complementizer remains unchecked if the *wh*-phrases in situ are associated

with the interrogative complementizer through unselective binding. To check the strong [+wh] feature, the *wh*-phrases must enter into a checking relation with the complementizer, which is done through LF adjunction to the complementizer, after the complementizer is inserted into the structure.

A question that arises now is why (23a,b), with an in-situ adjunct, are ungrammatical.

(23) a. *Qui a réparé la voiture comment?
b. *Who fixed the car how?

Under the standard preminimalist analysis, (23a,b) are ruled out by the Empty Category Principle (ECP). Since all *wh*-phrases are assumed to be located in [Spec, CP] in LF, the *wh*-phrases in situ in (23a,b) must undergo LF movement to [Spec, CP]. The constructions are then assumed to be ruled out because the trace left by the LF *wh*-movement of the *wh*-adjuncts is not properly governed. This account is clearly not available in the minimalist framework. Furthermore, it is empirically inadequate.

Rudin (1988) shows convincingly that in Bulgarian constructions such as (24), both *wh*-phrases are located in [Spec, CP] with the adjunct *wh*-phrase being adjoined to [Spec, CP], since the subject *wh*-phrase moves to [Spec, CP] before the adjunct to satisfy the Superiority Condition. (The condition requires that given two *wh*-phrases in situ, the higher one must move to [Spec, CP]. In the Bulgarian case the higher *wh*-phrase moves first.) 18

(24) [CP[[Koj] kak] [C' e kupil kniga]]?
who how is bought book
'Who bought the book how?'

The S-Structure representation of (24) is thus the same as the LF representation of (23) in the preminimalist framework. The grammaticality of (24), then, provides evidence that an element adjoined to [Spec. CP] can license its trace, contrary to the standard assumption that antecedent government from the [Spec, CP]-adjoined position is not possible.19

Furthermore, Haider (1986) and Müller and Sternefeld (1996) show that in German a *wh*-adverb can remain in situ overtly even in constructions in which the specifier of the interrogative complementizer in which the adverb is interpreted is filled by another *wh*-phrase. ((25a) is from Haider 1986, and (25b) from Müller and Sternefeld 1996.)

(25) a. Wer ist weshalb weggegangen?
who is why away-walked
'Who left why?'
b. Wer ist warum gekommen?
who is why came
'Who came why?'

As Haider points out, the standard ECP account of (23) incorrectly rules out (25).

Finally, notice that if the adjuncts in (23a,b) were to undergo LF *wh*-movement, under the feature movement theory the movement would be feature adjunction to the interrogative complementizer and not movement to [Spec, CP]. It seems that this would be a perfectly legitimate instance of feature movement. In fact, the movement would not differ in any respect from the movement the adjunct must undergo in the grammatical (9) to check the strong [+wh] feature of the complementizer, which is inserted in LF. Clearly, we need a new account of the ungrammaticality of (23a,b).

Recall that I have assumed that a *wh*-phrase must be associated with an interrogative complementizer in order to ensure proper interpretation. The association can be established in two ways: either by moving the *wh*-phrase to a position within the projection of the complementizer, or by having the complementizer unselectively bind the *wh*-phrase in situ. Now, Tsai (1994) and Reinhart (1995) argue that *wh*-adjunctsor, to be more precise, *wh*-adverbscannot be unselectively bound owing to the lack of a variable (i.e., an open position). 20 As a result, Tsai and Reinhart argue, *wh*-adverbs cannot be interpreted in situ. If this is true, the only way for the *wh*-adverbs in (23) to be associated with the interrogative complementizer is for them to undergo LF movement to the interrogative complementizer.21 As noted above, this appears to be a perfectly well formed instance of feature movement that does not violate any locality conditions on movement or licensing of traces. Notice, however, that the movement has no morphological motivation. Neither the complementizer nor the *wh*-adverb contains any features that need to be checked prior to the movement. The structures that (23a,b) have prior to the movement of the adjuncts to the interrogative complementizer are fully legitimate LF representationsall features that need to be checked are checked. As a result, the movement of the adverbs to the interrogative complementizer violates Last Resort. The problem with this movement is that it is driven by what Chomsky (1993) calls "the search for intelligibility" rather than

the satisfaction of formal requirements (feature checking), which is the only legitimate driving force for movement according to Chomsky (1993, 1995). I conclude, therefore, that (23a,b) converge as gibberish without LF movement of the *wh*-adverbs, which would ensure proper interpretation but is banned by Last Resort.

Let us now see how we can account for the German examples (25a,b), which contrast with the French and English examples (23a,b). The grammaticality of the German constructions indicates that the *wh*-adverbs are succeeding in establishing a relation with the interrogative complementizer. We could account for this by assuming that German *wh*-adverbs differ from their French and English counterparts in that they can be unselectively bound by C. This seems unlikely, given that, as Tsai (1994) and Reinhart (1995) argue, the possibility of unselective binding depends on the availability of an open position (i.e., a variable), which exists in NPs but not in adverbs (in this regard, see also Higginbotham 1983, 1985). It then seems more appropriate to account for the grammaticality of (25a,b) by having the *wh*-adverbs undergo LF feature movement to the matrix complementizer. The driving force for the movement has to be independent of the [+wh] feature; that is, the feature driving it cannot be the [+wh] feature and should not be available in English and French. It is tempting to relate this movement to the availability of partial *wh*-movement in German, which indicates that German *wh*-phrases can move into a CP projection independently of the [+wh] feature, the relevant C not being specified as [+wh]. That relating the relevant feature to the one involved in partial *wh*-movement may be on the right track is indicated by the fact that partial *wh*-movement is not available in English and French. 22 It is worth noting in this context that Stjepanovic* (1995) and I (Boskovic* 1997b) show that SC *wh*-phrases can also be fronted independently of the [+wh] feature. (We also show that the movement is not an instance of scrambling.) In fact, in Boskovic 1998b I argue with respect to Bulgarian, a language that allows multiply filled interrogative [Spec, CP]s in overt syntax, that even movement of *wh*-phrases to an interrogative [Spec, CP] does not have to be driven by [+wh] feature checking; that is, it can have feature-checking motivation independent of the [+wh] feature. Interestingly, Müller and Sternefeld (1996) observe that, unlike in short-distance questions, German *wh*-adverbs cannot be left in situ in long-distance questions. This is exactly what is expected if these adverbs are undergoing LF feature movement to C. The ungrammaticality of (26) then can be accounted for in the same way as the ungrammaticality of (19a).

(26) a. * Wer hat gesagt daß Fritz warum ein Buch gelesen hat?
who has said that Fritz why a book read has
'Who said that Fritz has read a book why?'
b. *Wen hast du empfohlen daß man wie bestrafen soll?
whom have you recommended that one how punish should
'Who did you recommend that one should punish how?'

Whatever the precise identity of the feature is that is responsible for the LF movement of *wh*-adverbs in German, it cannot be obligatorily present in German *wh*-phrases or the interrogative complementizer. 23 This is indicated by the fact, observed by Müller and Sternefeld (1996), that argument *wh*-phrases can be left in situ in long-distance questions in German, which in turn indicates that not all German *wh*-phrases in situ are forced to move to C in LF.

(27) Wer hat gesagt daß Fritz was lesen soll?
who has said that Fritz what read should
'Who said that Fritz should read what?'

In contrast to *wh*-adverbs, the *wh*-phrase in (27) can be unselectively bound and therefore does not depend on undergoing movement to the interrogative complementizer in order to be associated with it.

A question that arises now is how to account for pure *wh*-in-situ languages such as Japanese. Consider (28).

(28) a. Anata-ga dare-o mita ka?
you-NOM who-ACC saw Q
'Who did you see?'
b. Peter-wa [anata-ga dare-o mita ka] tazuneta?
Peter-TOP you-NOM who-ACC saw Q asked
'Peter asked who you saw.'

Notice first that (28a) should not receive the same analysis as French (5b); that is, it should not be analyzed as involving LF insertion of a complementizer with a strong [+wh] feature since, unlike in French, embedded questions with *wh*-phrases in situ are allowed in Japanese. If Japanese were to be given the same analysis as French, we would incorrectly predict (28b) to be ungrammatical. Notice also that the interrogative complementizer in Japanese (28) is lexically realized. As a result, in contrast to the interrogative complementizer in French (5b), the interrogative complementizer in Japanese (28) clearly cannot be introduced into the structure in LF.

It is well known that long-distance in-situ questions are also good in Japanese. Thus, both (29a) and (29b) are acceptable on the matrix true question, nonecho reading of the *wh*-phrase, and both contrast with French (19a) and (20a). Negated questions are also acceptable on the nonecho reading, again in contrast with French (cf. (21a)).

(29) a. John to Mary-wa [Peter-ga dare-o mita to]
John and Mary-TOP Peter-NOM who-ACC saw that
sinziteiru ka?
believe Q
'Who do John and Mary believe that Peter saw?'
b. (?)John to Mary-wa [Peter-ga dare-o mita kadooka]
John and Mary-TOP Peter-NOM who-ACC saw whether
siritagatteiru ka?
want-to-know Q
'Who do John and Mary want to know whether Peter saw?'
c. John-wa nani-o tabenakat-ta ka?
John-TOP what-ACC ate-NEG Q
'What didn't John eat?'

The data in (28) (29) indicate that, unlike French *wh*-phrases, Japanese *wh*-phrases do not undergo LF feature movement. If they did, it would be difficult to account for the contrasts between Japanese and French noted above. However, it is well known that Japanese *wh*-phrases are not completely insensitive to constraints on movement. Thus, an adjunct *wh*-phrase cannot be separated from the interrogative complementizer by an island, as illustrated in (30). According to Watanabe (1992), some degradation is found even in argument questions such as (29b), though this claim is somewhat controversial.

(30) *John to Mary-wa [Peter-ga naze kuruma-o naosita kadooka]
John and Mary-TOP Peter-NOM why car-ACC fixed whether
siritagatteiru ka?
want-to-know Q
'Why do John and Mary want to know whether Peter fixed the car?'

Given (30), some kind of movement must be taking place in Japanese questions. As noted above, the movement cannot be covert; if it were, we would not be able to account for the contrast between (29) and (19a), (20a), and (21a). The movement, then, must be overt. In other words, the

movement in question seems to behave like XP-movement, rather than X0-movement. *Wh*-phrases themselves are clearly not undergoing *wh*-movement in the constructions under consideration, since they are located in situ. Therefore, the only possibility seems to be that a null *wh*-operator is undergoing movement. This is exactly what is proposed by Watanabe (1992) and Aoun and Li (1993), who argue that the difference between languages such as English and languages such as Japanese with respect to what moves in *wh*-questions is a result of the shape of *wh*-phrases and their quantificational force in the languages in question. Under both Watanabe's and Aoun and Li's analyses, interrogative [Spec, CP]s are always overtly filled in languages such as Japanese. Under Watanabe's analysis the *wh*-operator always moves to [Spec, CP], whereas under Aoun and Li's analysis sometimes it is base-generated there and sometimes it moves, a dual behavior that is intended to account for island effects. Both analyses imply that the [+wh] feature in Japanese is strong and therefore requires the presence of an element with a [+wh] feature in [Spec, CP] overtly. In the current system we seem to be led to an analysis along the lines of Watanabe's and of Aoun and Li's. There are still many wrinkles left to iron out, though doing so lies well beyond the scope of this chapter. For relevant discussion, see Watanabe 1992 and Aoun and Li 1993. 24

2.4
Conclusion and Some Theoretical Consequences of the Proposed Analysis

In this chapter I have examined when *wh*-movement must take place in French and offered a minimalist account of the limited distribution of the in-situ strategy in French. I have argued that in French constructions involving *wh*-in-situ such as (5b), the interrogative CP projection is inserted only in LF. This is a necessary condition for the in-situ strategy to be available in French. *Wh*-movement then does not take place overtly in (5b) for a trivial reason: no CP projection is present in the S-Structure representation of (5b). Where the in-situ option is not allowed, either LF insertion of the interrogative CP projection is blocked or the insertion leads to a violation of locality constraints on movement. To be more precise, with embedded questions LF insertion of the interrogative complementizer is blocked because it violates the requirement that Merge expand the structure, and with long-distance and negative questions the insertion leads to a violation of locality constraints on movement. The LF

C-insertion analysis has interesting theoretical consequences. It provides evidence that lexical insertion, in fact even lexical insertion of elements with strong features, can take place in LF under certain well-defined conditions, namely, when the insertion takes place at the root of the tree and when the element undergoing insertion is phonologically null. This in turn provides evidence that Chomsky's (1995) conception of strong features is superior to earlier conceptions (e.g., the one in Chomsky 1994), since it allows LF insertion of elements with strong features.

The current analysis also provides evidence that feature movement is subject to at least some locality constraints on movement. In fact, it seems to be subject to the same constraints as overt movement to X0 positions, which confirms Chomsky's claim that feature movement involves X0-adjunction. Note that the analysis presented here is crucially based on the assumption that all LF movement involves adjunction to X0 elements. (LF *wh*-movement then involves adjunction to C0, rather than movement to [Spec, CP].) As a result, to the extent that it is successful, the analysis presented here also provides evidence for Chomsky's feature movement hypothesis.

Under the current analysis we would expect to regularly find *wh*-elements in situ in matrix clauses in languages with a strong [+wh] feature as long as the interrogative complementizer is phonologically null and is not lexically specified as a phonological affix, a reflex of which could be subject-aux inversion or a *do*-support type process. (A phonological affix C must be PF-adjacent to its host.) However, subject-aux inversion is crosslinguistically a very common phenomenon in questions. This is not surprising, given Pesetsky's (1992) conjecture that null morphemes are by and large specified as phonological affixes. Another interfering factor is that the presence of a fronted *wh*-phrase does not necessarily indicate that *wh*-movement has taken place. For example, as discussed in Boskovic * 1997b and illustrated briefly above, SC *wh*-phrases, including *wh*-phrases with echo interpretation, are forced to undergo fronting quite independently of the [+wh] feature. (Stjepanovic* (1995) argues that the fronting is motivated by a focusing requirement on *wh*-phrases in this language.) Fronting of a *wh*-phrase in SC thus does not necessarily provide evidence that *wh*-movement is taking place. In fact, as discussed above, even though it displays obligatory fronting of *wh*-phrases, SC turns out to behave like French with respect to when *wh*-movement must occur. Apparently, care should be taken in diagnosing *wh*-movement; fronting of *wh*-phrases

cannot be taken as the sole diagnostic of the process. The same point is made in Cheng 1991 with respect to several languages.

One potential empirical consequence of the analysis developed above remains to be discussed. We have seen that nothing in the minimalist system prevents a phonologically null complementizer with a strong [+wh] feature from being inserted in LF at the top of the tree in shortdistance root questions. Now we must ask whether other phonologically null elements can also be inserted in LF if they are inserted at the top of the tree. If so, the possibility arises that matrix clauses can be bare VPs in certain circumstances in overt syntax, the rest of the structure being built in covert syntax. First consider this question with respect to English constructions such as (31).

(31) John wanted a house.

Can the category I, and with it the IP projection, be inserted in the LF representation of (31), as a result of which the construction would be a bare VP in the overt syntax? 25 It appears that this would be possible if the I of (31) is phonologically null and has no PF requirementsfor example, it is not specified as a PF affix. However, Halle and Marantz (1993), Bobaljik (1994), and Lasnik (1995) have revived Chomsky's (1957) analysis of finite clauses by arguing that English finite I is a PF affix. In fact, according to these authors, I in (31) is not simply specified as a phonological affix; it even has phonological content. They argue that *ed* is actually located under I and undergoes affix hopping to the main verbor, in more recent terms, undergoes morphological merger (which should not be confused with Merge) with the main verb in PF under adjacency. If this is indeed the case, I, and with it the IP projection, would have to be present in the overt syntax in (31).

Next consider the French construction in (32).

(32) Jean voulait une maison.
Jean wanted a house

Lasnik (1995) argues that, in contrast to English finite I, French finite I is neither phonologically realized nor specified as a phonological affix. It is simply a bundle of features checked against the features of the verb after the verb raises to it. (In other words, whereas in English *want* but not *wanted* is present in the lexicon, *ed* being a phonological realization of I, in French *voulait* as a whole is present in the lexicon.) If I in French (32) is indeed phonologically null, the question arises whether it could be

inserted into the structure in LF, which in turn opens up the possibility that the construction could be a bare VP at S-Structure. 26 It is very difficult to tell empirically whether this is a possibility since, as discussed above with respect to the phonologically null complementizer, the derivation in which I is inserted in the overt syntax is also available in (32). Allowing LF insertion of I could not then rule out any constructions, though it might rule in some constructions that were ungrammatical on the insertion-of-IP-in-overt-syntax derivation. (Recall that this was the case with allowing the possibility of inserting the interrogative complementizer in LF, which did rule in some constructions that were ungrammatical on the insertion-of-CP-in-overt-syntax derivation.) At this point I have no empirical evidence that could conclusively settle this issue. What makes our job particularly difficult here is that, to ensure that I can be inserted in LF, it is not enough that I is phonologically null and has no PF requirements of its own. We need to ensure that no element present in the structure is required to move to I, or an element within the IP projection, overtly. This probably makes the ungrammaticality of negative sentences such as (33a,b) irrelevant here, since it seems plausible that French negation is required to move to I overtly (see, e.g., Acquaviva 1994, Belletti 1990, Haegeman 1992, Moritz 1989, Pollock 1989).

(33) a. *(Ne) pas Jean aime Marie.
NEG NEG Jean loves Marie
b. *(N')aime pas Jean Marie.

If the negative marker *ne* indeed must end up in I overtly, the movement should probably be driven by some inadequacy of negation, since non-negative sentences where negation is not present in I are fine. (I assume that in negative constructions where *ne* is missing, *ne* is deleted in PF.) This means that the presence of negation requires overt insertion of I, which in turn triggers subject and verb-movement to [Spec, IP] and I, respectively.

There is a similar interfering factor in (34), another potentially relevant construction.

(34) *Complètement, Jean oubliera cette histoire.
completely Jean will-forget this story

(34) contains what is traditionally called a VP adverb. (See, however, Belletti 1990, where it is suggested that *complètement* can be either a VP or a TP adverb.) If *complètement* is indeed a VP adverb, the question

arises of why (34) is illicit on the derivation in which I is inserted covertly. Kayne (1994) argues that the distinction between specifiers and adjuncts should be abolished. As noted by Ian Roberts (personal communication), in Kayne's system the adverb in (34) has to occupy the specifier position of some functional head, which is phonologically null. Suppose now that the phonologically null head, which is clearly phonologically weak, is a clitic and that French clitics in general must move to I (as suggested in Belletti 1990 and Kayne 1991, among many others), neither of which is an unreasonable assumption. Given these assumptions, (34) would be ruled out in the same way as (33). The presence of the functional head that the adverb introduces would require overt insertion of I, which in turn would require overt subject and verb-movement to [Spec, IP] and V, respectively. (Note that, as shown in Belletti 1990, *complètement* cannot undergo topicalization, which rules out the option of analyzing (34) by saying that the subject is in [Spec, IP], V is in I, and *complètement* undergoes topicalization.) 27 The ungrammaticality of (34) thus cannot help us draw a definite conclusion concerning the possibility of covert I-insertion. Covert I-insertion may actually be quite generally banned for independent reasons. Consider again the derivation in which I is inserted overtly in (32), with the verb moving to adjoin to I and the subject moving to [Spec, IP]. If these movements are driven by strong features of I, it appears that I could still be inserted covertly in (32). However, if these movements are driven by some inadequacy of the moved elements, which would require them to be located in the I-adjoined position and [Spec, IP], respectively, then the option of inserting I and the IP projection in LF would be independently ruled out. The option would also be ruled out if the Extended Projection Principle (EPP) effect is a result of some general requirement that every sentence must contain a filled [Spec, IP]. It is difficult to find empirical evidence concerning what drives the movements in question. Chomsky (1995) assumes without real empirical evidence that the movements are driven by strong features of I, in which case it seems that there are no obstacles to LF insertion of I in (32). However, elsewhere (class lectures, MIT, 1995) he suggests that the EPP is a result of some universal thematization requirement, which crosslinguistically requires overtly filled [Spec, IP]s. The requirement could be interpreted in a way that would completely rule out the option of inserting I in LF and with it the option of inserting covertly any projection lower than I, given that Merge must expand the structure.28

Where does this leave us? We seem to have good empirical and conceptual evidence (the theory allows it) that under certain circumstances C and the CP projection can be inserted into the structure covertly, though the covert insertion option does not rule out the overt insertion option. As for I and the IP projection, the situation is unclear. At this moment I have no conclusive empirical evidence one way or the other, and it is unclear whether the option is allowed theoretically. In other words, firm conceptual evidence of the relevant sort (this is allowed by the theory) is still not available, mainly because it is unclear what is responsible for movements to positions within the IP projection. What is clear is that the covert lexical insertion option is theoretically very limited: it can apply only at the root of the tree, to a phonologically null element that furthermore has no PF requirements and does not have to be present overtly to provide a landing site for another element that must move overtly. The fact that we have been able to find one instance of LF insertion that exactly fits these conditions confirms the theoretical framework adopted in this chapter.

Notes

I wish to deeply thank Howard Lasnik for invaluable comments and suggestions on the whole of my linguistic work, including this chapter. It is a pleasure to be able to dedicate this chapter to Howard as a small token of appreciation for everything he has done for me and the field of linguistics.

For helpful comments and suggestions on this chapter, I also thank Roger Martin and audiences at seminars at the University of Connecticut, Formal Approaches to Slavic Linguistics 5 hosted by Indiana University and Wabash College, Kanda University, Nanzan University, and the Linguistic Symposium on Romance Linguistics 27 hosted by the University of California at Irvine, where portions of this material were presented. Particular thanks are due to Jun Abe, Masao Ochi, Mamoru Saito, Yuji Takano, Juan Uriagereka, and Akira Watanabe. Thanks are also due to a number of people, particularly Michèle Bacholle, Cédric Boeckx, Viviane Déprez, and Géraldine Legendre, for their help with judgments. Some material from Boskovic * 1997b has served as a starting point for sections 2.1 and 2.2, where this material is considerably expanded and revised.

1. The same holds for Chomsky 1993. There Chomsky considers strong features illegitimate PF objects, which does not in principle rule out LF insertion of elements with strong features.

2. Chomsky (1995) also considers the possibility that the null complementizer in constructions such as (5b) is inserted in LF, but ultimately rejects it for reasons to be discussed below.

3. *Pourquoi* 'why' is actually an exception in that it cannot be left in situ. Rizzi (1990), however, shows convincingly that the reason for this is trivial: *pourquoi* must be base-generated within the CP projection.

4. Aoun (1986) offers an account of the ungrammaticality of (6b) based on his proposal that selectional restrictions must be satisfied at S-Structure in French. The construction is then ruled out because *demander* 'ask' does not take a [+wh] complement at S-Structure. This account is not available in the current theoretical framework, which does not attach any syntactic significance to S-Structure. The fact that the current analysis does not invoke S-Structure to account for the ungrammaticality of (6b) can in fact be considered an advantage of the analysis. Under this analysis there is no need to check whether selectional restrictions are satisfied in French before LF. Note also that in order to account for the contrast between French (6b) and constructions such as Japanese (7), Aoun proposes that, in contrast to what happens in French, selectional restrictions in Japanese must be satisfied only at LF. The difference between French and Japanese concerning when selectional restrictions must be satisfied is inconsistent with the current view of crosslinguistic variation, where languages are considered to differ only in their morphological properties. (I return to pure *wh*-in-situ languages below.) As will become obvious, Aoun's analysis is also empirically inadequate. It leaves several facts concerning the availability of the in-situ strategy in French unaccounted for.

5. Note that SC indirect questions have the same form as matrix questions. As a result, special care must be taken to rule out the possibility of analyzing a SC indirect question as a superficial direct question, with the superficial matrix clause being analyzed as an adsentential. To rule out this possibility, no pause should be made before the embedded question. (Note that the matrix clause needs to be given more content, which means avoiding a matrix clause that contains only a verb, to make it less likely to be analyzed as an adsentential.) No danger of analyzing the embedded question as a superficial matrix question arises in SC correlative constructions (see (11c,d)) and existential constructions such as (11e,f), whose *wh*-clause has all the formal properties of questions (see Izvorski 1996, 1998). Note also that I ignore the irrelevant echo question reading of the *wh*-phrases in (11).

6. *Li* is used in regular yes-no questions. With *wh*-phrases, it is used in emphatic questions.

7. Pesetsky (1987) also assumes that the interrogative complementizer in English is a phonological affix. For relevant discussion see also Pesetsky 1992, where it is conjectured that null morphemes are by and large specified as phonological affixes.

8. An alternative that may be worth pursuing is to assume that the interrogative complementizer in embedded questions adjoins to the higher verb overtly, thus satisfying its affixal requirement. This option is not available in matrix questions such as (15) because it would involve overt lowering.

9. This seems to allow for the possibility of having a number of separate trees in the overt syntax, which are then assembled in LF. This possibility is put to use by Hoffman (1995), who argues that this is exactly what happens in free-word-order languages. However, it seems that the possibility should be ruled out in PF, since it is difficult to see how some PF operations (e.g., intonational phrasing) can apply properly unless the whole structure is assembled into one tree before it enters PF.

10. It is true that the notion of strength was originally posited to drive overt movement. However, the way it is implemented in Chomsky 1995 leaves room for strength to drive covert movement in certain well-defined configurations. I see no reason why this possibility should not be taken advantage of, especially when this move is empirically well motivated, as I attempt to show here. (See also the discussion in section 2.2.)

11. To rule out LF insertion of phonologically null elements with strong features, Chomsky has to adopt an additional assumption, given in (i).

(i) a enters the numeration only if it has an effect on output.

Chomsky argues that strength in an interrogative complementizer has neither an LF nor a PF effect if the complementizer is introduced into the structure covertly. As a result, given (i), strength cannot be present in the numeration, which is defined as an array of lexical items that is mapped by the computational system into a linguistic expression (a PF-LF pair), to start with. The condition in (i) is conceptually very problematic. In order to determine the effects of (i), we need to know PF and LF outputs. However, the numeration, which is determined by (i), must be present in the initial stage of the derivation. Thus, (i) raises serious issues of globality. (i) also has seemingly undesirable empirical consequences that are not explored in Chomsky 1995. For example, (i) appears to ban constructions containing I with a strong D-feature and null subjects (*pro*, PRO, null operator), since the movement of the null subject to [Spec, IP], motivated by the strength of I, has no effect on the outputs. That is, it seems that languages that have null subjects of any kind cannot have an I with a strong D-feature in the contexts in which these null subjects appear. This may be too strong.

The basic idea behind (i) is that superfluous projections are not allowed in representations. Several ways of implementing this idea have been proposed that do not face the problems that (i) faces and do not have the effect of altogether ruling out covert insertion of elements with strong features (see Boskovic * 1997c and references therein). In Boskovic 1997c I in fact show that with small modifications in the notion of numeration and in the way lexical insertion works, the ban on superfluous projections in representations can be made to follow from the independently needed principle of Last Resort, which bans superfluous steps in derivations. That is, I show that it is possible to use Last Resort to prevent creation of representations with superfluous projections, which makes any condition that is intended to specifically rule out such projections redundant and therefore eliminable.

12. My informants' judgments for (19a) range from fully unacceptable as a true (nonecho) question to clearly worse than (5b). Note that (19a) sounds best with stress on the *wh*-phrase in situ, an indication of the echo question reading. (A similar situation holds for (21a), discussed below, which has the same status as (19a).)

It is worth noting here that (19a) is not as bad as (6b) on the true question reading. This can be accounted for under the analysis presented below, since (6b) and (19a) violate different conditions.

Note that the direct object *wh*-phrase in (20a) has matrix scope on the relevant reading. The construction is clearly worse than a Subjacency violation on the matrix true question reading of the direct object *wh*-phrase.

13. In this chapter I concentrate on finite *wh*-in-situ complements. Infinitival *wh*-in-situ complements are discussed in Boskovic * 1998a.

14. It is worth noting here that Chomsky (class lectures, MIT, 1995) suggests that when X undergoes overt XP-movement to [Spec, YP], two movements are actually involved: first Move F adjoins formal features of X to Y for feature checking, and then the rest of X undergoes XP-movement to [Spec, YP] ("pied-piping"), followed by a repair strategy that makes X pronounceable. Under this analysis it is not possible to make LF movement more constrained than overt movement by appealing to Move F, which I attempt to do here on the basis of the data under consideration (for relevant data, see also section 2.3). Therefore, if the discussion here is on the right track, the two-movements analysis cannot be correct: the decision to "pied-pipe" must be made immediately so that only one *actual* movement takes place (XP moves to [Spec, YP]), as originally suggested by Chomsky (class lectures, MIT, 1994; see also Chomsky 1995). (In the latter work Chomsky is somewhat ambivalent on this issue. However, he crucially assumes throughout the chapter that the checking configuration is specifier-head for overt syntax and FF(adjoined to head)-head for covert syntax, which goes against the spirit of the two-movements analysis.)

15. Takahashi (1997) also concludes that feature movement is subject to locality restrictions on movement. He claims that elements from which feature movement takes place can block movement of their own features via Huang's (1982) Condition on Extraction Domain (CED)that is, if they are subjects or adjuncts. The data considered here, however, provide evidence against this claim. Recall, for example, that French (9) must involve feature movement of the *wh*-phrase to C, after C is inserted in LF. Since the *wh*-phrase is an adjunct, (9) must involve feature movement out of an adjunct. Its grammaticality, then, provides evidence that phrases do not block movement of their own features. The same conclusion is reached in Boskovic 1997c. (It is worth pointing out here that, as noted by Takahashi (1997), Takahashi's (1994) minimalist account of CED effects does not rule out feature movement ofor, to be more precise, out ofadjuncts and subjects.)

16. Actually, the Attract system fails to account even for the full range of *Wh*-Island effects, which are supposed to be its showcase. In this system the *Wh*-Island effect is captured by appealing to feature checking instead of the A/$\bar{A}$ distinction. Thus, (i) is ruled out because the matrix C, which needs to check its [+wh] feature, fails to attract the closest [+wh]-feature-bearing element (*where*).

(i) ??Which booki do you wonder wherej John put *ti**tj*?

This seems to leave (ii) unaccounted for.

(ii) ??(Peter thinks that) That booki you wonder wherej John put *ti**tj*.

It is not at all clear why the [+wh] feature should be relevant in attracting topics. A similar problem arises with respect to a number of other constructions, for

example, relativization out of *wh*-islands (e.g., ??*the book that you wonder where John put*) and *tough*-movement out of *wh*-islands (e.g., ??*This car is tough to ask Peter when to repair*). Chomsky's (1995) system, where feature checking is intended to do the job of the A/$\bar{A}$ distinction with respect to Relativized Minimality, thus fails to account for the full range of *Wh*-Island effects. Several other types of Relativized Minimality effects with $\bar{A}$-movement also appear to remain unaccounted for in this system (e.g., Rizzi's (1990) pseudo-opacity effects and inner island effects). It is my belief that a fully successful way of capturing the effects of the A/$\bar{A}$ distinction on Relativized Minimality type phenomena would also cover the cases discussed in the text.

17. Note that since the relevant *wh*-phrases do not undergo LF *wh*-movement, we can now dispense with the stipulation that Subjacency does not apply in LF, a position I adopt here.

18. The construction is unacceptable if the order of the *wh*-phrases is reversed, which, as argued in Rudin 1988 and Boskovic * 1997a, is a Superiority effect.

(i) *[CP[[Kak] koj] [C' e kupil kniga]]?
how who is bought book

Note that, as argued extensively in Rudin 1988 and Boskovic 1998a, adjunction to [Spec, CP] in Bulgarian proceeds to the right, so that the *wh*-phrase that comes first in the linear order is the one that moves first to [Spec, CP]. For an account of the full range of Superiority effects in Bulgarian based on the economy approach to Superiority, see Boskovic 1997a, 1998b.

19. The assumption was rather strange, given that c-command seems to hold between the adjoined element and its trace and given that at the same time it was apparently assumed that antecedent government is possible from the [Spec, IP]-adjoined position. This was necessary to account for expletive *there* constructions, where the associate of the expletive was assumed to undergo adjunction to the expletive in LF (see Chomsky 1991).

20. They show that the standard argument-adjunct asymmetry is actually a noun-adverb asymmetry (for relevant discussion, see also Huang 1982), resulting from the presence of an open position in NPs, but not adverbs. The contrast between **Who left how?* and *Who left what way?* is instructive here.

21. Tsai and Reinhart also argue that *wh*-adverbs must be located within the interrogative CP projection in order to be properly interpreted.

22. Admittedly, the account adopted here of the different behavior of German and English/French *wh*-adverbs with respect to the possibility of remaining in situ in questions is not very deep. However, neither are the alternative accounts I am aware of. Thus, Müller and Sternefeld (1996; see also the references therein) account for this fact by stipulating that German differs from English in that in German, but not in English, antecedent government is possible from the [Spec, CP]-adjoined position. Haider (1986), on the other hand, suggests that the trace left by LF *wh*-movement of adverbs is lexically governed in German, but not in English.

23. This is in line with the assumption that the same kind of feature, with a potential difference in strength (the situation is not clear here given that, as argued in Boskovic * 1997c, strength is not the only thing that can force overt *wh*-movement), is involved in partial *wh*-movement, which is not an obligatory operation.

24. A potential problem here is that the movement of the *wh*-operator could be feature movement in spite of taking place overtly. Since the *wh*-operator is phonologically null, there seems to be no need for category pied-piping to ensure proper PF interpretation. This is un undesirable result, since we clearly want the movement in question to at least have the option of being XP-movement, in order to avoid strict locality restrictions on X0-movement. There are two ways of resolving the potential problem. One is to say that the *wh*-operator is actually not phonologically null. The operator moves to [Spec, CP], after which it undergoes PF deletion, perhaps as a reflex of the traditional Doubly Filled Comp Filter (see Chomsky and Lasnik 1977). To be able to undergo PF deletion, the operator would have to enter PF as a whole, rather than "broken" into features. This is actually exactly what Takahashi (1997) proposes happens in English constructions such as *the reason that John said that Mary left*, which, according to Takahashi, who essentially follows Chomsky and Lasnik (1977), involves movement of a phonologically overt element to [Spec, CP], followed by PF deletion of the element. (Note that the movement in question clearly has properties of XP-movement.)

> An alternative analysis is available that seems more principled to me. Chomsky (1995) argues that when one formal feature (FF) undergoes movement, all FFs are affected by the movement. Movement minimally affects all FFs. Returning now to the *wh*-operator, it is not clear that the operator contains anything but FFs. In fact, [+wh] may be the only feature the operator has. FF (*wh*-operator) would then be the whole *wh*-operator, (i.e., FF would not be a part of the lexical item in question; rather, it would be the whole lexical item). Given the contextual definition of Xmax and Xmin that Chomsky adopts (for evidence for this approach to the Xmax/Xmin distinction, see Boskovic 1997c), FF(*wh*-operator) would also be a maximal projection (it is a lexical item that does not project any more) as well as a minimal projection (it is a lexical item that is not a projection at all). Chomsky suggests that such elements are able to undergo both XP-and X0-movement. This would enable FF(*wh*-operator) to undergo XP-movement, which is a less constrained option. (It is not clear whether the final landing site of the movement could still be the C-adjoined position, and not only [Spec, CP].) In fact, as Masao Ochi (personal communication) points out, the null operator could undergo XP-movement even if it were to move in LF (by definition, Xmax would be undergoing movement) so that, for our purposes, it is actually not necessary for the null operator movement to be overt.

25. As for C and the CP projection, it is not clear whether they would ever be present in (31). For arguments that there is no C or CP projection in *that*-less declarative clauses, see Boskovic 1992, 1996, 1997c.

26. Whether (32) would be a bare a VP if the IP projection could be inserted in LF is not clear, since the answer depends on whether there are any projections

between VP and IP. A number of proposals concerning the existence of such "intermediate" projections have been made in the literature. If there are such projections, and if any of them is specified as a phonological affix, it would have to be present overtly. Note that what I called the IP projection could be split into a number of functional heads, some of which could be PF affixes and would there-fore have to be present overtly. Recall also that Pesetsky (1992) argues that phonologically null heads are by and large crosslinguistically PF affixes, in which case any phonologically null head above the VP in (32) would likely be an affix.

27. There is another potentially interfering factor in (34). As noted by Belletti (1990) and illustrated in (ia), even sentential adverbs in French cannot precede a subject, though this should be a base-generated option for them even when the subject is located in [Spec, IP] (cf. English (ib)).

(i) a. ?*Probablement Jean vendra ces livres.
probably Jean will-sell these books
b. Probably, John will sell these books.

It is possible that adverb-subject-verb sequences are for some reason ruled out in French; that is, it is possible that whatever rules out (ia) rules out (34). This is plausible if the ungrammaticality of (ia) is not due to structural reasons. (See, however, Belletti 1990 for an attempt to explain the ungrammaticality of (ia) in structural terms.) Notice, however, that some adverbs can appear before the subject in French, as illustrated in (ii). (The adverb must be heavily stressed, though.)

(ii) Souvent Pierre se trompe.
often Pierre makes-mistakes

28. Chomsky (class lectures, MIT, 1995) leaves open what drives movement of V to I, a process that seems to fall completely outside his system.

References

Acquaviva, Paolo. 1994. The representation of operator-variable dependencies in sentential negation. *Studia Linguistica* 48, 91 132.

Aoun, Joseph. 1986. *Generalized binding*. Dordrecht: Foris.

Aoun, Joseph, and Yen-hui Audrey Li. 1993. *Wh*-elements in situ: Syntax or LF? *Linguistic Inquiry* 24, 199 238.

Baker, C.L. 1970. Note on the description of English questions: The role of an abstract question morpheme. *Foundations of Language* 6, 197 219.

Belletti, Adriana. 1990. *Generalized verb movement*. Turin: Rosenberg and Sellier.

Bobaljik, Jonathan. 1994. What does adjacency do? In Heidi Harley and Colin Phillips, eds., *The morphology-syntax connection*, 1 32. (MIT Working Papers in Linguistics 22.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Boskovic, Zeljko. * 1992. Clausal selection, subjacency, and minimality. Ms., University of Connecticut, Storrs.

Boskovic, Zeljko. * 1996. Selection and the categorial status of infinitival complements. *Natural Language & Linguistic Theory* 14, 269 304.

Boskovic, Zeljko. 1997a. On certain violations of the Superiority Condition, AgrO, and economy of derivation. *Journal of Linguistics* 33, 227 254.

Boskovic, Zeljko. 1997b. Superiority effects with multiple *wh*-fronting in Serbo-Croatian. *Lingua* 102, 1 20.

Boskovic, Zeljko. 1997c. *The syntax of nonfinite complementation: An economy approach*. Cambridge, Mass.: MIT Press.

Boskovic, Zeljko. 1998a. LF movement and the Minimalist Program. In Pius N. Tamanji and Kiyomi Kusumoto, eds., *NELS 28*, 43 57. GLSA, University of Massachusetts, Amherst.

Boskovic, Zeljko. 1998b. Multiple *wh*-fronting and economy of derivation. In Emily Curtis, James Lyle, and Gabriel Webster, eds., *Proceedings of the 16th West Coast Conference on Formal Linguistics*, 49 63. Stanford, Calif.: CSLI Publications. [Distributed by Cambridge University Press.]

Cheng, Lisa. 1991. On the typology of *wh*-questions. Doctoral dissertation, MIT, Cambridge, Mass.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson and Paul Kiparsky, eds., *A festschrift for Morris Halle*, 232 286. New York: Holt, Rinehart, and Winston.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freidin, ed., *Principles and parameters in comparative grammar*, 417 454. Cambridge, Mass.: MIT Press. [Also in Noam Chomsky, *The Minimalist Program*, 129 166, Cambridge, Mass.: MIT Press (1995).]

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale and Samuel Jay Keyser, eds., *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 1 52. Cambridge, Mass.: MIT Press. [Also in Noam Chomsky, *The Minimalist Program*, 167 217. Cambridge, Mass.: MIT Press (1995).]

Chomsky, Noam. 1994. Bare phrase structure. (MIT Occasional Papers in Linguistics 5.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass. [Also in Gert Webelhuth, ed., *Government and Binding Theory and the Minimalist Program*, 383 439. Oxford: Blackwell (1995).]

Chomsky, Noam. 1995. Categories and transformations. In *The Minimalist Program*, 219 394. Cambridge, Mass.: MIT Press.

Chomsky, Noam, and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8, 425 504.

Chomsky, Noam, and Howard Lasnik. 1993. The theory of principles and parameters. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, eds., *Syntax: An international handbook of contemporary*

research, 506 569. Berlin: Walter de Gruyter. [Reprinted in Noam Chomsky, *The Minimalist Program*, 13 127. Cambridge, Mass.: MIT Press (1995).]

Collins, Chris. 1997. *Local economy*. Cambridge, Mass.: MIT Press.

Haegeman, Liliane. 1992. Sentential negation in Italian and the Neg Criterion. In *Geneva generative papers 0*, 1 38. Département de Linguistique Générale, Université de Genève.

Haider, Hubert. 1986. Affect a: A reply to Lasnik and Saito, On the nature of proper government. *Linguistic Inquiry* 17, 113 126.

Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale and Samuel Jay Keyser, eds., *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 53 110. Cambridge, Mass.: MIT Press.

Higginbotham, James. 1983. Logical Form, binding, and nominals. *Linguistic Inquiry* 14, 395 420.

Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16, 547 593.

Hoffman, Joel. 1995. Syntactic and paratactic word order effects. Doctoral dissertation, University of Maryland, College Park.

Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT, Cambridge, Mass.

Izvorski, Roumyana. 1996. The syntax and semantics of correlative proforms. In Kiyomi Kusumoto, ed., *NELS 26*, 133 147. GLSA, University of Massachusetts, Amherst.

Izvorski, Roumyana. 1998. Non-indicative *wh*-complements of possessive and existential predicates. In Pius N. Tamanji and Kiyomi Kusumoto, eds., *NELS 28*, 159 173. GLSA, University of Massachusetts. Amherst.

Kayne, Richard. 1991. Romance clitics, verb movement, and PRO. *Linguistic Inquiry* 22, 647 686.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Lasnik, Howard. 1995. Verbal morphology: *Syntactic Structures* meets the Minimalist Program. In Paula Kempchinsky and Héctor Campos, eds., *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero*, 251 275. Washington, D.C.: Georgetown University Press.

Moritz, Luc. 1989. Aperçu de la syntaxe de la négation en français et en anglais (Overview of the syntax of negation in French and in English). Mémoire de licence, Université de Genève.

Müller, Gereon, and Wolfgang Sternefeld. 1996. Ā-chain formation and economy of derivation. *Linguistic Inquiry* 27, 480 511.

Pesetsky, David. 1987. *Wh*-in-situ: Movement and unselective binding. In Eric J. Reuland and Alice G.B. ter Meulen, eds., *The representation of (in) definiteness*, 98 129. Cambridge, Mass.: MIT Press.

Pesetsky, David. 1992. Zero syntax, vol. 2. Ms., MIT, Cambridge, Mass.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20, 365 424.

Reinhart, Tanya. 1995. Interface strategies. (OTS Working Papers of Theoretical Linguistics 95-002.) Utrecht Institute of Linguistics OTS, Utrecht University.

Rivero, María-Luisa. 1991. Long head movement and negation: Serbo-Croatian vs. Slovak and Czech. *The Linguistic Review* 8, 319 351.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.

Roberts, Ian. 1992. *Verbs and diachronic syntax*. Dordrecht: Kluwer.

Rudin, Catherine. 1988. On multiple questions and multiple *wh*-fronting. *Natural Language & Linguistic Theory* 6, 455 501.

Stjepanović Sandra. 1995. Short-distance movement of *wh*-phrases in Serbo-Croatian matrix clauses. Ms., University of Connecticut, Storrs.

Takahashi, Daiko. 1994. Minimality of movement. Doctoral dissertation, University of Connecticut, Storrs.

Takahashi, Daiko. 1997. Move-F and null operator movement. *The Linguistic Review* 14, 181 196.

Tsai, Wei-tien Dylan. 1994. On economizing the theory of A-bar dependencies. Doctoral dissertation, MIT, Cambridge, Mass.

Watanabe, Akira. 1992. *Wh*-in-situ, subjacency, and chain formation. (MIT Occasional Papers in Linguistics 2.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Williams, Edwin. 1986. A reassignment of the functions of LF. *Linguistic Inquiry* 17, 265 299.

Chapter 3
Minimalist Inquiries:
The Framework

Noam Chomsky

The remarks that follow are "inquiries," a term intended to stress their tentative character. They are "minimalist" in the sense of the "Minimalist Program," itself exploratory as the term indicates, and in its short career already developing in partially conflicting and attractive directions. What appears here is the first part of an unfinished manuscript. Here, I will keep to general considerations, rethinking the issues and concerns that motivate the program and attempting to give a clearer account and further development of them from one point of view, taking as a starting point the final sections of Chomsky 1995b (henceforth *MP*). 1 That collection reflects an evolution over several years, with fairly radical changes along the way. Much like earlier stages, this one reflects a collective effort, incorporating ideas and proposals of many students and colleagues with no serious attempt at attribution, in fact no way to achieve it. I should, however, like to take the occasion to express my very special indebtedness to Howard Lasnik for many years of close collaboration, which has been extremely rewarding for me and is most inadequately recorded in print, though well known to participants in these enterprises.

3.1
Background

Let us begin by reviewing briefly a series of assumptions, discussed and qualified elsewhere.

First, there is a faculty of language FL, a component of the human mind/brain dedicated to language. Given this endowment, a human infant, but not her pet kitten, will reflexively categorize parts of the confusion around her as "linguistic" and develop rich and highly articulated capacities to enter into these peculiar modes of human thought and action. In contrast, the infant and kitten will, it seems, develop along a

rather similar path in acquiring capacities to deal with many other aspects of the world. 2 FL can be regarded as a "language organ," in the informal sense in which the visual system, the immune system, and the circulatory system are commonly described as organs of the body: not objects that can be removed leaving the rest intact, but subsystems of a more complex structure that we hope to understand by investigating parts that have distinctive characteristics, and their interactions. Despite explicit denials and apparent controversy, this much seems to be generally assumed, at least tacitly.3

Like other organs, FL has an "initial state" S0 that is an expression of the genes. To a good first approximation, it is uniform for the species, apparently also biologically isolated in essential respects and a very recent evolutionary development.4 FL undergoes state changes under triggering and shaping influences of the environment. If Jones's FL is in state L, we say that Jones has (speaks, knows, . . .) the (I-)language L.5 Two immediate tasks of a theory of language are to characterize the languages (states) attained and the shared initial state: the tasks of "descriptive adequacy" and "explanatory adequacy," respectively. We understand Universal Grammar (UG) to be the theory of the initial state, and particular grammars to be theories of attained states.

The language L includes a cognitive system that stores information: roughly, information about sound, meaning, and structural organization. Performance systems access this information and put it to use. Empirical questions arise at once: in particular, to what extent are the performance systems part of FL, that is, language-dedicated, specifically adapted for language?6 On the "sound side," the answer is unclear and disputed; on the "meaning side," the questions are much harder and more obscure for obvious reasons, and judgments can only be highly tentative. A standard working assumption is that performance systems are external to FL. That is a simplifying assumption, not definitely known to be false, though it may well be, perhaps in important ways.7 The issues have had little effect on empirical inquiry into questions of descriptive and explanatory adequacy, but come to the fore within the Minimalist Program.

I will adopt usual conventions for present purposes, recognizing that they are not innocent. I therefore take L to be a cognitive system alone.

I will assume further that L provides information to the performance systems in the form of "levels of representation," in the technical sense.8 The performance systems access these "interface levels." Assume further

that performance systems are of two kinds: sensorimotor systems and systems of thought (to give a name to something very poorly understood). Let us take them (tentatively) to be unitary and distinct, in the sense that all sensorimotor systems access one interface level, and all systems of thought access a distinct interface level. On these assumptions we understand L to be a device that generates expressions Exp = <Phon, Sem>, where Phon provides the "instructions" for sensorimotor systems and Sem the "instructions" for systems of thoughtinformation about sound and meaning, respectively, where "sound" and "meaning" are understood in internalist terms, "externalizable" for language use by the performance systems. 9 Theories of PF and LF seek to spell out the nature of Phon and Sem. I will assume some version of standard theories to be adequate for present purposes, using the conventional term *features* for the properties of language that enter into PF, LF, and the computational system that generates them.

Again, the assumptions are not innocent. Thus, Epstein et al. (1998) pursue a strong derivational approach in which performance systems access the computation itself, dispensing with levels of representation. That articulatory and perceptual systems access the same information (PF) is also far from self-evident, corresponding assumptions on the meaning side even less so. And there are many other questions.

To say that phonetic features are "instructions" to sensorimotor systems at the interface is not to say that they have the form "Move the tongue in such-and-such a way" or "Perform such-and-such analysis of signals." Rather, it expresses the hypothesis that the features provide information in the form required for the sensorimotor systems to function in language-independent ways. Similar observations hold on the (far more obscure) meaning side. The framework imposes a distinction between (a) linguistic expressions Exp = <PF, LF> that are internal to the mind/brain, and (b) observable events, utterances, and actionsexternalization of (mentally constructed) speech acts. No questions arise about the ontological status of the set of expressions {Exp} generated by L; its status is somewhat like that of potential visual images or plans for limb motions.10

Finally, I will assume that the principles-and-parameters (P&P) approach is in important respects on the right track. Within any version of it, the major problem is to discover the principles and parameters, and to show how a particular choice of parameter values and lexicon enters into fixing a language L11and to proceed beyond, to the study of use,

acquisition, pathology, cellular mechanisms, and a wide range of other questions having to do with the place of language in the biological and social worlds.

Whatever its ultimate fate, the crystallization of the P&P approach has contributed to substantial progress in several of these areas. The approach also opens some new questions. Concern for descriptive and explanatory adequacy is as old as the study of language. As soon as the two traditional goals were reformulated within modern generative grammar, serious tension arose between them: the search for descriptive adequacy seems to lead to ever greater complexity of rule systems, varying among grammatical constructions and across languages, whereas the search for explanatory adequacy leads to the conclusion that language structure is largely invariant. It is this tension that has driven the research inquiry of generative grammar from its inception. The P&P framework suggests a way to resolve the tension, thus offering some conception of the form that a genuine theory might take.

It therefore becomes possible to consider some new questions about FL. In particular, we may ask the question, how well is FL designed? Suppose that a super-engineer were given design specifications for language: "Here are the conditions that FL must satisfy; your task is to design a device that satisfies these conditions in some optimal manner (the solution might not be unique)." The question is, how close does language come to such optimal design?

If the question is real, and subject to inquiry, then the P&P approach might turn out to be an even more radical break from the tradition than it seemed to be. Not only does it abandon traditional conceptions of "rule of grammar" and "grammatical construction" that were carried over in some form into generative grammar, but it may also set the stage for asking novel questions that have no real counterpart in the earlier study of language. 12

The Minimalist Program is the attempt to formulate and study such questions. One should bear in mind that it is a *program*, not a theory, even less so than the P&P approach. There are minimalist questions, but no minimalist answers, apart from those found in pursuing the program: perhaps that it makes no sense, or that it makes sense but is premature. The program presupposes the common goal of all inquiry into languageto discover the right theoryand asks further why language is that way. More narrowly, it seeks to discover to what extent minimal conditions of adequacy suffice to determine the nature of the right theory.13

Questions of this kind are not often studied and might not be appropriate at the current level of understanding, which is, after all, still quite thin in a young and rapidly changing approach to the study of a central component of the human brain, perhaps the most complex object in the world, and not well understood beyond its most elementary properties.

The program is recent, and it is too early to assess it with any confidence. My own tentative judgment has two aspects, one methodological, the other substantive.

At the methodological level, the program has a certain heuristic and therapeutic value. It brings to light what might be fundamental problems, where empirical evidence and minimalist expectations conflict. And it encourages us to distinguish genuine explanations from "engineering solutions"a term that I do not mean in any disparaging sense. Problems of descriptive and explanatory adequacy are vast and largely obscure. One tries to overcome them somehow, with special assumptions that are often not independently well motivated, hoping to reformulate the problems in ways that will facilitate further inquiry. Take, say, the study of conditions on extraction of subjects in terms of government and licensing, or attempts to account for the verb-second phenomenon or linear ordering in terms of X-bar theory, with its standard stipulations. Various solutions have been proposed that are useful and enlightening, but we can ask whether they are of roughly the order of complexity of the original problem. If so, it would be wrong to conclude that such proposals lack value; on the contrary, they may and often have opened the way to considerable progress. But we can still ask whether they are genuine solutions. Or consider the target of verb-raising. Evidence has accumulated that the verb can raise to a position higher than T but lower than C, differentiating languages by the position of a functional category a, on current assumptions. 14 But there is reason to doubt that such a can exist; or, to put it differently, if it does, then departures are needed from what appears to be the simplest and most principled form of phrase structure theory. Again questions arise as to whether there is some better way to conceive the matter. The Minimalist Program helps to focus attention on such issues, and perhaps to address them by showing that elimination of descriptive technology yields empirical results that are as good as, possibly even better than, before.

The substantive thesis is that language design may really be optimal in some respects, approaching a "perfect solution" to minimal design specifications. The conclusion would be surprising, hence interesting if true.

3.2
Design Specifications

Proceeding along the course just outlined, we face two questions: (a) What is "good design"? (b) What are the minimal design specifications for FL? The former we may put to the side, not because it is unimportant, but because questions of this nature are common to all rational inquiry. The appropriate place to seek answers is in the hard sciences, where understanding is far deeper and intuitions are far more firmly grounded. Caution is in order in appealing to such considerations. Given some empirically supported conclusion, it is often possible to construct plausible conceptual grounds for it, and for alternatives.

To clarify the problem of design specifications, let us invent an evolutionary fable, keeping it highly simplified. 15 Imagine some primate with the human mental architecture and sensorimotor apparatus in place, but no language organ. It has our modes of perceptual organization, our propositional attitudes (beliefs, desires, hopes, fears, etc.) insofar as these are not mediated by language, perhaps a "language of thought" in Jerry Fodor's sense, but no way to express its thoughts by means of linguistic expressions, so that they remain largely inaccessible to it, and to others. Suppose some event reorganizes the brain in such a way as, in effect, to insert FL. To be usable, the new organ has to meet certain "legibility conditions."16 Other systems of the mind/brain have to be able to access expressions generated by states of FL ((I-)languages), to "read" them and use them as "instructions" for thought and action. We can try to formulate clearlyand if possible answerthe question of how good a solution FL is to the legibility conditions, and these alone. That is essentially the topic of the Minimalist Program.

We have assumed two external systems: sensorimotor systems and systems of thought, each with its own characteristics independent of FL. The former can only use information presented in a specific form: with temporal order, prosodic and syllable structure, certain phonetic properties and relations. The systems of thought require information about units they can interpret and the relations among them: certain arrays of semantic features, event and quantificational structure, and so on. Insofar as we can discover the properties of these external systems (an empirical problem, however difficult), we can ask how well the language organ satisfies the design specifications they impose, providing legible representations at the interface levels. That is the minimal condition FL must satisfy to be usable at all.17

To introduce some terminology of *MP*, we say that a computation of an expression Exp *converges at an interface level IL* if Exp is legible at IL, consisting solely of elements that provide instructions to the external systems at IL and arranged so that these systems can make use of them; otherwise, it *crashes* at IL. The computation *converges* if it converges at all interfaces. Call the expression Exp so formed *convergent* as well. As in *MP*, we keep here to a restricted version of the concept of convergence, setting aside the matter of legible arrangement (which raises all sorts of complex issues), and tentatively assuming it to be irrelevantno slight simplification. Certain features of lexical items are *interpretable*, that is, legible to the external systems at the interface; others are *uninterpretable*. We assume, then, that if an expression contains only features interpretable at IL, it converges at IL. 18

The property *converges at IL* may hold of an expression formed in the course of a derivation that then proceeds on to IL. If, say, particles or adverbs have only LF-interpretable features, then they converge at LF when extracted from the lexicon and at every subsequent stage of derivation to LF. Similarly, an embedded clause may converge, for example, the bracketed subpart of *John thinks [it is raining]*. The phrase "converge at an interface" should not mislead: convergence is an internal property of an expression, detectable by inspection.19

Suppose that in state L, FL generates expressions Exp = <PF, LF>. Then L determines sound-meaning associations: the sounds and meanings determined by PF and LF, respectively, are associated in Exp. These are matters of fact that lie well beyond legibility conditions.20 Take such standard examples as these:

(1) a. PF1: John is impossible to forgive
b. PF2: John is impossible to be forgiven

Suppose L assigns to both a semantic representation LF on the model of *John is unlikely to forgive*. Then the generated expressions converge but with the wrong pairings. Suppose L assigns to both a representation LF' corresponding (closely enough) to that of *it is impossible to forgive John*. Then the association is right for Exp1 = <PF1, LF¢> and wrong for Exp2 = <PF2, LF¢>, though both converge. The assignment is wrong because it does not indicate the deviance of Exp2, a crucial property distinguishing it from Exp1. The conclusion holds even if the language user in state L assigns LF¢ to PF2 by some interpretive mechanism, using L but presumably going beyond.

Suppose all "best ways" to satisfy legibility conditions yield incorrect associations. Then departure from optimal design is required. If there are some that consistently yield the right sound-meaning relations, then we have reason to believe that language design is optimal in nontrivial respects.

Suppose that FL satisfying legibility conditions in an optimal way satisfies all other empirical conditions too: acquisition, processing, neurology, language change, and so on. Then the language organ is a perfect solution to minimal design specifications. That is, a system that satisfies a very narrow subset of empirical conditions in an optimal waythose it must satisfy to be usable at allturns out to satisfy all empirical conditions. Whatever is learned about other matters will not change the conclusions about FL. That would be a strange and surprising result, therefore interesting to whatever extent it might be true. The Minimalist Program explores the possibility that language approaches "good design" in this sense. The strongest minimalist thesis would be this:

(2) Language is an optimal solution to legibility conditions.

Insofar as the thesis is true, information about other matters (sound-meaning connections, neurophysiology, etc.) may be helpful in practiceeven indispensablefor discovering the nature of FL and its states. But it is irrelevant in principle. The tasks of biology of language remain as before, but become even more intriguing and difficult, because a new problem arises: how did the structure of the brain and the course of evolution happen to yield the outcome (2)? The internalist study of languagesyntax in the broad sensebecomes much harder, hence more interesting and significant, because a standard of explanation is set that is very difficult to meet: descriptive machinery must satisfy stringent conditions, imposed by (2). Issues relating to the interface become of central concern. The problem of discovering whether, and if so how, considerations of economy enter into language design also gains new prominence, along with questions about their role in language acquisition. In general, all questions become harder, hence more interesting and significantinsofar as there is some truth to the strong thesis.

In these terms we might also be able to devise an interpretation for a thesis about language and psychology that seems to make little sense, but that has been so widely held that one might suspect that some significant intuition may lie behind it. The thesis is that linguists are to study "linguistic evidence" and "linguistic intuitions," but the results of their work,

however revealing and far-reaching, do not bear on "reality," sometimes called "psychological reality." Other kinds of evidence are required to find out about reality. To take a classic example, Sapir provided rich "linguistic evidence" for phonological analyses he proposed, and went on to adduce much weaker "psychological evidence" to demonstrate their "psychological reality"; this was considered an audacious and controversial move, mainly on grounds that even the psychological evidence doesn't bear on reality.

In the recent period such ideas have appeared often in critical discussion of the program of generative grammar. Similarly, it is sometimes held that conclusions based on linguistic evidence must be confirmed by "converging" evidence from other sources, though conclusions based on these sources stand on their own. Another variant, thankfully put to rest after too many years, is that it is the task of psychologists to test (verify, refute) the theories of linguists, which are based on "linguistic evidence," not to contribute directly to these theories.

The linguistic evidence is generally understood to consist of informant judgments about sound and meaning and their relations. 21 The proposal is odd, as has often been pointed out: evidence does not come with a mark saying "I do or do not bear on reality." Judgments about (la,b) have no different status than other kinds of evidence with regard to the nature of FL, as part of the (physical) world. As in the study of vision and other domains, these judgments are discovered by experiment, typically informal in this case, though they can be carried out as carefully as is necessary to advance understanding. Uncontroversially, one seeks the widest possible range of relevant evidence, converging or conflicting. But there is no principled asymmetry between categories of evidence in this regard.

In terms of the preceding discussion, we might replace these proposals by a substantive (but extraordinarily strong) empirical hypothesis, namely, the thesis (2): an optimal solution to legibility conditions satisfies all other empirical tests as well. The reformulated thesis replaces the obscure notion of "linguistic evidence" by the meaningful notion: satisfaction of interface conditions. Relevant evidence is very limited. The thesis is even stranger than the requirement that "linguistic evidence" suffices to determine grammars, but it has empirical content. One might perhaps suggest it as a kind of rational reconstruction of dubious ideas about linguistic evidence and psychological reality that have appeared in one or another form.

Suppose we understood external systems well enough to have clear ideas about the legibility conditions they impose. Then the task at hand

would be fairly straightforward at least to formulate: construct an optimal device to satisfy just these conditions, and see how well it satisfies other empirical conditions. If all such efforts fail, then add "imperfections" as required. But life is never that simple. The external systems are not well understood. Progress in understanding them goes hand in hand with progress in discovering the language systems that interact with them. So the task is simultaneously to set the conditions of the problem and to try to satisfy them, with the conditions changing as we learn more about how to do so. That is not surprising. It is much what we expect when trying to understand some complex system. We proceed with tentative proposals that seem reasonably firm, expecting the ground to shift as more is learned.

3.3 Architecture

We are taking L to be the recursive definition of a set of expressions Exp = <PF, LF>. We can now raise a questionat least, an apparent questionabout the interpretation of the recursive definition.

One might construe L as a step-by-step procedure for constructing Exps, suggesting that this is how things work as a real property of the brain, not temporally but as part of its structural design. Assumptions of this nature constitute a *derivational approach* to L. The *strong* derivational approach dispenses with the expression altogether, assuming that information is provided to interface systems "dynamically" (see p. 91). A *weak* derivational approach assumes that interface levels exist, allowing "postcyclic" operations that apply to them in whole or in part (deleting the tail of a chain, imposing metrical structure, determining ellipsis and scope, etc.). There are many options.

With richer set-theoretic assumptions, a recursive definition can be restated as a direct definition, in this case, of the following form: E is an expression of L iff . . . E . . . , where . . . . . is some condition on E. One might, then, take L to be a direct definition of the set {Exp}, adopting a *representational approach*. Again there are weaker varieties, for example, the assumption that the set of LF representations is given (universally, or by L), PF being derived from LF by some computational procedure.

The issue is reminiscent of old questions about morphological processes ("item-and-process" vs. "item-and-arrangement," etc.) and grammatical transformations. Thus, does a transformation map an input structure to

an output structure, or is it an operation on the "output" that expresses properties of the "input"? It is unclear whether these are real questions; on the surface they look like the question whether 25 = 52 or 5 = Ö25. If the questions are real, they are subtle. They have elicited no little passion over the years, but it is out of place. 22 The apparent alternatives seem to be mostly intertranslatable, and it is not easy to tease out empirical differences, if there are any.

Surprisingly, there is reason to believe that the questions may be real.23 The evidence that has been adduced is far from conclusive and often conflicting. I will adopt the derivational approach as an expository device, though I suspect it may be more than that. If so, that would be a curious and puzzling fact about the nature of the mind/brain.24

Suppose that the issue is real, and the derivational approach in fact correct. Then further questions arise. For example, we might inquire into the complexity of the generative procedure. Such questions have arisen over the years, in one or another form. One category concerns "least effort" conditions, which seek to eliminate anything unnecessary: (a) superfluous elements in representations, (b) superfluous steps in derivations. The tacit assumption is that failure to meet these conditions imposes deviant interpretations, in principle an empirical issue though often not an easy one to resolve. Subcategory (a) involves legibility conditions and convergence ("full interpretation"); (b) holds that operations are allowed only if there is some reason for them. In the terms we are exploring, reasons are reduced to effects at the interface. Possibilities that have been investigated (if not in these terms) include constraints that bar PF-vacuous overt movement and others that seek to limit effects on PF (Procrastinate). An LF counterpart is that covert operations are allowed only if they have an effect on interpretation at LF. Another category seeks to reduce "search space" for computation: "Shortest Movement/Attract." successive-cyclic movement (Relativized Minimality, Subjacency), restriction of search to c-command or minimal domains, and so on. Yet another imposes "local determinability" conditions (barring "look-ahead," "backtracking," or comparison of alternatives). I will assume these ideas to be generally on the right track and pursue them further below.25

Some of these notions have analogues in formal complexity theory. Most are the kinds of intuitive ideas about "operative complexity" that enter commonly into the cognitive sciences26 and design considerations generally. Suppose automobiles lacked fuel storage, so that each one had

to carry along a petroleum-processing plant. That would add only bounded "complexity," but would be considered rather poor design. Something similar might well be true for language.

Let's consider a few such proposals, beginning with conventional ones and proceeding to others that are more controversial; it's worth bearing in mind, however, that the logic is similar throughout.

UG makes available a set *F* of features (linguistic properties) and operations *CHL* (the computational procedure for human language) that access F to generate expressions. The language L maps F to a particular set of expressions *Exp*. Operative complexity is reduced if L makes a one-time selection of a subset [F] of F, dispensing with further access to F. It is reduced further if L includes a one-time operation that assembles elements of [F] into a lexicon *Lex*, with no new assembly as computation proceeds. On these (fairly conventional) assumptions, acquiring a language involves at least selection of the features [F], construction of lexical items Lex, and refinement of CHL in one of the possible waysparameter setting. 27 One could offer a conceptual argument that conventional assumptions are mistaken, on the grounds that a theory lacking certain concepts (here [F], Lex, and the operations forming [F] and Lex) is better than an otherwise identical one that employs them. But if operative complexity matters, the argument loses force. Conceptual arguments can be given either way. The issues are empirical and can be settled only by investigating consequences of alternative conceptions, considered so obvious in this case that the question has not arisen.

We assume, then, that a language L maps ([F], Lex) to Exp. The next natural simplification would be to reduce access to the domain ([F], Lex) of L. Consider [F]. In the computation of LFwhat we may call *narrow syntax*it seems that [F] is not accessed, only Lex (and features of its items). The restriction does not extend to phonology, however: features are introduced in the course of computation, and in different ways for different languages, whatever approach one takes to computation of PF. Keeping to narrow syntax, then, we may take CHL to be a mapping of Lex to the LF representations of Exp.

Is it also possible to reduce access to Lex, the second component of the domain of L? The obvious proposal is that derivations make a one-time selection of a *lexical array* LA from Lex,28 then map LA to expressions, dispensing with further access to Lex. That simplifies computation far more than the preceding steps. If the derivation accesses the lexicon at every point, it must carry along this huge beast, rather like cars that con-

stantly have to replenish their fuel supply. 29 Derivations that map LA to expressions require lexical access only once, thus reducing operative complexity in a way that might well matter for optimal design.

Again, conceptual arguments can be given either way, but they carry little weight. The questions are empirical. Investigating them, we can hope to discover whether (and if so how) what might reasonably be considered complexity/economy considerations enter into language design.

If FL operates with the economy principles just reviewed, then a language L follows procedures (3a) and (3b) to specify the language (apart from parameter setting), then applies (3c) and (3d) to derive a particular Exp.

(3) a. Select [F] from the universal feature set F.
b. Select Lex, assembling features from [F].
c. Select LA from Lex.
d. Map LA to Exp, with no recourse to [F] for narrow syntax.

We will return to further steps along the same path, but let us first look more closely at general properties of Lex and narrow syntax (the recursive part of L).

First, what operations enter into this component of CHL? One is indispensable in some form for any language-like system: the operation *Merge*, which takes two syntactic objects (a, b) and forms K(a, b) from them. A second is an operation we can call *Agree*, which establishes a relation (agreement, Case checking) between an LI a and a feature F in some restricted search space (its *domain*). Unlike Merge, this operation is language-specific, never built into special-purpose symbolic systems and apparently without significant analogue elsewhere. We are therefore led to speculate that it relates to the design conditions for human language. A third operation is *Move*, combining Merge and Agree. The operation Move establishes agreement between a and F and merges P(F) to aP, where P(F) is a phrase determined by F (perhaps but not necessarily its maximal projection) and aP is a projection headed by a. P(F) becomes the specifier (Spec) of a ([Spec, a]). Let us refer to Move of P to [Spec, f] as A-*movement*, where f is an agreement feature (f-feature); other cases of Move are $\bar{A}$-movement.

Plainly Move is more complex than its subcomponents Merge and Agree, or even the combination of the two, since it involves the extra step of determining P(F) (generalized "pied-piping"). Good design conditions would lead us to expect that simpler operations are preferred to more

complex ones, so that Merge or Agree (or their combination) preempts Move, which is a "last resort," chosen when nothing else is possible. Preference for Agree over Move yields much of the empirical basis for Procrastinate and has other consequences, as do the other preferences. 30

Let us turn next to the lexicon Lex, adopting some fairly common assumptions along with more controversial ones, and keeping to simple cases.

LIs fall into two main categories, substantive and functional; we are concerned now mainly with the latter. Take the *core functional categories* CFCs to be C (expressing force/mood), T (tense/event structure), and n (the "light verb" head of transitive constructions). All CFCs may have f-features (obligatory for T, n). These are uninterpretable, constituting the core of the systems of (structural) Case agreement and "dislocation" (Move). Neither T nor n assigns inherent Case; other light verbs may, as may substantive categories.31

Consider the selectional properties of CFCs, beginning with semantic (s-)selection.32 Assume that C can be unselected (root), whereas n and T cannot. C is selected by substantive categories, n only by a functional category. T is selected by C or V. If selected by C, it has a full complement of f-features; if by V, it is *defective* (Tdef). C selects T; T and n select verbal elements. n may also select a nominal phrase NP/DP as its *external argument* EA = [Spec, n].33

Each CFC also allows an extra Spec beyond its s-selection: for C, a raised *wh*-phrase; for T, the surface subject; for n, the phrase raised by object shift (OS). For T, the property of allowing an extra Spec is the Extended Projection Principle (EPP). By analogy, we can call the corresponding properties of C and n *EPP-features*, determining positions not forced by the Projection Principle. I will restrict attention to XP positions, though a fuller picture might add X0 as another case of the EPP (see note 92). EPP-features are uninterpretable (nonsemantic, hence the name), though the configuration they establish has effects for interpretation.

Basic structural properties of CFCs are illustrated in the configuration (4), where H is the CFC, XP is the extra Spec selected by its EPP-feature, and EA is the external argument selected by H = n.

(4) a = [XP [(EA) H YP]]

The following properties hold among these elements:

(5) a. If H is n/C, XP is not introduced by pure Merge.
b. In the configuration [b Tb . . .a], b minimal,

i. if H is C, Tb is independent of a;
ii. if H is n, Tb agrees with EA, which may raise to [Spec, Tb] though XP cannot;
iii. if H is Tdef, XP raises to [Spec, Tb] if there is no closer candidate g for raising. 34

Pure Merge is Merge that is not part of Move. The relevant properties of Tb have to do with Case/agreement and the EPP. In (5bii), if EA does not raise, [Spec, Tb] is introduced by pure Merge to satisfy the EPP. The case of H = nondefective T is omitted in (5b): if (5bi) holds for C, it holds for (nondefective) T selected by C. In fuller generality, b in (5b) should be taken to be the minimal b containing a headed by any CFC Hb, which would therefore be either T or n. The relations of Tb to a extend partially to n as wellspecifically, exceptional case marking in (5biii) as compared to raising to [Spec, T]. I put this aside temporarily, for ease of exposition and because of some differences to which I will return.

Property (5a) follows in part from the q-theoretic principle (6). which is implicit in the conception of q-roles as a relation between two syntactic objects, a configuration and an expression selected by its head.35

(6) Pure Merge in q-position is required of (and restricted to) arguments.

We can therefore restrict attention in (5a) to expletives (*Expl*): for arguments it follows from (6). (6) also applies to (5b), restricting pure Merge in [Spec, Tb] to Expl.

The property (5bi) is illustrated in (7), (5bii) in (8), and (5biii) in (9).

(7) a. there are questions about [a what C [TP John read *t*]]
b. there is a possibility [a that proofs will be discovered]

(8) Tb [a[DO the book] [[Subj many students] [read *t*DO]]]

(9) a. Tb-is likely [a there to be a proof discovered]
b. Tb [nP I expected [a there to be a proof discovered]]

In (7) a = CP with H = C, and the relevant properties of matrix T (Tb) are independent of a in accord with (5bi). a is a closed system with regard to Case/agreement properties, determined internally without outside effect.

In (8) a is an OS construction with H = n and DO = XP of (4)/(5bii). Under Holmberg's Generalization V raises to Tb, which agrees with EA = Subj. EA can then raise to [Spec, Tb] (*many students read-pl the book (never)*) or remain in situ with Merge of Expl to satisfy the EPP (*there read-pl the book (never) any students*), illustrating the two options

for (5bii). The positions of DO and Subj can be determined by left-edge adverbs. 36

In (9) a = TP with H = Tdef and XP = *there*. XP raises to [Spec, Tb] in (9a), but not in (9b) with intervening g = *I*, yielding (10a) and (10b), respectively, which illustrate (5biii).

(10) a. there is likely to be a proof discovered
b. I expected there to be a proof discovered

Properties of matrix T (= Tb) depend on a in (9a) but not in (9b), where it is the head of nP (not Tb) that is related to a. The relation of Tb to a in (9a) is analogous to that of n to a in (9b), but there are differences to which we will return.

The order object-subject in (8) follows from the assumption that Merge preempts the more complex operation Move. When the derivation has reached the stage (11), two operations have to take place for an OS construction, each creating Spec: Merge of EA and raising of DO (Move).

(11) [n[V DO]]

If Merge applies first, the order is as in (8).
The examples in (9)/(10) contrast with those in (12).

(12) a. *there is likely [a a proof to be discovered]
b. *I expected [a *t* to be a proof discovered]
c. I expected [a a proof to be discovered]

Suppose the derivation has reached reached the stage (13), analogous to (11), with Ta = Tdef.

(13) [Ta [be a proof discovered]]

The EPP requires that something occupy [Spec, Ta]. Two options are available: merge *there* or move *a proof*. Preference of Merge over Move selects the former. Accordingly, (9a)/(10a) is permitted and (12a) is barred. But Merge of an argument in [Spec, Ta] violates the q-theoretic condition (6). Therefore, (12b) is barred. Either an expletive is merged, yielding (10b), or Move applies, yielding (12c). The choice depends on whether or not an expletive is available in the initial lexical arraythe first time step (3c) enters.

Four kinds of complexity considerations enter into this account:

(14) a. Simple operations preempt more complex ones.
b. Search space is limited (locality).

c. Access to the feature set F is restricted by (3).
d. Computation is locally determined (no look-ahead).

Conclusion (14d) follows from (14a c) and (6), all plausible principles of some generality; in the background is the derivational approach (cyclicity). Alternative analyses have been proposed, but to my knowledge they are not locally determinable or they introduce special or dubious assumptions.

By the reasoning just reviewed, other constructions similar on the surface to (12c) could be legitimatefor example, (15), where V is a *seem*-type verb with an optional Spec 37 that is raised to the matrix Spec in (15a) and missing in (15b).

(15) a. me(dat) V [TP several people to be in the room]
b. there V [TP several people to be in the room]

(15a) is similar to (12c): Move applies in the embedded TP in preference to Merge of *me*, in accord with (6); *me* then merges as Spec of the *seem* construction and raises to matrix subject. Case (15b) could arise if the embedded clause is a multiple-subject construction, in which both Merge (of Expl) and Move (of *several people*) have applied, Expl then raising to the matrix Spec. Both possibilities appear to be legitimate in languages with the relevant properties.38

Control infinitivals, I have assumed, fall together with finite clauses, headed by C selecting nondefective T (with tense-modal structure and a full complement of f-features). Like other CPs, they generally undergo movement and clefting and can appear as root expressions (typically with *wh*-phrase Spec or as discourse fragments), and structural Case is assigned to the subject of T. These properties are common to CPs and distinguish them from raising/ECM infinitivals headed by Tdef, lacking C and tense structure and assigning no Case to subject, and lacking the distributional freedom of CP.39 We also find convergent constructions like those in (16), analogous to (7) and contrasting with (12a).

(16) a. it's fun [a PRO to [*t* go to the beach]]
b. it's about time [a PRO to [leave *t*]]
c. it was decided [a PRO to be executed *t* at dawn]

Raising is possible throughout in the closed system a (as in (7)). In (16c), at least, PRO is controlled by an implicit argument: it can mean that the prisoners decided that they would be executed at dawn, but not that we decided that they would be.

A problem throughout the whole account is why raising is *ever* possible, if Agree and Merge preempt Move. The question is answered in part by the q-theoretic principle (6), which bars pure Merge of arguments in non-q-positions and correspondingly restricts Move to such positions. Choice of Move over Agree follows from presence of EPP-features, where pure Merge is inapplicable. The remaining question, then, is why Merge of Expl does not always bar Move. That question is partly answered by the initial choice of lexical array: it may or may not make Expl available. But that cannot be the whole story, as illustrated in (7) and (16), where Expl is available in the lexical array but Move takes place in the embedded phrase a. 40

A straightforward solution would be to take the derivational approach still more seriously and further extend the procedures in (3) that reduce access to the domain of L. Suppose we select LA as before, under (3c); the computation need no longer access the lexicon. Suppose further that at each stage of the derivation a subset LA*i* is extracted, placed in active memory (the "workspace"), and submitted to the procedure L. When LA*i* is exhausted, the computation may proceed if possible; or it may return to LA and extract LA*j*, proceeding as before. The process continues until it terminates. Operative complexity in some natural sense is reduced, with each stage of the derivation accessing only part of LA. If the subarray in active memory does not contain Expl, then Move can take place in the corresponding stage; if it does, Merge of Expl preempts Move.41

The next step is to characterize the subarrays LA*i* that can be selected for active memory. LA*i* should determine a natural syntactic object SO, an object that is relatively independent in terms of interface properties. On the "meaning side," perhaps the simplest and most principled choice is to take SO to be the closest syntactic counterpart to a proposition: either a verb phrase in which all q-roles are assigned or a full clause including tense and force. Call these objects *propositional*. Considerations on the "sound side" support the choice, given properties of the kind mentioned earlier distinguishing CP from TP, which extend to nP (fronting, extraposition, pseudoclefting, response fragments, etc.).

LA*i* can then be selected straightforwardly: LA*i* contains an occurrence of C or of n, determining clause or verb phraseexactly one occurrence if it is restricted as narrowly as possible, in accordance with the guiding intuitions. Take a *phase* of a derivation to be an SO derived in this way by choice of LA*i*. A phase is CP or nP, but not TP or a verbal phrase headed by H lacking f-features and therefore not entering into Case/agreement

checking: neither finite TP nor unaccusative/passive verbal phrase is a phase. Suppose phases satisfy a still stronger cyclicity condition:

(17) The head of a phase is "inert" after the phase is completed,
triggering no further operations.

A phase head cannot trigger Merge or Attract in a later phase, and we can restrict attention to phases in which all selectional requirements are satisfied, including the EPP for T (by virtue of cyclicity) and for n/C, and selection of EA for n if required; otherwise, the derivation crashes at the phase level.

Derivations proceed phase by phase: (18), for example, has the four phases shown in brackets.

(18) [John [*t* thinks [Tom will [*t* win the prize]]]]

An alternative that has been suggested (class lectures, MIT, 1995, and various talks and papers) is to define phases in terms of convergence. The two options are then as follows:

(19) a. Phases are propositional.
b. Phases are convergent.

Under (19a) LA*i* is determined by a single choice of C or n. Under (19b) local determination is not possible. Complexity considerations therefore favor option (19a), and again the empirical evidence supports the same conclusion.

The two options have similar (though not identical) consequences in such cases as (18), but are clearly distinguished elsewhere. One case is A-movement, as in (20). 42

(20) which article is there some hope [a that John will read *t*wh]

For reasons to which we will return, assume that the *wh*-phrase has an uninterpretable feature analogous to structural Case for nouns, which requires it to move to its final position in an appropriate C. Then a is a phase under option (19a) but not option (19b) (it does not converge, containing an uninterpretable feature). The only phase is (20) itself; merger of *there* blocks raising of *John* to [Spec, TP] within a, so (20) is underivable without look-ahead. That is unnecessary under option (19a), where a is derived from LA*i* lacking Expl.

The descriptive typology of movement, a leading research topic for years,43 offers other reasons to suspect that phases are real, understood under option (19a). There are several categories: movement can be

feature-driven or not, and in the former case can be directly or indirectly feature-driven. Typical cases include raising to subject (directly feature-driven), the nonfinal stages of successive-cyclic movement (indirectly feature-driven), and QR and "stylistic movement" (perhaps not feature-driven). 44

Indirect feature-driven movement (IFM) subdivides into types depending on the attracting head H in the final stage: (a) A-movement when H has f-features (yielding the Case/agreement system), or (b) $\bar{A}$-movement when H has *P-features* of the peripheral system (force, topic, focus, etc.).45 The intuitive argument for IFM has always been that locality conditions require "short movement" in successive stages, leading to convergence in the final stage. We can express a version of this idea as a "phase-impenetrability condition," strengthening further the notion of cyclic derivation. Given HP = [a [H b]], take b to be the *domain* of H and a (a hierarchy of one or more Specs) to be its *edge*. The thesis under consideration is (21).

(21) *Phase-Impenetrability Condition*
In phase a with head H, the domain of H is not accessible to operations outside a, only H and its edge are accessible to such operations.

The cycle is so strict that operations cannot "look into" a phase a below its head H. H itself must be visible for selection and head movement; hence, its Specs must be as well. The Phase-Impenetrability Condition yields a strong form of Subjacency.46 For A-movement, it should follow from the theories of Case/agreement and locality.47 The stipulation is for clausal $\bar{A}$-movement, the basic question from the earliest study of these topics. We will return to some speculation about reducibility to economy conditions.

The intended consequences do not follow if phases are determined by convergence, but the propositional alternative (option (19a)) accommodates them, verbal phases being limited to transitive n with f-features and EA. The Phase-Impenetrability Condition requires that $\bar{A}$-movement, target the edge of every phase, CP and nP. There is evidence from reconstruction effects and parasitic gap constructions that this may be true.48

The idea that IFM applies only if needed to guarantee eventual convergence appears to raise questions of look-ahead. These are obviated if the Phase-Impenetrability Condition holds. Local determination is straightforward, and an uninterpretable feature in the domain determines at the phase level that the derivation will crash.49

Let us return to the basic configuration (4) for CFCs, repeated as (22), where XP is the extra Spec determined by the EPP-feature of the attracting head H.

(22) a = [XP [(EA) H YP]]

Typical examples of (22) are raising to subject (yielding (23a)), OS (yielding (23b), with XP = DO and *t* its trace), and overt Ā-movement (yielding (23c), with H = C and XP a *wh*-phrase; see note 31).

(23) a. XP [T YP]
b. XP [Subj [n [V *t*]]]
c. XP [C YP]

The EPP-feature of T might be universal. For the phase heads n/C, it varies parametrically among languages and if available is optional.

Cases (23b) and (23c) are alike, and unlike (23a), in other respects. One is that the EPP-feature can be satisfied by Merge of an Expl in (23a), but not in (23b) and (23c) (see (5a)). This fact might be unexpected for (23b), because (23b) and (23a) enter into the Case/agreement system in much the same way; and the relation of [Spec, H], H, and a related phrase b in the complement of H appears to be similar for T and n.

Phases are determined by a choice of C/n, not T, which suggests a basis for the similarities and asymmetry. The fact that the EPP-feature when available is optional for C/n suggests that it is a property of the phase Ph.

(24) The head H of phase Ph may be assigned an EPP-feature. 50

Once Ph is completed, exhausting the lexical subarray from which it is derived, (24) may optionally apply, assigning an EPP-feature to H. From the strong cyclicity condition that renders H inert beyond the phase itself (see (17)), it follows that the EPP must be satisfied by raising within Ph: pure Merge from outside Ph is barred. In (23b,c) XP is raised from within the domain of H = n/C, completing the account for case (5a).51 The remaining properties of (5) should follow from the theories of Case/agreement and locality, to which we will turn shortly.

The picture might be extended to incorporate QR, if alongside the f-and P-features that drive movement, there are also QU-features, attracting quantifiers that pied-pipe an appropriate phrase. One might also explore a generalization of the idea that operations can apply only if they have an effect on outcome; see text at note 25. This would translate here into restricting (24) to the condition in which it permits IFM or specific

interpretations associated with peripheral positions (e.g., specificity and informational conditions on OS). 52 There are many problems and ramifications, which I will leave to the side.

(24) yields A- or $\bar{A}$-movement depending on whether the phase head has f- or P-features. It might have both. Suppose that in the construction (25) all four phase heads are assigned an extra Spec by (24), associated with P-features for C and n2 but not for n1.

(25) [Spec, C2] . . . [Spec, n2] . . . [Spec, C1] . . . [Spec, n1] . . . XP

XP raises through the Specs in succession, landing finally in [Spec, C2]. The result is the four-membered $\bar{A}$-chain ([Spec, C2], [Spec, n2], [Spec, C1], [Spec, n1]) and the two-membered A-chain ([Spec, n1], XP) (formed by OS). [Spec, n2] is an $\bar{A}$-position, by virtue of the P-feature associated with the extra Spec introduced by (24); n2 also had f-features involved in object Case/agreement but these would have been deleted phase-internally before (24) assigns the extra Spec.53

There are more complex cases to consider, along with a variety of other issues.54 But it seems that the cyclic approach to accessing lexical arrays is plausible on conceptual and empirical grounds, along with option (19a) and the condition (24) on extra Specs.

Nothing has been said yet about the fact that CHL forms syntactic objects in parallel, according to the derivational approach adapted from *MP*. Consider, for example, the expression (26).

(26) the demonstration that glaciers are receding showed that global warming must be taken seriously

The prefinal phases of the derivation are the syntactic objects corresponding to (27a c).55

(27) a. P1 = [CP that global warming must be taken seriously]
b. P2 = [CP that glaciers are receding]
c. P3 = [nP [the demonstration P2 [show P1]]]

For each new phase, a subarray provides the lexical material required and the operations proceed in the manner already sketched, with P1/P2 unordered. Step (27c), for example, is formed from the subarray {*the, demonstration, show*}; repeated Merge yields DP = *[the [demonstration* P2]] and then P3 = [DP [*show* P1]], with q-roles assigned. The next subarray adds T and C, permitting the derivation to continue to P4 = (26).

We have found evidence that computational complexity enters into language design, but one might ask whether the argument can be strengthened

on purely conceptual grounds. Consider theses (28a d), where (28b d) presuppose a positive answer to (28a).

(28) a. Computational complexity matters for a cognitive system.
b. The solution must be comprehensive, with a guarantee of "quick decision" for all derivations.
c. Complexity should not be allowed to grow "too fast."
d. Decisions in computation attend only to principles of UG.

In recent discussion such theses are sometimes adopted as virtual conceptual necessity. 56 That is a questionable move, however.

One reason is that the theses seem to presuppose that the derivational approach is correct, which is hardly obvious (if the question is even meaningful). That aside, it is not clear why thesis (28a) should be true. Thesis (28b) raises further questions. Languages L have expressions that satisfy all conditions of UG and L but are "unusable" for some reason; see note 17. It could turn out that among these are expressions that cannot be derived efficiently, a result that would be interesting if true (see Chomsky 1991).

Interpretation of thesis (28c) in the technical sense of complexity theory also calls for justification. Why should we expect such properties to be relevant to natural language? Thesis (28d) is intended to bar look-ahead by requiring that at each choice point in a derivation, an irrevocable decision must be made in terms of principles of UG. The intuitive idea is that only "fundamental principles" of UG can be invoked, not consequences of these principles (however easily determined). Again, that is not an obvious conclusion.57

Theses (28c) and (28d) are often held to undermine the account of (9) (10)/(12)/(15)/(16) reviewed earlier on the grounds that it involves look-ahead. If that were true, the validity of the theses would be called into question, not the analysis, unless some justification can be found for them. Notice how difficult it is even to give a clear formulation of thesis (28d), hence of the argument that the proposed analysis of these constructions even raises look-ahead issues.58 The assumption is that thesis (28d) is violated by resort to the q-theoretic principle (6), though apparently not by preference of Merge over Move. Why is unclear: selection of Merge over Move or conversely is determined (trivially) by principles of UG. Furthermore, how do we know that (6) is a "theorem," not a principle of language design (part of q-theory), so that resorting to it in fact satisfies the intuition that lies behind this application of thesis (28d)?

Again, it seems that we should seek a resolution of the issues on empirical grounds. My suspicion is that thesis (28a) might be correct and perhaps (28b), along with versions of (28c,d) that restrict choices in derivations to elementary principles of UG and bar even narrowly bounded look-ahead. But the questions are empirical. If the theses of (28) hold in some form, that would be a surprising empirical fact about language design; evidence is required to establish it. The best evidence I know is twofold: observed cases seem to support thesis (28a) and to be consistent with (28b). Inquiry seems to reveal further that postulated properties of language that induce computational complexity were incorrectly formulated and that when they are improved, undue complexity is overcome and computation is "local," suggesting that thesis (28c) may be valid. 59 Furthermore, there seems no need to invoke principles beyond the most elementary ones at any point. If such results are consistently found, they would provide confirming evidence for versions of (28), and for the derivational approach, which provides the framework within which they are formulated.

3.4
The Strong Thesis

Let us now consider the strongest minimalist thesis (2) and ask where it fails. The research strategy is to seek "imperfections" of language, properties that language should not have, assuming (2). Apparent imperfections come in several varieties. Given some apparent property P of language, the following outcomes are possible:

(29) a. P is real, and an imperfection.
b. P is not real, contrary to what had been supposed.
c. P is real, but not an imperfection; it is part of a "best way" (perhaps not unique) to meet design specifications.

The most likelyhence least interestingoutcome is (29a). Conclusion (29b) is less likely, therefore more interesting; efforts within the Minimalist Program have sought to establish it in particular cases. Conclusion (29c) is the most interesting possibility. The question is novel, perhaps prematureor unreasonable, because the strongest minimalist thesis (2), or even weaker versions, is not correct, as one might expect for some complex biological system.

Adopting (2) as a point of departure, assume that FL provides no machinery beyond what is needed to satisfy minimal requirements of

legibility and that it functions in as simple a way as possible. We would like to establish such conclusions as these: 60

(30) a. The only linguistically significant levels are the interface levels.
b. The *Interpretability Condition*: LIs have no features other than those interpreted at the interface, properties of sound and meaning.
c. The *Inclusiveness Condition*: No new features are introduced by CHL.
d. Relations that enter into CHL either (i) are imposed by legibility conditions or (ii) fall out in some natural way from the computational process.61

Condition (30a) requires that there is no Deep or Surface Structure, or other levels that have been proposed. It holds that everything accounted for in these terms has been misdescribed and is better understood in terms of legibility conditions at the interface: that includes the Projection Principle, binding theory, Case theory, the Chain Condition, and so on.

Condition (30b) is transparently false.

Condition (30c) permits rearrangement of LIs and of elements constructed in the course of derivation, and deletion of features of LIbut optimally, nothing more.

Condition (30d) has to be spelled out. Properties induced by legibility conditions might include adjacency, argument structure, scope, and the like. Those of category (ii) should include at least (perhaps at most) the relations provided directly by the indispensable computational operation Merge. But there should be no government, no stipulated properties of chains, no binding relations internal to language, no interactions of other kinds. It is hardly necessary to observe that all of this is highly unlikely. There is substantial empirical evidence supporting the opposite conclusion at every point. Furthermore, a basic assumption of the work in the P&P framework, with its impressive achievements, is that everything just suggested is false: that language is highly "imperfect" in these respects. It would be no small task, then, to show that this apparatus is a kind of descriptive technology, and that if we abandon it, we can maintain or even extend descriptive and explanatory force. Nevertheless, recent work suggests that such conclusions, which seemed out of the question a few years ago, are at least plausible, possibly correct in nontrivial respects.

Much of *MP* is devoted to establishing parts of (30) in terms of bare phrase structure, with a cyclic notion of generalized transformations and

reinterpretations of binding and Case theory. I will assume that these are the right directions, though many serious problems remain.

Consider condition (30c). It requires that there be no phrasal categories or bar levels, hence no X-bar theory or other theory of phrase structure, apart from bare phrase structure, which will be simplified further below. It also rules out introduction of traces, indices, l-operators, and other new elements in the course of operation of CHL. Recourse to such devices could be innocuous (e.g., if used for convenience to annotate properties that can be determined by inspection at LF), but questions arise if they enter into interpretation and function significantly within the computationfor example, percolation of indices, or operations that apply specifically to trace.

In *MP* indices are introduced, but not within CHL itselfrather, in extending lexical arrays to numerations (see note 28). This narrow departure from condition (30c) could be eliminated by weakening the requirement that an item of a lexical array be removed when accessed in computation, leaving this as an option. Whether that is the correct move depends on the consequences. A minor matter is that it would require modification of interpretive procedures at the LF interface. More important, it would require a new notion of chain. That seems to be the only significant consequence, and it merits some thought.

If a in the syntactic object SO is merged somewhere else (by the operation Move) to form SO', then the two occurrences of a constitute a chain, the original occurrence called the *trace or copy* of the new one. The terminology is misleading, for several reasons. First, each of the elements is a "copy" of the other. Second, copy theory is the simplest version of transformational grammar, making use only of Merge, not Merge followed by an operation that deletes the originaland, under trace theory, a further operation that creates a new kind of element, trace, a serious violation of the Inclusiveness Condition. These are "imperfections," to be avoided unless shown to be necessary. I will continue to use the terminology, but only for expository convenience, adopting the copy theory as the null hypothesis.

62

If LAs are extended to numerations Num, a chain can be defined as a sequence of identical a'smore accurately, a sequence of *occurrences* of a single a. That seems to be the simplest way to characterize chains; hence, it is to be adopted unless shown to be empirically inadequate. Thus, in (12c), repeated as (31), there are two occurrences of *I* and two occurrences of *a proof* (one where the terms are spelled out and another at the trace).

(31) Ii T [*t*i expect [[a proof]j to be discovered *t*j]]

Chains so-defined can be formed only by movement, given LA = Num. Basic properties of chains should then follow from elementary derivational principles. That would be a good result, eliminating stipulated properties of chains and explaining why these hold. If achievable, it would also bear on broader questions that arose in section 3.3, providing support for the derivational interpretation of the recursive function CHL, hence for the surprising conclusion that there is even an empirical issue concerning alternative interpretations.

What exactly do we mean by an "occurrence of a"? To ensure that occurrences in the intended sense are distinguished, we can take an occurrence of a in K to be the full context of a in K. 63 In *MP* a simpler notion is proposed: an occurrence of a is a sister of a. Then in (31) the occurrence of matrix subject *I* is I1 = *T . . . discovered tj* and the occurrence of its embedded copy is I2 = *expect . . . discovered tj* (actually, the syntactic objects corresponding to them). The occurrences of *a proof* and its trace are P1 = *to be discovered tj* and P2 = *discovered*, respectively.

In (31) two chains are defined: CI = <I1, I2> and CP = <P1, P2>. In informal description CI is the chain <*I, t*i> and CP the chain <*a proof, t*j>, where *I* and *a proof* are called the *heads* of the respective chains.

In (31) occurrences are properly identified if taken to be sisters, but that might not always be the case.64 The simplified definition in terms of sisterhood is based on the assumption that other properties of CHL guarantee that no problems arise. I will assume that to be true.

We can simplify chains from sequences to sets, relying on the fact that a "higher" occurrence of a properly contains lower ones. Thus, in CI of (31) I1 properly contains I2 (*I* c-commands its trace). Chains formed by successive-cyclic movement fall into place directly. Consider (32).

(32) a. a proof is likely [a *t* to be [discovered *t*]]
b. who did you say [*t* has [*t*' discovered the proof]]

In (32a) each *t* = *a proof*, and the chain headed by the matrix subject is {P1, P2, P3}, the three occurrences of *a proof* (P1 = *is likely* a, etc.). In (32b), on the assumptions of *MP*, the external system at the LF interface requires an operator-argument analysis, with *who* "reconstructed" to an operator binding the (also "reconstructed") A-chain argument headed by *t*.65 A straightforward implementation can be based on analysis into two chains: the $\bar{A}$-chain (*who, t*) and the A-chain (*t, t*'). The three occurrences of *who* are analyzed into the chains {Q1, Q2} and {R1, R2}, where Q1 =

did you say . . . the proof, Q2 = R1 = *has . . . the proof*, and R2 = *discovered the proof*. This is one of the permissible analyses into chains; it need only be permitted, not forced. In (32a) there is a three-membered chain; in the (interpretable) analysis of (32b) there are two two-membered chains, sharing one occurrence of *who*. 66

A chain, then, is a set of occurrences of an object a in a constructed syntactic object K. It would make sense to rethink all notions involving chains in these terms, including their interpretation at the interface. If we do, some apparent problems disappear. Suppose that raising of DP to [Spec, T] checks and deletes its uninterpretable Case feature. We want DP and its trace *t* to be identical, so the feature must also delete in the trace. But what guarantees that the feature is deleted throughout the chain? The question does not arise if we think of a chain, more properly, as a set of occurrences of a in K: the feature is deleted in the single element a, unproblematically.

Should the notions of occurrence and chain be extended to proper subparts of LI that are not elements of the lexical array: features and sets of features? Let's put the question aside for the moment and turn to the relations permissible under condition (30dii). Consider the operation Merge (indispensable in some form). Merge takes two objects a and b and forms a new object K(a, b). The operation provides two relations directly: *sisterhood*, which holds of (a, b), and *immediately contain*, which holds of (K, a), (K, b), and (K, K) (taking it to be reflexive). Suppose we permit ourselves the elementary operation of composition of relations. Applying it in all possible ways, we derive three new relations: the transitive closure *contain of immediately contain, identity* (= (sister(sister)), and *c-command* (= sister(contain)). Thus, K contains a if K immediately contains a or immediately contains L that contains a; conversely, a is a *term* of K if K contains a. And a c-commands b if a is the sister of K that contains b.67

The relation of c-command is available, and expected, on very weak assumptions. The relation has played a large role in syntactic theory, though it may be that it does not function within narrow syntax but only in interpretation of the information it providesthat is, in mapping it to syntactic objects that belong to mental systems external to the language faculty itself (see note 44). That might be expected if external systems access representations (PF and LF) to which the notions "sister" and "contain" apply.

The sisterhood relation is significant primarily (perhaps only) for heads, that is, LIs and modified LIs (MLIs) formed from them. Further-

more, sisterhood relations presumably remain if LI is modified to MLI: if VP is sister of T, for example, it should remain so even if uninterpretable features are deleted from T. More generally, LI and its modifications are not distinguished with regard to the fundamental relations defined in terms of Merge. For a an LI, then, we extend the relations defined for (a, b) to (MLI(a), b), MLI constructed from a (typicallyperhaps onlyby deletion of uninterpretable features). We extend the same convention to a feature F(a) of a head a. Thus, if LI a is the sister of b or c-commands b, then MLI(a) and F(a) do as well.

Questions have arisen about the interpretation of these notions for adjunction, particularly head adjunction of H' to H. With no further elaboration, H and H' would be sisters and neither would c-command outside. Whether this matters is unclear. Failure of H' to c-command its trace seems to have no significance in the present framework; failure of H to c-command into its former c-command domain would be problematic only if H functions after adjunction in implementing agreement and movement, but that does not seem necessary. Pending some good reason to sharpen the sisterhood relation for this case, I will leave it as is. 68

Though varieties of government would be "imperfections," to be avoided if possible, the closer-to-primitive notion of L-marking should pass muster, hence also notions of barrier that are based on nothing more than L-marking. Here numerous questions arise about island conditions within a minimalist framework, about which I have nothing useful to say.69

3.5 Imperfections

There are some respects in which the strong thesis seems untenable, and we find what appear to be "design flaws" that are not necessary for language-like systems. The most obvious involve the phonological component, which takes syntactic objects constructed by the computational operations CHL and converts them to representations at the PF interface. Here there are radical violations of the Interpretability and Inclusiveness Conditions (30b,c). The Inclusiveness Condition is violated by operations that introduce such new elements as prosodic structure and narrow phonetics. The Interpretability Condition is violated by the discrepancy between the phonological properties of LIs ("morphophonemes," "phonological units," etc., within various frameworks and terminologies) and the narrow phonetic instantiations of combinations of such elements. It

may be that phonological features of LIs do not even appear at the level PF, that the "input" and "output" of the phonological component are in different "vocabularies." In that case the Interpretability Condition would be maximally violated by the operations of the phonological component, and the Inclusiveness Condition is clearly inoperative.

We therefore turn to the alternatives of (29). Option (29b) seems implausible; the properties appear to be real, whatever their proper expression. Of the two remaining possibilities, the more interesting by far is (29c). The properties are real, but not imperfections: the super-engineer, called upon to map independently motivated syntactic objects to PF, would hit upon the phonological component as an optimal solution. I have no idea whether this can be formulated as a sensible research task. As for option (29a), it might be argued that departures from "good design" are not surprising in this domain. Direct evidence about sound systems is available for language acquisition, permitting a degree of complexity and variation. And the subsystem reflects special properties of the sensorimotor systems, which are in a certain sense "extraneous" to language, relating to externalization by systems with nonlinguistic properties and capable of much variation while FL remains fixed, as in sign languages. Symbolic systems designed for special purposes (metamathematics, computers, etc.) dispense entirely with a phonological component, not facing the need to meet the legibility conditions for human language at the sensorimotor interface.

The strongest sustainable inclusiveness/interpretability requirement, then, is (33).

(33) Inclusiveness holds of narrow syntax, and each feature is interpreted at the level LF or associated with phonetic features by the phonological component.

The phonological component is generally assumed to be isolated in even stronger respects: there are *true* phonological features that are visible only to the phonological component and form a separate subsystem of FL, with its own special properties. Assume this to be true. Then in the course of construction of LF, an operation Spell-Out delivers the structure already formed to the phonological component, which converts it to PF. If LIs express Saussurean arbitrariness in the conventional way, then Spell-Out "strips away" the true phonological features, so that the derivation can converge at LF; it will crash if later operations introduce

LIs with phonological features. On the assumptions of Distributed Morphology, the phonological features are introduced after Spell-Out by phonological operations applying to LIs lacking them. I will assume some instantiation of this array of options to be correct.

Narrow syntax also involves devices that are imperfections unless shown to be unreal (option (29b), which again seems implausible) or to be motivated by design specifications (option (29c), the most interesting possibility). Consider two striking examples:

(34) a. Uninterpretable features of lexical items
b. The "dislocation" property

Under (34a) we find features that receive no interpretation at LF and need receive none at PF, hence violating any reasonable version of the Interpretability Condition (30b). The example that has played the most important role ever since Jean-Roger Vergnaud's famous unpublished letter written twenty years ago is structural Case. The picture is more complex for agreement features: semantically interpretable for nouns, but not for verbs/adjectives, and phonetically optional throughout. 70 External manifestation of inflectional features appears to be the locus of much of the variety languages display, a topic that has gained prominence within the P&P framework.

These observations presuppose that occurrences of features can be distinguished, a question raised earlier but put off: occurrences of agreement features are distinguished by category, some interpretable, some not. That falls well short of truly identifying occurrences. To do so would be necessary if feature chains exist.71 In the absence of clear evidence to the contrary, I will assume that feature chains do not exist, hence that features cannot move or be attracted.72

The dislocation property (34b) is another apparent imperfection. In (35), for example, the phrase *an unpopular candidate* is in the natural position to be interpreted as object of *elect* in (35a,b) but not in (35c,d), though the interpretation is in relevant respects the same.

(35) a. they [elected an unpopular candidate]
b. there was [elected an unpopular candidate]
c. an unpopular candidate was elected
d. there was an unpopular candidate elected

In (35c,d) the surface phonetic relations are dissociated from the semantic ones.73 Such phenomena are pervasive. They have to be accommodated

by some device in any adequate theory of language, whether it is called "transformational" or something else.

Dislocation of a yields a chain (a, *t*)more accurately, a chain {X, Y}, where X and Y are occurrences of a. The raised element typically c-commands its trace in the original position, but where true, that follows from independent properties of CHL. Further operations might lead to violation of c-command and of locality relations between the two positions, as in multiple head raising or independent XP-dislocation.

(36) a. [[Vj-T]i-C] [. . . *t*i . . . [VP . . . *t*j . . .]]
b. [written *t*j for children]i, [those books]j couldn't possibly be *t*i]

In (36) the indices are redundant, the chains being determined by constitution of the trace. That need not be so, however, as in (37).

(37) whomi did everyone talk to whomj about whomk

Chains can be formed with i = j or i = k, both consistent with locality conditions; and the full range of interpretations seems to be available, either way.

In the approach we are pursuing here, the chains at LF are determined by identity throughout, the ambiguity of (37) being resolved by the derivation, given the initial numeration. Other means would be required if we were to eliminate this device along lines discussed earlier. In a strict derivational approach semantic interpretation is cyclic and the problem of interpreting (36) and (37) arises in a somewhat different form. In representational approaches chains are determined by an algorithm A operating in a "search space" of possible LFs; the burden of accounting for locality and other conditions on chains then rests on A. Here we return to issues of general architecture raised in section 3.3.

However these matters are resolved, we have two "imperfections" to consider: uninterpretable features and the dislocation property. These properties (in fact, morphology altogether) are never built into special-purpose symbolic systems. We might suspect, then, that they have to do with externally imposed legibility conditions. With regard to dislocation, that has been suggested from the earliest days of modern generative grammar, with speculations about facilitation of processing (on the sound side) and the dissociation of "deep" and "surface" interpretive principles (on the meaning side). The boundaries are not clear, 74 nor are the mechanisms to express them. One approach to the array of problems was to distinguish the role of deep and surface structure (D- and S-Structure) in

semantic interpretation: the former enters into determining quasi-logical properties such as entailment and q-structure; the latter such properties such as topic-comment, presupposition, focus, specificity, new/old information, agentive force, and others that are often considered more discourse-oriented and appear to involve the "edge" of constructions. Theories of LF and other approaches sought to capture the distinctions in other ways. The "deep" (LF) properties are of the general kind found in language-like systems; the "surface" properties appear to be specific to human language. If the distinction is real, we would expect to find that language design marks it in some systematic wayperhaps by the dislocation property, at least in part.

To the extent that such ideas can be given substance, it would follow that the dislocation property is required; it falls within the design specifications given to the super-engineer seeking an optimal solution to conditions imposed by the external systems.

This line of argument might provide motivation for the dislocation property, but it would remain to find the mechanisms employed to implement it. The distinction is familiar. We may say that the function of the eye is to see, but it remains to determine the implementation (a particular protein in the lens that refracts light, etc.). Similarly, certain semantic properties may involve dislocated structures, but we want to discover the mechanisms that force dislocation. Minimalist intuitions lead us to look at the other major imperfection, the uninterpretable inflectional features. Perhaps these devices are used to yield the dislocation property. If so, then the two imperfections might reduce to one, the dislocation property. But the latter might itself be required by design specifications. That would be an optimal conclusion, falling under option (29c).

To establish any such conclusion is no simple matter. We are entering terrain that is mostly unexplored. One approach is suggested by the observation that for convergence, uninterpretable features must be deleted in the course of computation of LF. Consider the dislocated example (35c), repeated here.

(38) an unpopular candidate T-was elected *t*

There are three kinds of uninterpretable features in this structure: (a) the agreement features of T (taking them as a unit, the set of f-features), (b) the EPP-feature of T that requires "second Merge," and (c) the structural Case feature of *an unpopular candidate*. The f-set (a) identifies T as a target of dislocation; the EPP-feature (b) requires that something be merged in

this position; the Case feature (c) identifies *an unpopular candidate* as a candidate for such merger (hence dislocation). Successful implementation of the operation erases all of the uninterpretable features, forming MLIs with a reduced set of features. The approach is optimal, if indeed uninterpretable features are the mechanism for dislocation. 75

Assuming so, let us look more closely. Suppose that the derivation has constructed the SO (39), having merged T with the copula-headed phrase.

(39) T be elected an unpopular candidate

The new element T has uninterpretable features of two types: its f-set and its selectional feature EPP. Like other selectional features, EPP seeks an XP to merge with the category it heads. The f-set we can think of as a *probe* that seeks a *goal*, namely, "matching" features that establish agreement. The relation of the probe of T to its goal is the *T-associate* relation.

For the f-set of T in (39), there is only one choice of matching features: the f-set of *candidate*. Locating this goal, the probe erases under matching. Taking structural Case to be a reflex of an uninterpretable f-set,76 it too erases under matching with the probe. The erasure of uninterpretable features of probe and goal is the operation we called *Agree*. But the EPP-feature of T must also be satisfiedin this case by "pied-piping" of a phrase P(G) determined by the goal of T's probe, which merges with (39), becoming [Spec, T]. The combination of selection of P(G), Merge of P(G), and feature deletion under matching (Agree) is the composite operation Move, which dislocates *an unpopular candidate*, eliminating all uninterpretable features.

Matching is a relation that holds of a probe P and a goal G. Not every matching pair induces Agree. To do so, G must (at least) be in the *domain* D(P) of P and satisfy locality conditions. The simplest assumptions for the probe-goal system are shown in (40).

(40) a. Matching is feature identity.
b. D(P) is the sister of P.
c. Locality reduces to "closest c-command."

Thus, D(P) is the c-command domain of P, and a matching feature G is *closest to* P if there is no G' in D(P) matching P such that G is in D(G').

In the absence of evidence to the contrary, we adopt (40), with a qualification taken over from earlier work.77

(41) Terms of the same minimal domain are "equidistant" to probes.

The minimal domain of a head H is the set of terms immediately contained in projections of H.

With matching restricted to identity, Case and lexical category cannot enter into Agree or Move, since the probes do not manifest these features. And OS must be an implementation of (here invisible) object agreement, with ancillary Case checking.

If uninterpretable features serve to implement operations, we expect that it is structural Case that enables the closest goal G to select P(G) to satisfy the EPP by Merge. Thus, if structural Case has already been checked (deleted), the phrase P(G) is "frozen in place," unable to move further to satisfy the EPP in a higher position. More generally, uninterpretable features render the goal *active*, able to implement an operation: to select a phrase for Merge (pied-piping) or to delete the probe. The operations Agree and Move require a goal that is both local and active. 78

We therefore have the possibility of *defective intervention constraints* in a structure (42), where > is c-command, b and g match the probe a, but b is inactive so that the effects of matching are blocked.

(42) a > b > g

We will return to some illustrations.79

In *MP* Agree is analyzed in terms of feature movement (Attract) and a concept of matching that is left unclear. Here we take matching to be identity and dispense with Attract, with complications it induces about extended MLIs, feature chains, and other matters. Checking reduces to deletion under matching with an active local goal and ancillary deletion of the uninterpretable feature that rendered the goal active. I will use the terms *checking* and *attract* only for convenience.

Suppose that the EPP-feature of T could be satisfied more simply than by the full operation Move. That is the case in (35b), repeated here.

(43) there [a T-was elected an unpopular candidate]

Here the lexical array includes the expletive *there*. At stage a of the derivation (= (39)), the independent operations Agree and pure Merge suffice: Agree deletes the f-set of T and the structural Case of *candidate*, and Merge (of *there*) satisfies the EPP-feature of T. The more complex operation Move is preempted; dislocation does not take place, though we have long-distance agreement of T and its goal (its associate).80

Manifestation of structural Case depends on interpretable features of the probe: finite T (nominative), n (accusative), control T (null), on our

earlier assumptions. We may therefore regard structural Case as a single undifferentiated feature. The same would be expected for the uninterpretable f-set of the probe. Its manifestation depends on interpretable features (namely, f-features) of the goal, so that it too can be taken to be undifferentiated with respect to the value of the individual features of the f-set ([+/ plural], etc.). For both probe and goal, the form of the uninterpretable features is determined by Agree. To rephrase in traditional terms, verbs agree with nouns, not conversely, and Case is assigned.

We therefore understand "feature identity" in (40a) to be identity of the choice of feature, not of value. More important, defective intervention effects are induced whether or not b and g of (42) are identical in f-feature value (singular blocks plural agreement, etc.). This lends theory-internal support to the earlier observation that f-features are interpretable only for N; their value is specified only in this case. Notice also that only the most underspecified element, PRO, can have null Case, so raising of a ¹ PRO to [Spec, T] causes the derivation to crash when T is a control infinitival.

We take deletion to be a "one fell swoop" operation, dealing with the f-set as a unit. Its features cannot selectively delete: either all delete, or none. The f-features of T do not agree with different NPs, for example. In the same spirit, we assume that only a probe with a full complement of f-features is capable of deleting the feature that activates the matched goal. Suppose that the probe for participial (like adjectival) a is a f-set lacking the feature [person] and that G is the closest matching goal in its search space: P(G) = DP may be attracted to [Spec, a], deleting the probe of a (participial agreement), but the operation will not delete structural Case in DP, which can move on to [Spec, T], deleting the probe of T and the Case of DP (subject agreement). n and nondefective T, with a full complement of f-features, delete the uninterpretable feature that activates the matched goal (raised or not). 81

How would noncontrol infinitivals (Tdef) and weak expletives Expl of the *there*-type fit into this picture? The former category falls into place if T always has at least a minimal feature complement, perhaps only [person] for Tdef. If so, Move of a to [Spec, Tdef] will delete the f-set of T (= uninterpretable [person]) but not the structural Case feature of a, so that a can undergo further movement and agreement. The phase head n/C have no counterpart to Tdef with a reduced f-set and therefore do not provide an "escape hatch" for successive-cyclic A-movement.

Weak Expl shares the basic movement/attraction properties of nominals. That is expected if Expl has an uninterpretable feature F that activates it until erased and a f-set G that matches a probe in T. But G is uninterpretable for Expl, so a distinct F is unnecessary, obviating the need for structural Case in Expl. The composition of G is determined by two conditions: (a) Expl can raise to [Spec, Tdef]; (b) Expl cannot delete the probe of nondefective T. Condition (a) requires that G contain a feature to match the probe of Tdef ([person], if what precedes is correct). From (b) it follows that G must be less than a full f-set, hence optimally just [person]. That (b) holds is shown by long-distance agreement structures such as (44b d), (44b) surfacing commonly as (44c), or in English more naturally as (44d), as noted.

(44) a. they declared [three men guilty]
b. there were declared [three men guilty]
c. there were declared guilty three men
d. there were three men declared guilty

If the matching feature of the probe were deleted by the operation, it would not be available for associate matching and the nominative Case of the associate would remain unchecked because of the lack of a full complement of features in T (compare participial agreement). The problem does not arise if (b) holds and uninterpretable features delete in an "all or none" fashion, not selectively. In (44b d) the full complement of f-features of T deletes the uninterpretable feature G of *there*, barring further raising. When Expl raises to [Spec, Tdef], the probe (a single feature) deletes under matching as before, but G does not, because deletion requires matching with a full complement of f-features of the probe. Therefore, successive-cyclic raising through [Spec, Tdef] is possible. 82

Reinterpretation of Attract in terms of Agree eliminates the need to introduce "checking domains." That is a step forward. The notion is complex, and furthermore unnatural in minimalist terms; feature checking should involve features, nothing more, and there is no simpler relation than identity. More important, the notion is irrelevant for the core cases: elements merge in checking domains for reasons independent of feature checking; and feature checking takes place without dislocation to a checking domain.

As discussed, both properties are illustrated in expletive constructions. Much work on the topic has taken long-distance effects to be a property

of these constructions, hence of an expletive-associate relation; various ideas have been explored regarding how that relation is established. In *MP*, chap. 4, a different approach is suggested: the long-distance effects are attributed to a T-associate relation that involves features only and is independent of the expletive. The reasons were theory-internal, but a broader range of cases adds empirical support. Long-distance effects are found without expletives in such constructions as (15a), 83 the EPP being satisfied by raising of quirky Case; and expletive subjects are found without T-associate agreement when there is no accessible nominative. We will return to some illustrations. The general conclusions are these:

(45) a. Long-distance agreement is a T-associate (probe-goal) relation.
b. The EPP can be satisfied by
i. Merge of expletive
ii. Merge of associate
iii. Merge of a closer to T than the associate

Case (45bi) is illustrated by T-associate agreement, with the definiteness effect. Case (45bii) exhibits agreement of [Spec, T] and T, but that is ancillary to the T-associate relation. In case (45biii) there is no definiteness effect and long-distance T-associate agreement holds with embedded accessible nominative; or, if such an associate is lacking, T is default.

More generally, we should not expect Spec-head relations to have any special status. Within bare phrase structure, we cannot, for example, take the result of first Merge to a to be sometimes a specifier and sometimes a complement, as in an X-bar-theoretic analysis that takes the object of a to be its complement (*see John, proud [of John]*) but the subject of objectless a to be its specifier (base forms of *John eat, John proud*). The restriction to a single specifier is also questionable: rather, we would expect first Merge, second Merge, and so on, with no stipulated limit.

3.6 Syntactic Objects

We are taking the elements that enter into derivations to be features and objects constructed from them in a restricted way:

(46) a. Lexical items LI
b. Modified lexical items MLI
c. Sets K constructed from given elements a, b

An MLI is an LI with uninterpretable features deleted. In case (46c) K corresponds to the subtree dominated by a node in a standard phrase

structure diagram; there are no objects corresponding to the nodes, and no nonbranching projections.

q-structure and similar semantic roles are based on pure Merge of XP to substantive LIs or their projections. Checking theory, in contrast, involves uninterpretable features of functional categories and is reduced to feature matching under conditions (40a c). The two theories differ in nature and implementationplausibly, since they are conceptually quite distinct. Something like q-theory is a property of any language-like system, whereas checking theory is specific to human language, motivated (we are speculating) by legibility conditions. The Chain Condition is an expression of this duality and should fall out as a descriptive observation, along with other properties of chains. 84 Apart from Merge of selected XP, narrow syntax involves only feature deletion to form reduced MLIs, sometimes associated with Merge in $\bar{\theta}$ (non-q) positions yielding dislocation.

According to this conception, agreement (hence movement) is driven by uninterpretable features of the probe, which must be deleted for legibility. The operation Greed of *MP*, in contrast, was driven by uninterpretable features of the goal. The principle we are now entertaining is what Lasnik (1995a,b) calls Enlightened Self-Interest, with the further requirement that the matched probe delete: we might call the principle Suicidal Greed. Suicidal Greed does not have the "look-ahead" property of Greed, a complexity reduction that could be significant, as discussed.

With this shift of perspective, structural Case is demoted in significance. The Case Filter still functions indirectly in the manner of Vergnaud's original proposal, to determine the distribution of noun phrases. But what matters primarily are the probes, including f-features of T, n. That reverses much of the recent history of inquiry into these topics and also brings out more clearly the question of why Case exists at all. The question arises still more sharply if matching is just identity, so that Case can never be attracted; operations are not induced by Case-checking requirements. Recall that lexical category also cannot be attracted and does not induce operations, raising the same questions.85 For Case, a plausible answer is the one already suggested: uninterpretable features activate the goal of a probe, allowing it to implement some operation (Agree or Move).86 It follows that after structural Case of DP is deleted, the phrase cannot move further to an A position and its f-set cannot induce deletion (though it is still "visible" to a probe, allowing defective intervention effects as in (42)). Suppose quirky Case is (q-related) inherent Case with an additional structural Case feature, as often suggested in one or another

form. Then it too is immobile once it reaches a Case-checking position. 87 If the f-features of T that check the structural Case of raised quirky Case themselves delete, we have default T; if they remain, we have remote agreement with some lower accessible nominative.

The descriptive observations seem generally accurate. In these terms, the visibility thesis and the Chain Condition reduce to LF convergence.

In his detailed review of Icelandic agreement, Sigurðsson (1996) concludes that remote nominative Case allows number agreement but not first/second person agreement. That would follow if the [person] feature of T reduces to [3person] (the default choice) when it attracts quirky Case or Expl to [Spec, T].88

Suppose expletive Expl is merged in [Spec, T] without movement. Agreement is manifested for *it*-type Expl with a full complement of f-features, and *there*-type Expl becomes inactive, indicating that its uninterpretable feature deletes ([person], I have assumed). But these are properties of Agree, not Merge.89 The results are expected if Expl is an X0 head and its [person] feature is uninterpretable, therefore able to probe its domain T' (= D(Expl)), locating the f-set of T as the closest goal. The uninterpretable probe deletes, and the f-set of T as well if Expl has a full f-set. We therefore have agreement but not via merger: full or partial agreement depending on the f-set of Expl, which becomes inactive.90

The head of an A-chain can undergo Ā-movement, of course, with different features accessed. Take *wh*-movement. This would be point by point analogous to A-movement if the *wh*-phrase has an uninterpretable feature [wh] and an interpretable feature [Q], which matches the uninterpretable probe [Q] of a complementizer in the final stage; successive cyclicity could then function in the manner discussed.91 The *wh*-phrase is active until [wh] is checked and deleted. The *Wh*-Island Constraint is then a defective intervention effect of the type (42): [Q] of the already checked *wh*-phrase (b in (42)) bars attraction of lower [Q], but cannot move or check the uninterpretable feature of the probe. A possible analysis of *wh*-in-situ constructions is that [wh] pied-pipes only the head (overtly or covertly).92

The reasoning extends to such constructions as (47a d) ((47b) based on OS).

(47) a. *[John to seem [*t*I is intelligent]] (would be surprising)
b. *(we hoped) [PRO to be decided [*t*I to be killed at dawn]]

c. *[DO this book] seem [*t*DO to read [*t*DOI [never [[Subj any students] *t*read]]]]
d. *there seem [a[Subj several people]I are [Pred friends of yours]]

The EPP is satisfied throughout, and if local matching sufficed for agreement, the expressions should converge with uninterpretable features deleted. Appeal to such principles as "maximal checking" would not make the proper distinctions locally (e.g., barring (47c) in favor of subject raising). But in all cases the position superscripted *I* is inactive, hence unable to raise (47a c) or to delete the features of a matched probe (47d). Case (47d) illustrates a defective intervention effect of type (42): Subj is visible (barring Pred as goal) but inactive, unable to establish agreement with matrix T. 93 The same property holds in (48).

(48) a. *there were decided [a PRO to stay with friends]
b. *XP T-seem that [a it was told friends CP]

PRO and *it* are inactive, their structural Case feature having been checked and deleted in a. But their f-features remain visible and block association of matrix T to *friends*, both of which therefore retain uninterpretable features. Case (48b) is therefore barred with pure Merge of Expl, or raising of *it* or *friends* ("superraising").94

We are now in a position to derive the basic structural properties of CFCs ((5), extended to (50)), in the configuration (49).

(49) a = [XP [(EA) H YP]]

(50) a. If H is n/C, XP is not introduced by pure Merge.
b. In the configuration [b Hb . . . a]. Hb a CFC and b minimal,
i. if Ha is C, Hb is independent of a;
ii. if Ha is n, Hb = Tb agrees with EA, which may raise to [Spec, Tb] though XP cannot;
iii. if Ha is Tdef, if Hb is T then XP raises to [Spec, Tb] if there is no closer candidate g for raising; and if Hb is n then XP agrees with n (as may a lower associate if XP = Expl).

(50a) has already been dealt with (see discussion following (24)), so we can turn to (50b).

In case (50bi) a = [XP [C TP]], TP headed by nondefective Ta. We can limit attention to Ta with its f-set deleted; were the f-set not deleted, the derivation would have crashed at a. Hence, agreement is fully established with the closest associate Assoc that matched Ta; Assoc is either raised to

[Spec, Ta] or remains in situ in a long-distance agreement relation, its structural Case feature deleted in either case. Any structural Case feature remaining in a is inaccessible by virtue of the defective intervention effect induced by Assoc, again causing crash detectable at a. Case (50bi) follows: the f-set and EPP-feature of Hb have to be satisfied independently of a.

Consider case (50bii), with a = [XP [EA [n YP]]. XP is raised from within YP, checking and deleting the f-set of n (object agreement) and its own structural Case feature. XP is therefore inactive for A-movement and cannot check the f-features of Tb. These can only be deleted by agreement with EA, deleting its structural Case feature as well, with EA either raising to [Spec, Tb] or remaining in situ. Recall that XP does not induce a defective intervention effect that would bar the Tb-EA relation; see (41). 95

In case (50biii) a = [XP [Tdef YP]]. Whether merged (hence Expl) or raised, XP has an uninterpretable feature that renders it active for the Case/agreement system and must be checked outside a by Hb that is either n (ECM) or T (raising). Suppose Hb is n. If XP = Expl, its uninterpretable feature [person] deletes but the f-set of n remains, able to check the Case (accusative) of a lower associate (*I expect [there to be a proof discovered]*); there is no defective intervention effect. If XP is an argument, it is the associate of n and object agreement with (accusative) Case checking proceeds as before. In either case XP raises to [Spec, n] only when n has an EPP-feature (see note 95). If Hb is T, XP raises to [Spec, Tb] unless barred by a closer candidate g. If Tb is nondefective, it either agrees with XP and checks its Case (nominative) or shifts to default, as already discussed. If Tb too is defective, then XP raised to [Spec, Tb] will have to be associated with a still higher T or n, by raising or agreement in situ.

The basic properties (50) of CFCs therefore follow from simple and plausible assumptions. Without running through cases, it should be clear that the Phase-Impenetrability Condition (21) holds for A-movement for the same reasons (see note 47).

Further insight into these matters should derive from raising constructions with quirky Case moving to matrix subject, as in (51).96

(51) a. me(dat) thought(pl) [*t*me [they(pl,nom) be industrious]]
b. *me(dat) seem(pl) [*t*me [John(dat) to like horses(pl,nom)]]
c. *John seems(sg) me(dat) [*t*John to like horses]

The matrix verb agrees with the embedded nominative in (51a), but not in (51b), which requires default inflection because of the defective interven-

tion effect: the f-features of *John* block the T-associate relation between T-*seem* and nominative *horses*. In (51a) as well a phrase with f-features intervenes between matrix T and nominative, namely, the trace of the quirky dative *me*. But the latter is not the head of an A-chain, unlike in (51b) (also (47d) and (48)). Case (51c) is blocked by locality; quirky dative, with structural Case, is accessible. The conclusions are as before, but sharpened: it is only the head of the A-chain that blocks matching under the locality condition (40c). A-movement traces are "invisible" to the probe-associate relation; or from another perspective, the A-chain itself (regarded as a set of occurrences of a) constitutes the barrier.

This account relies on restricting basic operations to Merge and Agree, based on feature matching (identity) and driven by Suicidal Greed. Suppose that there is a Move a operation dissociated from matching and dependent on properties of a, with locality expressed as "Shortest Movement." In the defective intervention cases, matrix T is the closest target that can check the Case feature of the potential associate a. With Move a available and constrained only by locality, the derivation converges (incorrectly) with raising or agreement of a. Further constraints are required, then, if this option exists. 97

A crucial property of deletion is that a deleted feature is invisible at LF and inaccessible to CHL (the [±active] property), but accessible to the phonological component. This property poses a problem on the assumption made in *MP* that Spell-Out applies at a single point in a derivation: pre-Spell-Out, the probe must delete when checked yet remain until Spell-Out.98 The natural conclusion is that Spell-Out is associated with agreement. Deleted features are literally erased, but only after they are sent to the phonological component along with the rest of the structure Spossibly at the phase level. Spell-Out therefore applies cyclically in the course of the (narrow syntactic) derivation. I will assume that this approach, apparently the simplest and most principled one, is correct.99

The single Spell-Out thesis of *MP* retains the flavor of the Extended Standard Theory model, distinguishing overt from covert operationspre- and post-Spell-Out, respectively. If both overt and covert operations are cyclic, then there are two independent cyclesand if operations of the phonological component are cyclic, a third cycle as well. With cyclic Spell-Out, contingent on feature-checking operations, these distinctions collapse. There is a single cycle; all operations are cyclic. Within narrow syntax, operations that have or lack phonetic effects are interspersed. There is no distinct LF component within narrow syntax, and we can dispense

with troublesome questions about its apparently cyclic character. Agree alone, not combined with Merge in the operation Move, can precede overt operations, contrary to the assumptions made in *MP* and related work. Crucial cases are long-distance agreement, *wh*-in-situ, and others. Many questions arise, but they do not seem obviously unanswerable.

One question has to do with identification of chains by identity under cyclic Spell-Out. Suppose the uninterpretable features of a head H are checked at phase Ph with Ph transferred to the phonological component and H reduced to H', which is transferred at a later phase of derivation. We want to take <H', H> to be a chain, so that H will be unpronounced by general principles, but H, H' are not identical. We may therefore take chains to be determined by a relation of nondistinctness holding between (a, b) if they differ only in uninterpretable features. 100

The principle Procrastinate is no longer formulable (at least, as before), eliminating another case of look-ahead. The concept of strength, introduced to force violation of Procrastinate, appears to have no place. It remains to determine whether the effects can be fully captured in minimalist terms or remain as true imperfections.

To implement the program in a suitably Spartan fashion, we seek to restrict CHL to indispensable operations that satisfy minimalist conditions. There are two candidates: Merge, indispensable in some form, and Agree, which, we speculate, might ultimately be accounted for in terms of design specifications. We have to determine how these operations apply, seeking and questioning departures from optimal design.

The operations Merge and Agree must

(52) a. Find syntactic objects to which they apply,
b. Find a feature F that drives the operation,
c. Perform the operation, constructing a new object K.

An operation Op takes objects already constructed (perhaps in the lexicon) and forms from them a new object. Condition (52a) is optimally satisfied if Op applies to full syntactic objects already constructed, with no searchthat is, if CHL operates cyclically. It follows that derivations meet the condition (53).

(53) Properties of the probe/selector a must be satisfied before new elements of the lexical subarray are accessed to drive further operations.

If the properties of a are not satisfied, the derivation crashes because a can no longer be accessed. We continue to assume cyclic application of all

operations, returning to some ambiguities. Let us turn now to condition (52c).

The operation Merge forms K from a, b. Minimally, K should consist only of a and b, so K = {a, b}. More information is needed about K, however: its category (its *label*) and the nature of the merger, either substitution or adjunctionthe former at least not entering into narrow syntax, on the sparest assumptions, but needed for the phonological component and LF interpretation. 101 To ensure that every category has a label, let us say that label(a) = a, for a an LI.

Adjunction has an inherent asymmetry: X is adjoined to Y. Exploiting that property, let us take the distinction between substitution and adjunction to be the (minimal) distinction between the set {a, b} and the ordered pair <a, b>, a adjoined to b. The constructed objects K, then, are of the form {g, {a, b}} (substitution) or {g, <a, b>} (adjunction), where g is the label of K.

The term *substitution*, adapted from earlier work, is misleading within this framework. For clarity, let us refer to substitution as *Set-Merge* and adjunction as *Pair-Merge*. I will put aside here a number of questions that arise about adjunction.102

On minimal assumptions, the label g should be the label of either a or b. Hence, no matter how complex the object constructed, its label is an LI, the head selected from the lexicon that has "projected" through the derivation, or a reduced MLI. If the label is determined from a, b by general principles, then the result of merger of a, b is simply {a, b} or <a, b>.

Are labels predictable?

Consider pure Merge. There are two cases: set-Merge and Pair-Merge. The latter adjoins a to b to form <a, b>. Given the asymmetry, it is natural to conclude that the adjoined element a leaves the category type unchanged: the target b projects. Hence, adjunction of a to b forms K = {g, <a, b>}, where g is the label of b. Eliminating redundancy, the operation forms K = <a, b>.

As an operation, Set-Merge is symmetric, so one might expect either label to project. If so, the outcome would either be interpretable at LF or not. But here too properties of language design appear to determine the label without look-ahead. Set-Merge typically has an inherent asymmetry. When a, b merge, it is to satisfy (selectional) requirements of one (the *selector*) but not both. Fairly generally, furthermore, the selector is uniquely determined for a pair (a, b), as can be seen by reviewing cases.

Set-Merge of (a, b) has some of the properties of Agree: a feature F of one of the merged elements (say, a) must be satisfied for the operation to take place. Furthermore, F is in the label of a, hence detectable in an optimal way (satisfying condition (52b)). The selector F for Merge is analogous to the probe for Agree. Furthermore, F is the only element of a that enters into the operation, hence the only one available without further complication to determine the label of the merged elements. In this case too, then, the label is predictable and need not be indicated: the label of the selector projects.

The intuitive content of the Projection Principle is that for a substantive category a, the selector F is a semantic property of a, an interpretable feature. Hence, F does not delete (another difference between q-theory and checking theory). Further specification depends on how q-theory is understood. Take, say, transitivity of a verb V. If the property is implemented in terms of q-grids, then a feature of V selects the object. If it is implemented configurationally as a structure [n n [V V . . .]], then n too is a relevant selector and the n-VP structure enters into the interpretation. In either case one or another form of deviance (or crash) results from inappropriate merger. These and many other questions come to the fore as the framework is more carefully articulated.

The two cases of pure Merge differ in several respects. The asymmetrical operation Pair-Merge has no selector and is optional; the symmetrical operation Set-Merge has a selector (typically unique) and is obligatory. 103 In these respects language design is close to optimal, providing just the information necessary for an operation Op to project the label L: L is determined by Op itself if Op is asymmetrical, but a selector is needed to determine L if Op is symmetrical. Accordingly, Merge has a selector for Set-Merge but not Pair-Merge, which is therefore optional. The label is determined without look-ahead to check eventual convergence (contrary to *MP*), another case confirming conclusions about complexity suggested earlier.

Pure Merge, then, satisfies the conditions of (52) rather well. Consider the second elementary operation, Agree. We are now assuming cyclicity: a is a candidate for the operation only if it is the full syntactic object under inspection, so search is unnecessary, satisfying (52a). Agree requires a probe F in a. By condition (52b), F has to be readily detectable, hence optimally in the label L(a) of a, its sole designated element. F seeks a matching feature F'. Deletion takes place under conditions already dis-

cussed. The syntactic object a is otherwise unchanged. Its label remains as before. The same properties carry over to Move, constructed from Agree and Merge. 104

In all cases, then, the label is redundant. The syntactic objects are LIs, or sets {a, b} or <a, b> constructed from them. The label is determined and available for operations within CHL or for interpretation at the interface, but is indicated only for convenience.

Computation is driven by a probe/selector of a label, which projects. Hence, no operation can be contingent on application of earlier ones. That seems to be a valid descriptive generalization, which falls into place. Both label determination and operations are "first-order Markovian," requiring no information about earlier stages of derivation.

Conditions (52a) and (52b) are optimally satisfied, as is (52c) for Merge and partially (so far) for Agree. Continuing with Agree, (52c) requires that the matched goal G must be easily located. We want to identify a domain D(P) of the probe P, such that G is within D(P). There are two candidates for D(P): the smallest and the largest of the categories labeled by the label containing P. The former includes only the complement of P; the latter its specifiers as well. Search space is more limited if D(P) is the smallest category, as so far assumed; see (40b). It is restricted further by the "closest match" condition (40c). Natural complexity/economy conditions are again satisfied, along the lines of (3), (14), and their extensions; there is substantial evidence that G must be in the complement of the probe P, not its specifiers, and that locality conditions enter into choice of G.

The basic operations Merge and Agree satisfy reasonable "good design" conditions (52). The conclusions extend to the third operation, Move, insofar as it is constructed from the basic operations.

Move of b, targeting a, has three components.

(54) a. A probe P in the label L of a locates the closest matching G in its domain.
b. A feature G' of the label containing G selects a phrase b as a candidate for "pied-piping."
c. b is merged to a category K.

P and G' are uninterpretable. P deletes if G is active (Suicidal Greed). G' also deletes, but it cannot delete in step (54a) before carrying out its function in step (54b). There may be reasons to suppose that G cannot delete before step (54c), but I will defer the matter.

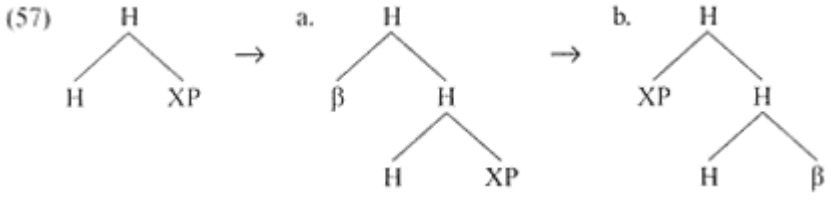
The new object K formed by Merge of b to a retains the label L of a, which projects. There are two reasonable possibilities, illustrating the ambiguity of cyclicity mentioned earlier:

(55) a. a is unchanged.
b. b is as close to L as possible.

Suppose we have the LI H with selectional feature F, and XP satisfying F. Then first Merge yields a = {XP, H}, with label H. Suppose we proceed to second Merge, merging b to a. In this case b is either extracted from XP (Move) or is a distinct syntactic object (pure Merge). There are two possible outcomes, depending on choice of K in (55).

(56) a. {b, {XP, H}} (as in (55a))
b. {XP, {b, H}} (as in (55b))

In tree notation the equivalents are (57a, b), respectively.



The desired outcome is (57a), not (57b); that has always been assumed without discussion. Thus, if H = T, XP = VP, and b is either an expletive merged to TP or a DP raised from XP, the result of Merge would be (58a), with b the Spec of T and VP remaining its complement, not (58b), with b becoming the complement of T and VP its Spec.

(58) a. [T b [T VP]]
b. [T VP [T b]]

But the reasons are not entirely obvious. Each outcome satisfies a reasonable condition: (58a) satisfies the familiar *Extension Condition* (55a); (58b) satisfies the condition of *Local Merge* (55b).

One possibility is to stipulate that the Extension Condition always holds: operations preserve existing structure. Weaker assumptions suffice to bar (57b) but still allow Local Merge under other conditions. Suppose that operations do not tamper with the basic relations involving the label that projects: the relations provided by Merge and composition, the relevant ones here being sisterhood and c-command. 105 Derivations then

observe the condition (59), a kind of economy condition, where R is a relevant basic relation.

(59) Given a choice of operations applying to a and projecting its label L, select one that preserves R(L, g).

In the case of (56) (57) the principle (59) selects (a) over (b). Basic relations of H are preserved in (a), but not in (b), which modifies sisterhood and c-command for H (in fact, the category a = {XP, H} disappears altogether).

One case permitted under (59) but barred by strict adherence to the Extension Condition is head adjunction. The standard assumption is that in this case Local Merge takes precedence over the Extension Condition. Thus, in (56) (57), (b) is chosen over (a) for b a head H'. Adherence to the Extension Condition would change the status of H' to an XP; V-raising to T, for example, would create a VP-TP structure, with VP the Spec of T, contrary to intention. The uniformity condition for chains (*MP*, sec. 4.1, (17)) bars that choice, leaving as the only option head adjunction yielding (b), in violation of the Extension Condition. (59) is satisfied vacuously. 106 Head adjunction thus provides some reason to weaken the Extension Condition as proposed, permitting Local Merge if (59) is satisfied.

For XP-merger, (59) eliminates the ambiguity of choice of K for second Merge, imposing the Extension Condition (55a) rather than Local Merge (55b). But the choice remains open for third Merge. Suppose we have the outcome (a) of (56) (57), as required, and now merge g to it (either by pure Merge, or with g extracted from XP). The two possible outcomes are (60a) (satisfying the Extension Condition) or (60b) (satisfying Local Merge).

(60) a. {g, {b, {XP, H}}}
b. {b, {g, {XP, H}}}

In (60a) the newly merged phrase g is the outer Spec; in (60b) it is the inner Spec. Either way, sisterhood and c-command are preserved for H, satisfying condition (59). The discussion so far has kept to the Extension Condition (55a) (e.g., (8b)). The question becomes more intricate when we cast a wider net, another issue deferred here.107

Both Local Merge and the Extension Condition yield cyclicity and hence satisfy the optimality condition (52a), but in slightly different ways. The distinction brings out an empirical difference between the derivational definition of c-command and the definition given earlier in terms of

composition of elementary relations (see note 45; also Collins 1997, 84). Combined with Kayne's Linear Correspondence Axiom (Kayne 1994), as in the work cited, the derivational approach entails that the Extension Condition is inviolable. The compositional approach still allows the option of Local Merge, as in head movement and Merge in inner Spec.

The steps toward minimalist goals discussed above improve the *MP* theory in other respects. It was argued there that Merge is preferred over Move, but on dubious grounds. Now the preference is immediate: Agree and Merge are each components of Move, so it is a simple matter of more versus less. In fact, because of the extra component (b) of (54), Move is more complex than combined application of Agree and Merge. Thus, if a derivation D has reached the stage (61) and an expletive is available, D must proceed to (62a) (Merge combined with Agree, the latter to establish the T-associate relation), not (62b) (Move, then backtracking to the alternative (62a) when the derivation crashes with an unused expletive).

(61) [TP T be [a proof discovered]]

(62) a. there was a proof discovered
b. a proof was discovered

Again, a look-ahead property disappears. This sharpens the analysis of (10a)/(12a): the desired outcome is not just an option, but in fact the only option at the stage (61) of the cycle.

Suppose we have reached the stage (61) and the only unused element of the lexical subarray is the complementizer C. Since Merge is preferred to Move, the *MP* theory wrongly predicts merger of C and TP, barring dislocation of *a proof* to the subject position to yield (62b) (Eduardo Raposo, personal communication). The problem is overcome when we recognize the asymmetry of Merge. T contains no selector that allows C to merge to TP. With selectional features of T satisfied in the strict cyclic derivation of full TP (see (53)), the derivation can move on to C, which selects TP, projecting C.

In *MP* it is speculated that categories lacking interpretable features should be disallowedspecifically Agr, consisting only of uninterpretable f-features (sec. 4.10). That conclusion is forced in this version. 108 Suppose a is an LI that consists of uninterpretable features only and selects b, yielding the syntactic object K = {a, b} with label a. In the course of a convergent derivation, a will disappear, leaving K and higher projections of a without a label. But terms without labels are not well-formed syn-

tactic objects. Accordingly such elements as Agr not only *might* not exist, but *cannot* exist, on rather plausible assumptions. The argument carries over to other cases, among them semantically null determiners Dnull. If true D relates to referentiality/specificity in some sense, then an indefinite nonspecific nominal phrase (*a lot of people*, *someone* that enters into scopal interactions, etc.) must be a pure NP, not DP with Dnull (and the EPP cannot be stated as a D-feature). Notice that the argument holds only for heads a that select b. If a enters a derivation by Merge to a selector, deletion of a leaves a legitimate object (thus expletives might, in principle, delete completely).

I have explored a number of lines of argument in the attempt to refine and improve the Minimalist Program, including (a) restriction of basic operations to Merge and Agree (the latter reducing to deletion of uninterpreted matched features), each satisfying conditions of "good design," each preempting the more complex operation Move; and (b) restriction of access to the features F provided by UG in successive steps ((3) and its extension to cyclic derivation based on lexical subarrays). Basic properties of CFCs are then accommodated. These and other cases discussed reduce operative complexity in a natural way, reinforcing the suspicion that there is something to the curious principles in (28) and the intuitions and architectural conceptions on which they are based. Basic relations are restricted to those provided by Merge and composition. A number of other ways of overcoming errors and defects of *MP* and advancing the project of *MP* and related work have been outlined. I have raised, but surely not solved, the problem of whether the most striking apparent imperfections of narrow syntax, violation of the Interpretability Condition and dislocation, are true imperfections or are reasonable ways of satisfying design conditions, perhaps with uninterpretable features serving as the mechanism to induce structural properties required by interpretive systems at the interface. The discussion has been largely theoretical, with only a few indications about how things turn out on "best possible" expectations. Many questions have been left dangling, innumerable others unmentioned. I hope to return to some of these topics in the continuation.

Notes

1. As discussed in the introduction of *MP*, the chapters are largely based on lecture-seminars at MIT, the last of them in fall 1994. What follows draws from discussions during the lecture-seminars of fall 1995, 1997. I will make no attempt to review the impressive range of recent work that bears directly on questions that

arise, a failure that leaves no slight lacunae, as does the failure to consider alternatives that have been developed (see, among others, Abraham et al. 1996, Brody 1995, Collins 1997, Epstein et al. 1998, Frampton and Guttmann 1998, Sportiche 1995, Zwart 1996). For comments on an earlier draft, many of them incorporated here, I am particularly indebted to Zeljko Boskovic, * Chris Collins, Sam Epstein, Howard Lasnik, and Juan Uriagereka.

2. See, for example, Hermer and Spelke 1996. More generally, I assume that mental capacities are "modular" in the sense of Chomsky 1975, with "learning theories" LT(O, D) that may vary for organism O and cognitive domain D. The resulting modules might then have input/output properties of the kind analyzed in Fodor 1983, while belonging to a "central" system more structured than Fodor assumes.

3. For some illustrations and discussion, see Jenkins 1997, 1999, Marcus 1998.

4. Which implies virtually nothing about the novelty of its component elements.

5. I assume here familiar idealizations, abstracting away from real-world interactions that yield complex and widely varying forms of multiple systems. The term *idealization* sometimes misleads; the process is a crucial part of the effort to determine reality.

6. As systems, that is; their components need not be. See note 4.

7. Processing systems vary with languages and language types, even for very young infants, enabling them to sort out distinct languages in the data to which they are exposed. See Bosch and Sebastián-Gallés 1997. Jusczyk 1997, Mehler and Dupoux 1994. Whether these important discoveries (which add new dimensions to "poverty of the stimulus" arguments) are consistent with the simplifying assumption depends on how sensorimotor processing is "modulated" by the target language. Whorfian ideas on the meaning side have a similar flavor. See Phillips 1996 for an intriguing approach that bears on some of these questions.

8. I am assuming here the basic framework of Chomsky 1955, though of course there have been radical changes since. Levels are systems of representations: representations formed in the course of derivation typically do not form part of a level. Note further that the term *representation* is a technical one, with no "representation" relation in the sense of representational theories of ideas, for example.

9. On my own views regarding these issues, see Chomsky 1975, 1995a, 1996, among others.

10. Some disagree, regarding the issues as problematic. See, for example. Carr 1997, and for comment on some related matters, George 1996. On features as "instructions" for vocal gestures, see Halle 1983.

11. L an I-language in the technical sense, here and below. One simplifying assumption is that L is literally deducible from a choice of parameter values and lexicon, so that acquisition is "as if instantaneous." That need not be the case (e.g., in the theory of acquisition proposed in Locke 1997). It therefore becomes interesting to ask how close to true the assumption is. None of this has anything to do with the existence of a "language acquisition device" (LAD). LAD is just S0,

under a particular construal, including whatever properties of S0 may manifest themselves in the course of development. Postulation of LAD is often described as questionable or wrong, but that can hardly be so, at least if language is an identifiable component of human cognitive structure in any respect.

12. The question of a perfect language, whether designed by God or humans, is of course an old one, but completely distinct. Note further that the question of optimal design has nothing to do with the issue of "best theory" for FL (however intricate and "imperfect" the design of the system).

13. It is a misunderstanding to contrast "minimalism and X," where X is some theoretical conception (Optimality Theory, Lexicalism, etc.). X may be pursued with minimalist goals, or not.

14. See Belletti 1990 and much subsequent work.

15. Complications can readily be added. Little is known about evolution of higher mental faculties, and it is not clear how much can be learned within the limits of contemporary understanding. For a skeptical appraisal, see Lewontin 1990, 1998; and for critical analysis of recent efforts, see Berwick 1997, Jenkins 1997, 1999, Orr 1997.

16. These are called "bare output conditions" in *MP*, "output" because they are conditions on interface levels (hence "outputs" on a derivational approach), "bare" to distinguish them from filters, ranked constraints, and other devices that are part of the computational system itself.

17. For significance, we might assume further that there is no (nonarbitrary) bound on the number of legible expressions. Note that FL satisfying this minimal condition mightand the real system in fact doespermit generation of expressions that are unusable (structure of memory, garden path, etc.).

18. Interpretability is not to be confused with intelligibility. A convergent expression may be complete gibberish, or unusable by performance systems for various reasons. See note 17. And performance systems typically assign interpretation to nonconvergent expressions.

19. Convergence is defined in terms of properties of the external systems; the concept is clear insofar as these properties are clear. Many questions arise about its role in interpretation of deviance and in economy conditionsspecifically, does crash "free up" alternative derivations, as assumed in *MP* but not as these notions are elaborated elsewhere (e.g., Collins 1997, and in what follows)?

20. Also, well beyond the minimal sound-meaning connection given by initial assembly of features in the lexicon, an inescapable step in fixing a language for familiar reasons.

21. I put aside a variant that restricts "linguistic evidence" to identification of "well-formed" ("grammatical") expressions, so that the linguist then faces the alleged problem of selecting among grammars that are extensionally equivalent over these objects. Such demands inherit (and, by the radical restriction of evidence, amplify) the incoherence of the other approaches, adding the further difficulty of deciding what this property might be, for natural language.

22. There are many such debates, often with an oddly one-sided character: criticism of a largely unspecified position, with no defense of it on the part of those who are alleged to hold it but who in fact do not see what the issue is. Examples include the "innateness hypothesis," "autonomy of syntax," and "formalist" approaches. For an effort to find some significance in the "functionalist-formalist controversy," see Lasnik 1999. For similar attempts with regard to the "autonomy" thesis, see Chomsky 1977. Critics of the "innateness hypothesis" may have in mind issues of modularity and species-specificity, though that is unclear, since proposals with any substance are highly modular and (so far as is known) species-specific. See Jenkins 1997, 1999, Marcus 1998.

23. On conceptual and empirical arguments, with varying conclusions, see, among other works, those mentioned in note 1, *MP*, and Chomsky 1998.

24. Any interpretation of L is computational in some sense, raising difficult and obscure questions about what this means for a cognitive system. These are not to be confused with problems of processing (parsing, production).

25. Among many other works, see Chomsky 1986a on vacuous movement, Fox 1995, 1998 and Reinhart 1993 on the LF counterpart, and Collins 1997 on local determination.

26. For example, integrated action/perception models motivated by computational savings over construction of the detailed properties of a presented scene. For a review, see Clark 1998.

27. The properties of features and assembly form a large part of the subject matter of traditional and modern linguistics; I will put these topics aside here, including questions about organization of assembled features within a lexical item LI. Also left to the side is the question whether LI is assembled in a single operation or at several stages of the derivation, as in Distributed Morphology (Halle and Marantz 1993). Rephrasing of the account just given in these terms is straight-forward. Recall that L is a *state* of FL; state changes, of course, may modify the lexicon.

28. Or, if we distinguish independent selections of a single lexical item, a numeration Num (as in *MP*), an extension I will put aside until it becomes relevant.

29. It would not suffice to say that constant memory can be accessed throughout the derivation. The lexicon is a distinct component of memory; for CHL, our beliefs about the stars don't matter, but the lexical properties of *star* do. However hard it may be to make the distinction properly, there is good reason to believe that it is real.

30. Much current work allows Move to raise a to [Spec, V] even when Merge or Agree would be options; see Lasnik 1995c and sources cited there. An interesting question is whether the evidence for this conclusion can be incorporated within the more restricted framework envisaged here under a proper interpretation of the qualification "when possible."

31. I am putting aside many questions concerning the substantive/functional distinction, adopting it only for heuristic purposes. Also omitted is the D head of DP (which seems to belong to a different system), more complex verbal construc-

tions, and the question whether nontransitive ones have a light verb head. Some mightfor example, *seem*, which c-commands the experiencer Spec even in nonraising languages like English, either because it selects a light verb with this Spec or because it raises from that position to a light verb. Ignored as well are the "peripheral" systems outside TP; I will use C and T as surrogates for richer systems. On these matters see Rizzi 1997, Cinque 1999, and many other studies on the CFC systems and others. The concepts "inherent/structural Case" are understood as in Chomsky 1981, 1986b: q-related versus structurally determined.

32. In the sense of Pesetsky (1982), modifying ideas of Grimshaw (1979).

33. Strengthening *may* to *must* stipulates part of Burzio's Generalization; the rest should follow from Case/agreement theory.

34. Irrelevant to our concerns here, for C (5a) might be parametrically contingent on other operations (e.g., partial *wh*-raising as in German). Other possible parameterization is put aside here.

35. In effect, Hale and Keyser's (1993) theory adapted in *MP* (sec. 4.6). One consequence of this conception of q-theory is that the q-Criterion cannot be satisfied by raising an argument to a q-position or by raising of "q-features" (the existence of such features aside, I will suggest below that feature movement may not be possible). Other conceptions reject these conclusions (see, e.g., Hornstein 1999, Manzini and Roussou (to appear)).

36. The examples in the text are translations of Icelandic examples given by Jonas (1996). It had been assumed that Subj is higher than Obj in these constructions, but Jonas found that the conclusions relied on improper choice of left-edge markers. This eliminates complications in *MP* about the issue.

37. The counterpart of the *to* phrase in *seems to-me [NP to . . .]*. This phrase is sometimes described as the complement of V, but I assume it to be a Spec in a Larsonian shell. See note 31.

38. On (15a), see Sigurðsson 1996. On (15b), see *MP*, sec. 4.9, (168). The structure of subjects in multiple-subject constructions raises a wide range of questions put aside here.

39. See *MP* and Martin 1996. Some of these distinctions have been attributed to trace government, but that mechanism is not available here.

40. Observe that (9), (10), and (12c) appear problematic because of raising of *a proof* from DO of *discover*. There are reasons to suppose that the actual structure at this stage is the expected *there to be [discovered a proof]* as in similar languages, the preferred English construction *a proof [discovered t]* being formed outside the system we are now considering. I will assume so, leaving the issue to the side here.

41. Why not dispense with LA, just selecting subarrays cyclically? Apart from the general considerations about access reduction already discussed, there is a more specific reason: chain properties can be reduced in significant part to identity if lexical arrays are enriched to numerations. To achieve the same result with cyclic choice of successive subarrays requires continual access to the full lexicon and memory of how many times each item has been selected.

42. An island effect, if it exists at all, is very weak with such structures.

43. Early work sought to establish the categories of A- and Ā-movement ("Move NP," "Move *Wh*"), later head movement, while parallel inquiries sought commonalities. Important outcomes were Rizzi's (1990) theory of Relativized Minimality and Lasnik and Saito's (1992) Move a theory. The distinctions mentioned here crosscut these categories.

44. "Stylistic" operations might fall within the phonological component (see *MP*, sec. 4.7.3, Kidwai 1996). Operations lacking overt counterparts and apparently not interacting with CHL might be among the principles of interpretation of LF, hence "postcyclic," inspecting a representational level in the manner of many other systems (including binding theory, on the assumptions of *MP*). If so, much of the very enlightening recent work on ellipsis and antecedent-contained deletion (along with event structure and other topics) could be understood as an exploration of the language-external systems at the border of the language faculty, roughly analogous to acoustic and articulatory phonetics on the sound side.

45. The categories might overlap, but unproblematically it seems. System design should preclude unwanted cases of improper movement. That seems attainable, but must be demonstrated. I will continue to restrict attention to raising of XP.

46. It also suggests a new approach to some Empty Category Principle (ECP) issues, such as subject extraction (Idan Landau, personal communication).

47. To clarify this and related conclusions and establish them in full generality requires a far more comprehensive review and analysis than is undertaken below. Similar qualifications hold throughout.

48. See Fox 1998 and Nissenbaum 1998. If adjunction is restricted as suggested in *MP* (sec.4.7.3), then movement to the edge will be to a Spec position for nP as well as CP. Phases might also be the target for QR, if this noncyclic operation targets C', merging the raised quantifier phrase between C and [Spec, C]; see note 107.

49. Convergence is not guaranteed, of course (it can fail in many ways)only permitted without look-ahead, the desideratum we are exploring. Conditions could be added to restrict crash, but they are redundant, simply restating properties of convergence, unless motivated in some other way. Questions arise about operations that appear to violate Subjacency (see references of note 92, among others). Note the restriction to uninterpretable features *in the domain.* Legitimacy of those at the edge (specifically, EA) will be determined at the next higher phase, a matter that opens interesting questions, put aside here.

50. Parametrically varying properties of H enter into the application of (24), which might be extended to head movement (see note 93 and text). I will call the EPP-feature a *P-feature* (*periphery feature*) if H does have an appropriate EPP-feature by virtue of its inherent properties (e.g., the Case/agreement properties of n, the Q-feature of interrogative C). The device is introduced to extend the general theory of movement beyond A-movement, but should raise warning flags.

51. In *MP* an unsatisfactory argument was required to bar expletives from merging in OS constructions. The analogous problem arose for CP, but was ignored.

52. For some speculations along similar lines, see *MP*, pp. 294 (condition (76)) and 377.

53. On some assumptions, though not here, IFM passing through [Spec, n2] is improper movement.

54. Among other questions, what is the status of small clauses, or relative clauses and other adjuncts? Possibly the latter are derived "in parallel," in the manner of multidimensional analyses of coordination or parentheticals, with their own LAs and with the ultimate status of the adjunct being determined in the larger structure in which it is inserted (as for other multidimensional structures).

55. See note 31. Many questions about the internal structure of the words are put aside.

56. See, for example, Collins 1997, Johnson and Lappin 1997, Yang 1997.

57. The reference to "fundamental principles" in some (obscure) sense is crucial. Otherwise, for properly selected categories of expressions (which may well exhaust the possibilities), look-ahead properties of computation at stage a might be overcome by resort to whatever aspects of UG determine that the wrong choice eventually crashes.

58. The issue is whether crash "frees up" derivational paths not selected earlier, as assumed in *MP* (incorrectly, I assume here). Computations can crash in endless ways, raising no complexity issue. See note 49.

59. See Chomsky 1998, Collins 1997, Frampton and Gutmann 1999.

60. These are conceptual "good design" conditions, but fairly trivial ones, based on the assumption that less machinery is better than more.

61. The significance of (ii) was brought out in Epstein's (1999) derivational analysis of c-command, which underlies much important work since, including Epstein et al. 1998.

62. That the copy theory is the simplest version is clear in the earliest formulation, in Chomsky 1955. The more complex operation Merge-Delete was adopted there on the assumption that T-markers are mapped to PF. When that picture was modified under the Extended Standard Theory, Delete was abandoned in favor of trace theory, trace being a new kind of element. That was an error (mine), and the copy theory, which restored the simplest case of the original approach, was mistakenly regarded as a further innovation. The divergence of history and logic has caused much confusion.

63. We could, for example, identify this as K' = K with the occurrence of a in question replaced by some designated element Occ distinct from anything in K. In Chomsky 1955 *occurrence* is defined in terms of linear order, adapting a device from Quine 1940; but that mechanism is not available here.

64. Suppose, for example, that the operation Move could raise the object of V to become the object of V', forming [V' Obj]. Then if V = V' (by virtue of V-raising), the sisters of Obj would be identical.

65. See *MP*, sec. 3.5, and for improved versions with many consequences, Fox 1998 and Sauerland 1998b. A question that might be raised is where the "reconstruction" operation takes place: within narrow syntax or on the other side of the interface, along with binding theory (I am assuming, following *MP*) and other interpretive systems. Similar questions arise about other covert operations; see note 44.

66. Attempts in *MP* to account for successive-cyclic movement in terms of linked chains and other devices can be eliminated: they were based on failure to take seriously enough the actual notion of chain. They were also in error in not recognizing the role of intermediate traces in interpretation and computation. The conclusions drawn there about trace invisibility no longer hold, though interesting aspects of the question remain.

67. The compositional definition of c-command is suggested by Epstein's (1999) derivational approach (see note 61): a c-commands b if a is merged with K containing b. There are empirical differences between the two approaches, to which I will return. The derivational definition also raises some questions: in particular, why does "containment" enter (i.e., why does X merged with Y c-command terms of Y)? The matter is addressed in Epstein et al. 1998, chap. 6, but inconclusively, as far as I can see. An argument for asymmetry of c-command is also presented there; I am assuming that there is no asymmetry, its effects being derived in other ways.

68. Sisterhood relations would be carried over under head adjunction if the result were taken to be an extended MLI. Many questions dissolve if head adjunction is part of the phonological component. There are, I think, independent reasons to suspect that this may be true, at least over an interesting range, but I will defer the question.

69. See Kitahara 1999, Lasnik 1995a,b, Uriagereka 1999 a.

70. There are important distinctions among these features, which I will largely ignore here. Problems of interpretation are also not trivial, as in (i) (v).

(i) animal languages (is, are) their main research interest

(ii) three books (is, are) too much to read in a week

(iii) we expected animal languages to be their main research interest

(iv) we found three books too much to read in a week

(v) animal languages raise(s) serious issues/seem(s) to be their main research interest

In (i) and (ii) the difference in meaning appears superficially to lie in the verbal inflection, but it carries over to forms lacking that inflection, as in (iii) and (iv). And though agreement properties of the copula (with surface subject or complement) are a factor, the matter is more convoluted ((iv) and (v)). See Reid 1991 for discussion from a "functionalist" perspective.

71. A possible case is control contingent on long-distance agreement. See *MP*, sec. 4.4.5, (40) (43), reviewing work of Anna Cardinaletti and Michal Starke. It

was assumed there that both binding and control are subject to these long-distance effects, but closer examination of Italian by Carlo Cecchetto (personal communication) indicates that the effects are limited to control alone (see also Cardinaletti 1997, Den Dikken 1995). That would explain the failure of binding in such examples as *there seem to each other [to be many men in the room]* and fits with other evidence that binding requires an "overt" (possibly PRO, trace, or *pro*) antecedent, not just an implicit argument; see Rizzi 1986. There are additional complications. The phenomena seem less clear for passives than unaccusatives (possibly because of interference from an implicit subject) and become blurred or disappear in "long-distance" cases; see Hornstein 1996, Lasnik 1997.

72. Contrary to *MP*. The extension of these notions to features raises difficulties, not insuperable but better avoided, as seems possible.

73. I continue to disregard (35d), abstracting to the expected form (35b); see note 40.

74. They have also shifted as inquiry has proceeded: thus, from the 1950s through the mid 1970s such matters as quantifier scope were commonly taken to be prototypical examples of "surface" interpretation, whereas more recent work generally takes them to be prototypical properties at LF.

75. The approach renders superfluous the intuitive motivation for pied-piping proposed in *MP*, in terms of PF crash. For argument that some such device nevertheless operates in languages with rich morphology, with consequences for null subject, barriers, and Spell-Out, see Uriagereka 1999 a. For an approach somewhat similar to what is outlined below, from a partially different perspective, see Frampton and Gutmann 1999.

76. If a reflex of an interpretable f-set, it would be erased in situ by the f-set of *candidate* itself.

77. See *MP*, Ura 1996. The condition stated there refers only to Specs of the same head. Whether the generalization (41) is appropriate depends on answers to questions about the structure of more complex constructions (double object verbs, etc.). These and other questions, including parameterization, are put aside here. See Boskovic * 1997 and McGinnis 1998, among others. The notion "feature occurrence" used implicitly here for expository convenience can be eliminated by restatement in terms of the heads to which features belong.

78. Among the problems that arise is the status of scrambling. The logic would suggest that for at least some cases, a scrambling feature induces pied-piping even after Case assignment, the pied-piped element being "attracted" by a higher probe, whereas other cases fall into a category distinct from feature-driven movement. For exploration of alternatives in a comparative study, see Sauerland 1998a.

79. For a similar configuration in phonology, with > = linear order, see the discussion in Halle 1995 of coronal assimilation in Sanskrit, barred by an intervening (nonassimilating) coronal.

80. On the assumptions made in *MP*, sec. 4.10, multiple-subject (including transitive expletive) options are parameterized in terms of deletion of EPP-features. See particularly Ura 1996, and for skepticism about the option, Zwart 1997. Agreement in the sense discussed here is to be distinguished from concord, with different properties.

81. The analysis of structural Case is along the lines proposed by George and Kornfilt (1981). As they observe, structural Case linked to f-features may be dissociated from finiteness. Matters become more complex when we consider ergative/absolutive and mixed systems, and languages in which f-features without finiteness do not suffice for nominative Case assignment (see Iatridou 1993).

82. For a different approach, on the assumption that *there* has structural Case, see Lasnik 1995a. For a different perspective on a wide range of related issues, see Moro 1997.

83. Similar conclusions are supported by locative and quotative inversion (see Collins 1997), though with restrictions and complications, and similarities to other poorly understood constructions (e.g., *still unclear remain (are, seem to be) the answers to those questions*).

84. Among these, the ECP, Subjacency, and other conditions. But many problems remain. See Kitahara 1999, Lasnik 1995a,b, Uriagereka 1999 a.

85. Perhaps substantive lexical categories do not exist, only bare roots. Configurational and morphological properties, along with interpretable noncategorial features of the root, would then determine relevant structural properties, as in Semitic. The possibility is suggested by work since the 1960s on derivational versus transformationally induced morphology. See Marantz 1997 for pertinent discussion.

86. A different motivation, based on the need to overcome ambiguity in the composite verbal element at LF, is developed by Uriagereka (1996). One can also think of various functional arguments: the familiar trade-off between order rigidity and richness of inflection, facilitation of search for attracted elements, and so on.

87. Pure inherent Case I take to be a distinct phenomenon, "invisible" to matching, as if inherent Case inactivates the f-set. See note 31, and for more on the topic, McGinnis 1998 and sources cited there.

88. Reformulation is needed if [3person] is lacking, as has been suggested. Sigurðsson concludes that third person nominative Case requires agreement in the monoclausal construction [Dat V-T [VP *t* Nom]] and allows it as an option, alternating with default T, in the biclausal construction [Dat V-T *t* [TP Nom . . .]] (*t* the trace of Dat; irrelevant properties omitted). These facts could be expressed in terms of optionality of [person] reduction in matrix T, leaving open the reasons.

89. If Merge could induce agreement, nP-internal subject would manifest object agreement and have accusative Case, which would delete, preventing raising to [Spec, T].

90. Questions remain about French-style *il*, with a definiteness effect and other properties. Suppose the [person] feature of T can delete or shift to default with *there*-type Expl, in the manner suggested for quirky Case. That might accommodate "list readings" (as in *there (is/*am, remains/*remain) only me, there (are, remain) only us (John and Bill)*, in response, say, to *who's still here to do the work*), with no person agreement or nominative Case assignment, in the absence of the full complement of f-features. Also to be accounted for is the fact that in n phases the external argument pronoun does not agree with the n head, possibly indicating that argument pronouns have true D-N structure, unlike expletives. See Cardinaletti and Starke 1994 and Uriagereka 1988 for some relevant considerations. Needless to say, these remarks barely touch on a rich array of questions.

91. To complete the analogy, C (and n with its f-set deleted) may have a non-specific P-feature analogous to [person] for Tdef, perhaps contingent on assignment of the EPP-feature to a phase; see discussion of (24), (25).

92. Following ideas of Watanabe (1992) and Hagstrom (1998). This is not to be confused with the distinction between pied-piping of a full XP and pied-piping of a minimal operator (e.g., interrogatives/raising relatives vs. comparatives/complex adjectivals/nonraising relatives), a choice fixed by context (with various questions about relatives; see Sauerland 1998b). The *wh*-island analysis extends to other constructions if the feature that drives movement shares properties with [wh] (assuming here a hierarchy of features); see Hagstrom 1998 for supporting evidence. Multiple overt *wh*-movement as in some Slavic languages might be analogous to multiple head options for A-movement along with a principle to overcome the Subjacency effect. See Richards 1997, adapting ideas in Brody 1995, and for a general critical review, Boskovic * 1998. As is well known, pied-piping in Ā-movement differs from the A-chain analogue, with variation among languages and constructions that is poorly understood.

93. The perennial troublemaker (i) falls into place if the (undeleted) [person] feature of embedded *there* bars association of matrix T to *three men*.

(i) *there seem there to be three men in the room

Groat (1997) points out further complications. Thus, whatever its status, (ii) is more acceptable.

(ii) there look as though there are three men (vs. *[a man]) in the room

That suggests that the [person] barrier may be overridden in some manner.

94. The *MP* analysis of superraising assumed that *it* raises to matrix T by locality and the derivation crashes because *friends* cannot be Case-checked. But covert attraction of the f-features of *friends* by matrix T should still be possible, with convergence (Eduardo Raposo, personal communication). See Raposo and Uriagereka 1996 for a different approach.

95. Furthermore, I am speculating that XP can only appear if forced (barring Expl-raising always, and argument raising except to implement IFM or LF interpretive effects); see discussion of (24), (25). I overlook questions of possible parameterization for OVS languages; see Ura 1996.

96. Translations of Icelandic examples: (51a), with a default variant, from Sigurðsson 1996; (51b,c) from Boeckx (to appear); these and other examples, and discussion, in Schütze 1997. See also (15), (45biii), and note 88.

97. Unwanted chains could be formed without further constraints: for example, given a and b, and g in a, g could be merged to b (by Move), which is then merged to a, forming an unwanted <g, g> chain violating certain conditions. Motivated by Pollock's (1989) split-inflection ideas, in chapter 2 of *MP* I adopted Attract a in addition to Move a. In section 4.5.6 I sought to eliminate Move a as an independent operation, taking it to be conditional on Attract, revised as feature attraction. Considerations reviewed here seem to me to support this general point of view (now much simplified).

98. The problem is noted in *MP* (p. 385, n. 50) but left unresolved.

99. The conception is similar to the strict derivational interpretation of Spell-Out proposed by Epstein et al. (1998); for similar suggestions, on different grounds, see Uriagereka 1996, 1999 b. The basic architecture resembles that of Bresnan (1971); her results on the interaction of phonological and transformational rules fall into place more directly than in the Extended Standard Theory model. See several papers in Abraham et al. 1996 and Yang 1997 for related discussion.

100. The issue is narrow and might be avoidable in other ways. It would be restricted to subjects in OS constructions if the "reconstruction" operation induced by the final step of $\bar{A}$-movement preempts deletion, so that strict identity holds throughout $\bar{A}$-chains.

101. The discussion of categories and labels here largely follows Collins 1997. Questions about categories raised in note 64 are not relevant here; the syntactic object has some distinctive property that must be recognized.

102. See Kayne 1994 and *MP*. Kayne assimilates adjunct and specifier. *MP* suggests that XP-adjunction may not be part of narrow syntax (sec. 4.7.3)possibly not a central range of cases of head adjunction either (see notes 32 and 46).

103. On the conventional and reasonable assumption that the selectional feature itself may be optional, as in the case of a verb that may or may not have IO (*send*, etc.).

104. An elaborate argument to guarantee projection of the target in *MP* is superfluous, under this reanalysis. The label is determined by the probe for the Agree component of Move, and by the selector for the Merge component. These are different features, but they yield the same choice of label. It seems a notational question whether we take the label to be the original LI or a reduced MLI.

105. New containment relations are defined whichever choice we make, in incommensurable ways; identity is irrelevant.

106. Whether sisterhood (hence c-command) is preserved depends on how (or if) the notion is defined for head adjunction. See note 46.

107. See Richards 1997 for evidence in support of Local Merge in the important case of multiple Move. See also note 33; the proposal there falls under Local Merge, with the Extension Condition obviated vacuously for "postcyclic" QR.

108. In *MP* it could be avoided only by recourse to the (dubious) distinction between deletion and erasure. Nothing follows about functional categories with semantic properties, as in Rizzi 1997, Cinque 1999.

References

Abraham, Werner, Samuel David Epstein, Höskuldur Thráinsson, and C. JanWouter Zwart, eds. 1996. *Minimal ideas*. Amsterdam: John Benjamins.

Belletti, Adriana. 1990. *Generalized verb movement*. Turin: Rosenberg and Sellier.

Berwick, Robert. 1997. Walking the walk, talking the talk. *Los Angeles Times Book Review*, Sept. 7.

Boeckx, Cédric. To appear. Quirky Agreement. *Studia Linguistica*.

Bosch, Laura, and Nuria Sebastián-Gallés 1997. Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition* 65, 33 69.

Boskovic Zeljko. * 1997. *The syntax of nonfinite complementation*. Cambridge, Mass.: MIT Press.

Boskovic Zeljko. 1998. *Wh*-phrases and *wh*-movement in Slavic. Presented at the Comparative Slavic Morphosyntax Workshop, Bloomington, IN. http://www.indiana.edu/~slavconf/linguistics/index.html

Bresnan, Joan. 1971. Sentence stress and syntactic transformations. *Language* 47, 257 281.

Brody, Michael. 1995. *Lexico-Logical Form: A radically minimalist theory*. Cambridge, Mass.: MIT Press.

Cardinaletti, Anna. 1997. Agreement and control in expletive constructions, *Linguistic Inquiry* 28, 521 533.

Cardinaletti, Anna, and Michal Starke. 1994. The typology of structural deficiency. Ms., University of Venice and University of Geneva.

Carr, Philip. 1997. Review of Chomsky et al., *Language and thought* (Warwick, R.I.: Moyer Bell, 1993). *Journal of Linguistics* 33, 215 218.

Chomsky, Noam. 1955. The Logical structure of linguistic theory. Ms., Harvard University and MIT, Cambridge, Mass. [Revised 1956 version published in large part in 1975 (New York: Plenum) and 1985 (Chicago: University of Chicago Press).]

Chomsky, Noam. 1975. *Reflections on language*. New York: Pantheon.

Chomsky, Noam. 1977. *Essays on form and interpretation*. New York: North-Holland.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1986a. *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1986b. *Knowledge of language*. New York: Praeger.

Chomsky, Noam. 1991. Linguistics and cognitive science: Problems and mysteries. In Asa Kasher, ed., *The Chomskyan turn*. Oxford: Blackwell.

Chomsky, Noam. 1995a. Language and nature. *Mind* 104, 1 61.

Chomsky, Noam. 1995b. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1996. Internalist Explorations. Ms., MIT, Cambridge, Mass.

Chomsky, Noam. 1998. Some observations on economy in generative grammar. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky, eds., *Is the best good enough?* Cambridge, Mass.: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads*. Oxford: Oxford University Press.

Clark, Andy. 1998. Where brain, body, and world collide. *Daedalus* 127.2 (special issue, "The Brain"), 257 280.

Collins, Chris. 1997. *Local economy*. Cambridge, Mass.: MIT Press.

Dikken, Marcel den. 1995. Binding, expletives, and levels. *Linguistic Inquiry* 26, 347 354.

Epstein, Samuel David. 1999. The derivation of syntactic relations. In Samuel David Epstein and Norbert Hornstein, eds., *Working Minimalism*. Cambridge, Mass.: MIT Press.

Epstein, Samuel David, Erich Groat, Ruriko Kawashima, and Hisatsugu Kitahara. 1998. *A derivational approach to syntactic relations*. Oxford: Oxford University Press.

Fodor, Jerry. 1983. *The modularity of mind*. Cambridge, Mass.: MIT Press.

Fox, Danny. 1995. Economy and scope. *Natural Language Semantics* 3, 283 341.

Fox, Danny. 1998. Economy and semantic interpretation. Doctoral dissertation, MIT, Cambridge, Mass.

Frampton, John, and Sam Gutmann. 1999. Cyclic computation: A computationally efficient minimalist syntax. *Syntax* 2, 1 27.

George, Alexander. 1996. Katz astray. *Mind & Language* 11, 295 305.

George, Leland, and Jaklin Kornfilt. 1981. Finiteness and boundedness in Turkish. In Frank Heny, ed., *Binding and filtering*. Cambridge, Mass.: MIT Press.

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10, 279 326.

Groat, Erich. 1997. A derivational program for syntactic theory. Doctoral dissertation, Harvard University, Cambridge, Mass.

Hagstrom, Paul. 1998. Decomposing questions. Doctoral dissertation, MIT, Cambridge, Mass.

Hale, Kenneth, and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Kenneth Hale and Samuel Jay Keyser, eds., *The view from Building 20*. Cambridge, Mass.: MIT Press.

Halle, Morris. 1983. On distinctive features and their articulatory implementation. *Natural Language & Linguistic Theory* 1, 91 105.

Halle, Morris. 1995. Feature geometry and feature spreading. *Linguistic Inquiry* 26, 1 46.

Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale and Samuel Jay Keyser, eds., *The view from Building 20*. Cambridge, Mass.: MIT Press.

Hermer, Linda, and Elizabeth Spelke. 1996. Modularity and development: The case of spatial reorientation. *Cognition* 61, 195 232.

Hornstein, Norbert. 1996. Existentials, A-chains, and reconstruction. Ms., University of Maryland, College Park.

Hornstein, Norbert. 1999. Movement and Control. *Linguistic Inquiry* 30, 69 96.

Iatridou, Sabine. 1993. On nominative Case assignment and a few related things. In Colin Phillips, ed., *Papers on Case and agreement II*. (MIT Working Papers in Linguistics 19.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Jenkins, Lyle. 1997. Biolinguistics: Structure, development and evolution of language. *Web Journal of Formal, Computational and Cognitive Linguistics* 1(2). http://www.kcn.ru/tat_en/science/fccl/gp008.txt.

Jenkins, Lyle. 1999. *Biolinguistics*. Cambridge: Cambridge University Press.

Johnson, David, and Shalom Lappin. 1997. A critique of the Minimalist Program. *Linguistics and Philosophy* 20, 272 333.

Jonas, Dianne. 1996. Clause structure and verb syntax in Scandinavian and English. Doctoral dissertation, Harvard University, Cambridge, Mass.

Jusczyk, Peter. 1997. *The discovery of spoken language*. Cambridge, Mass.: MIT Press.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Kidwai, Ayesha. 1996. Word order and focus positions in Universal Grammar. Ms., Jawaharlal Nehru University.

Kitahara, Hisatsugu. 1999. Eliminating * as a feature (of traces). In Samuel David Epstein and Norbert Hornstein, eds., *Working minimalism*. Cambridge, Mass.: MIT Press.

Lasnik, Howard. 1995a. Case and expletives revisited: On Greed and other human failings. *Linguistic Inquiry* 26, 615 633.

Lasnik, Howard. 1995b. Last Resort. In Shosuke Haraguchi and Michio Funaki, eds., *Minimalism and linguistic theory*. Tokyo: Hituzi Syobo.

Lasnik, Howard. 1995c. A note on pseudogapping. In Rob Pensalfini and Hiroyuki Ura, eds., *Papers on minimalist syntax*. (MIT Working Papers in Linguistics 27.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass. [Reprinted, with minor revisions, in Lasnik, Howard. 1999. *Minimalist analysis*. Oxford: Blackwell.]

Lasnik, Howard. 1997. Levels of representation and the elements of anaphora. In Hans Bennis, Pierre Pica, and Johan Rooryck, eds., *Atomism and binding*. Dordrecht: Foris.

Lasnik, Howard. 1999. On the locality of movement: Formalist syntax position paper. In Michael Darnell, Edith Moravcsik, Frederick Newmeyer, Michael

Noonan, and Kathleen Wheatley, eds., *Functionalism and formalism in linguistics*. Vol. 1, 33 54. Amsterdam: John Benjamins.

Lasnik, Howard, and Mamoru Saito. 1992. *Move* a. Cambridge, Mass.: MIT Press.

Lewontin, Richard. 1990. The evolution of cognition. In Daniel Osherson and Edward Smith, eds., *An invitation to cognitive science*, Vol. 3: *Thinking*. Cambridge, Mass.: MIT Press.

Lewontin, Richard. 1998. The evolution of cognition: Questions we will never answer. In Dan Scarborough and Saul Steinberg, eds., *An invitation to cognitive science*, Vol. 4: *Methods, models, and conceptual issues*. Cambridge, Mass.: MIT Press. [Update of Lewontin 1990.]

Locke, John. 1997. A theory of neurolinguistic development. *Brain and Language* 58, 265 326.

Manzini, Maria Rita, and Anna Roussou. To appear. A minimalist theory of A-movement and control. *Lingua*.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark, and Alexander Williams, eds., *Proceedings of the 21st Annual Penn Linguistics Colloquium*. (Pennsylvania Working Papers in Linguistics 4.2.) Penn Linguistics Club, University of Pennsylvania, Philadelphia.

Marcus, Gary. 1998. Can connectionism save constructivism? *Cognition* 66, 153 182.

Martin, Roger. 1996. A minimalist theory of PRO and control. Doctoral dissertation, University of Connecticut, Storrs.

McGinnis, Martha. 1998. Locality in A-movement. Doctoral dissertation, MIT, Cambridge, Mass.

Mehler, Jacques, and Emmanuel Dupoux. 1994. *What infants know*. Oxford: Blackwell.

Moro, Andrea. 1997. *The raising of predicates*. Cambridge: Cambridge University Press.

Nissenbaum, Jon. 1998. Movement and derived predicates: Evidence from parasitic gaps. In Uli Sauerland and Orin Percus, eds., *The interpretive tract*. (MIT Working Papers in Linguistics 25.) MITWPL, Department of Linguistics and Philosophy, MIT.

Orr, Allen. 1997. The softer side of sociobiology. *Boston Review*, Oct./Nov.

Pesetsky, David. 1982. Paths and categories. Doctoral dissertation, MIT, Cambridge, Mass.

Phillips, Colin. 1996. Order and structure. Doctoral dissertation, MIT, Cambridge, Mass.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20, 365 424.

Quine, Willard V.O. 1940. *Mathematical logic*. Cambridge, Mass.: Harvard University Press.

Raposo, Eduardo, and Juan Uriagereka. 1996. Indefinite *se*. *Natural Language & Linguistic Theory* 14, 749 810.

Reid, Wallace. 1991. *Verb and noun number in English*. London: Longman.

Reinhart, Tanya. 1993. *Wh*-in-situ in the framework of the Minimalist Program. Ms., TelAviv University. [Published in *Natural Language Semantics* 6, 3 28 (1997).]

Richards, Norvin. 1997. What moves where when in which language? Doctoral dissertation, MIT, Cambridge, Mass.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17, 501 557.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman, ed., *Elements of grammar*. Dordrecht: Kluwer.

Sauerland, Uli. 1998a. Erasability and interpretation. Ms., MIT, Cambridge, Mass.

Sauerland, Uli. 1998b. On the making and meaning of chains. Doctoral dissertation, MIT, Cambridge, Mass.

Schütze, Carson. 1997. INFL in child and adult language: Agreement, Case and licensing. Doctoral dissertation, MIT, Cambridge, Mass.

Sigurðsson, Halldór Ármann. 1996. Icelandic finite verb agreement. *Working Papers in Scandinavian Syntax* 57, 1 46.

Sportiche, Dominique. 1995. Sketch of a reductionist approach to syntactic variation and dependencies. In Héctor Campos and Paula Kempchinsky, eds., *Evolution and revolution in linguistic theory*. Washington, D.C.: Georgetown University Press.

Ura, Hiroyuki. 1996. Multiple feature-checking: A theory of grammatical function splitting. Doctoral dissertation, MIT, Cambridge, Mass.

Uriagereka, Juan. 1988. On government. Doctoral dissertation, University of Connecticut, Storrs.

Uriagereka, Juan. 1996. Formal and substantive elegance in the Minimalist Program. In Chris Wilder, Hans-Martin Gärtner, and Manfred Bierwisch, eds., *The role of economy principles in linguistic theory*. Berlin: Akademie-Verlag.

Uriagereka, Juan. 1999 a. Minimal restrictions on Basque movement. *Natural Language & Linguistic Theory* 17, 403 444.

Uriagereka, Juan. Forthcoming b. Multiple Spell-Out. In Samuel David Epstein and Norbert Hornstein, eds., *Working minimalism*. Cambridge, Mass.: MIT Press.

Watanabe, Akira. 1992. Subjacency and S-Structure movement of *wh*-in-situ. *Journal of East Asian Linguistics* 1, 255 291.

Yang, Charles. 1997. Minimal computation: Derivation of syntactic structures. Master's thesis, Artificial Intelligence Laboratory, MIT, Cambridge, Mass.

Zwart, C. Jan-Wouter. 1996. *Morphosyntax of verb movement*. Dordrecht: Kluwer.

Zwart, C. Jan-Wouter. 1997. Transitive expletive constructions and the evidence supporting the multiple specifier hypothesis. In W. Abraham and E. van Gelderen, eds., *German-Syntax Problems. Problematic Syntax*. Tübingen: Narr.

Chapter 4
The Primitives of Temporal Relations

Hamida Demirdache and Myriam Uribe-Etxebarria

This chapter presents a restrictive theory of temporal relations in natural languages that reduces the grammar of Tense and Aspect to the same set of universal substantive and structural primitives. The proposal that there is a single grammar for temporal relations allows us to uniformly derive the compositional interaction of Tense and (recursive) Aspect, crosslinguistically.

The theory proposed here is grounded in a semantic and a syntactic parallel that we establish between Tense and Aspect. We take as a point of departure Klein's (1995) proposal that both Tense and Aspect relate two times. We derive this proposal from the thesis that tenses and aspects are dyadic predicates of *spatiotemporal* ordering: they establish a topological relation between two time-denoting arguments.

We propose a uniform structural representation for temporal relations that captures the semantic parallel between Tense and Aspect syntactically. In particular, we extend the proposals in Zagona 1990 and Stowell 1993 by proposing that *both* Tense and Aspect are dyadic predicates that head a maximal projection in the syntax (TP and AspP, respectively) and take time-denoting phrases as arguments.

The strict parallel that we establish between the syntax and semantics of Tense and Aspect allows us to reduce their grammar to the same set of substantive primitivesTense and Aspect are both dyadic predicates of spatiotemporal ordering, as well as to the same set of structural primitivesthere is a single uniform phrase structure for temporal relations.

Based on Hale (1984), we define spatiotemporal relations in terms of a single basic semantic opposition: that of central versus noncentral coincidence. This proposal constrains the number of logically possible temporal relations that we expect to find in natural languages by restricting the topological relations that Tenses and Aspects establish between their

temporal arguments to three basic relations. At the same time it explains the pervasive use, crosslinguistically, of prepositions as well as locative, motion, directional, postural, and stance verbs to express temporal and aspectual relations. The theory of temporal relations developed here thus opens the way to a simple account of the diversity of temporal and aspectual systems instantiated in natural languages.

Our proposal bears on general questions concerning the traditional distinction between lexical and functional categories and the role of syntactic heads in the grammar. As a result, it has implications for minimalist theories of clausal architecture. Within our theory Tense and Aspect, which are standardly assumed to be canonical examples of functional categories, are heads that can play a dual role in the syntax, as has been argued for n and P (see Chomsky 1995). As lexical heads, they have argument structure that they project into the syntax (see also Zagona 1990, Stowell 1993). As functional heads, they can enter into feature-checking relations (i.e., Case, categorial, and/or f-features) with arguments generated in lower projections and displaced to their checking domain by Move a/Attract F. The analysis defended here thus provides further support for the claim that syntactic projections make use of multiple specifiers to fulfill the multiple functions of their heads, as first suggested by Koizumi (1994).

4.1 Tense

Tense has been classically analyzed as establishing an ordering relation between two times. Tense relates the time of utterance (UT-T) and the time at which the event (or state) denoted by the VP occurs (or holds) (EV-T). Thus, the standard assumption is that Past orders the event time BEFORE the utterance time, whereas Future orders the event time AFTER the utterance time, as illustrated in (1). 1

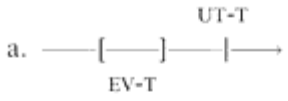
(1) a. Past: EV-T < UT-T
b. Future: UT-T < EV-T

Enç (1987), Zagona (1990), Giorgi and Pianesi (1991), Stowell (1993), Thompson (1995), and Uribe-Etxebarria (1994), among others, propose that the syntax and semantics of Tense can be derived from a theory of the structural representation of Tense together with independently motivated principles of Universal Grammar (see also Hornstein 1990 for the latter claim).

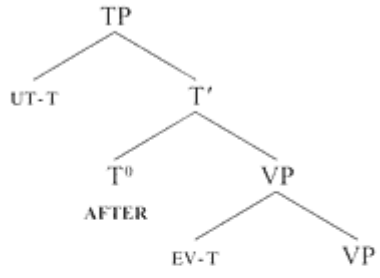
Zagona (1990) and Stowell (1993) propose to capture the idea that Tense relates two times syntactically by breaking down Tensestructurallyinto its semantic components. In particular, Zagona proposes that Tense is a head that projects a maximal projection TP and takes two time-denoting phrases as its arguments. The external argument of T0 is a Reference Time (typically, the UT-T in matrix clauses) and its internal argument is the EV-T. Stowell adopts this proposal and further argues that Tense is a dyadic predicate that establishes a temporal ordering relation between its two time-denoting arguments. Past Tense is analyzed as a temporal ordering predicate with the meaning of AFTER (UT-T AFTER EV-T) and Present as a temporal ordering predicate with the meaning of WITHIN (UT-T WITHIN EV-T). Stowell's (simplified) phrase structure for Tense is given in (2) and (3).

(2) *Past Tense*

a.

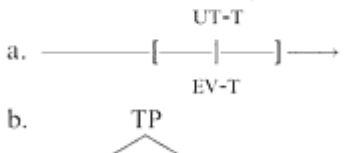


b.

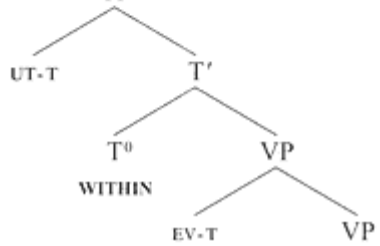


(3) *Present Tense*

a.



b.



Stowell identifies the EV-T with Kratzer's (1991) event argument. He departs from Kratzer, however, in analyzing this event argument as a time-denoting argument (rather than an argument of spatiotemporal location) and in assuming that all predicatesbe they stage- or individual-levelhave a (temporal) event argument. He follows Kratzer in analyzing this temporal argument as the true external argument of the verb. Since it is the highest (external) argument of the verb, it is base-generated in the highest subject position of a recursive VP shell structure, as shown in (2b) and (3b).

In what follows we adopt the proposal that Tense is a predicate projecting a maximal projection in the syntax and establishing an ordering relation between its two time-denoting arguments. We now turn to the role of Aspect.

4.2
(Grammatical) Aspect

Smith (1991) argues that Aspect conveys a viewpoint on the situation described by a sentencethat is, a temporal perspective that focuses all or parts of the situation. In Smith's words, "Aspectual viewpoints function like the lens of a camera, making objects visible to the receiver. Situations are the objects on which viewpoint lenses are trained. And just as the camera lens is necessary to make the object available for a picture, so viewpoints are necessary to make visible the situation talked about in a sentence" (p. 91). Thus, in a simple past tense sentence such as *Laïla wrote a book*, the event is presented as a whole, as completedas illustrated by the temporal schema in (4). (The slashes indicate the part of the event that is focused by the viewpoint; I and F are the initial and final bounds of the event.

(4) *Simple past sentence*
Laïla wrote a book. [I F]
////////////////////

In contrast, the progressive viewpoint in a sentence such as *Laïla was writing a book* focuses a subpart of the event of writing that includes neither its beginning nor its culmination. The event is presented without initial or final endpoints, as illustrated by the temporal schema in (5).

(5) *Progressive sentence*
Laïla was writing a book. [I F]
////////

Smith then argues that Aspect plays a crucial role in the semantic interpretation of a sentence: "Continuing the analogy of a viewpoint with the lens of a camera, we shall say that the part focused by a viewpoint is visible to semantic interpretation . . . . What is focused has a special status, which I will call visibility. Only what is visible is *asserted* . . ." (p. 99; emphasis added). Thus, in the progressive sentence in (5) only the time interval within the temporal contour of the event focused by Aspect is visible to the semantic interpretation. Since this time does not include the endpoints of the event, no assertion is made about whether the event of writing culminatedeven though the sentence is in the past. Thus, we can say *Laïla was writing a book, but she never finished it*. In contrast, in the simple past sentence in (4) the event is portrayed in its entiretyas including both its initial and final endpoints. Hence, *Laïla wrote a book, but she never finished it* is a contradictory statement.

In sum, the role of Aspect is to focus (pick out) an interval in the temporal contour of the event described by a sentence. Only the time interval focused by Aspect is visible to semantic interpretation. We call the time interval in the event time of the VP that Aspect focuses, the *Assertion Time* (AST-T), following Klein (1995, 687): "The Assertion Time is the time for which an assertion is made or to which the assertion is confined; for which the speaker makes a statement."

4.3
Reducing Tense and Aspect to the Same Set of Substantive and Structural Primitives

In order to reduce Tense and Aspect to the same set of substantive and structural primitives, and thus uniformly derive their compositional interaction crosslinguistically, we establish strict parallels between the *semantics* of Tense and Aspect on the one hand, and the *syntax* of Tense and Aspect on the other.

4.3.1
A Uniform Semantic Approach to Tense and Aspect

We reduce Tense and Aspect to the same set of semantic primitives by assuming that since Tense has been analyzed as relating two times, Aspect should also be analyzed as relating two times. This is precisely the proposal made by Klein (1995).

(6) Aspect, like Tense, relates (orders) two times. Aspect relates the Event-time to the Assertion-time. Tense relates the Assertion-time to Speech-time.

Note that for Klein, Tense does not directly order the Event-time with respect to the Speech-time. In this respect, he concurs with Reichenbach (1947), Hornstein (1990), Giorgi and Pianesi (1991), and Thompson (1995), for whom the relation between the Speech Time (S) and the Event Time (E) is always mediated by a Reference Time (R).

4.3.2
A Uniform Syntactic Approach to Tense and Aspect

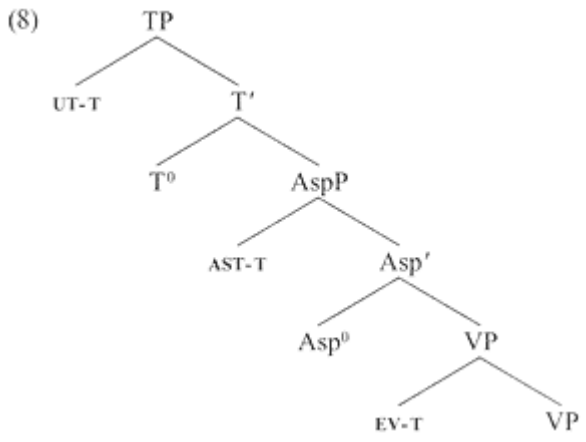
Having reduced Tense and Aspect to the same set of semantic primitives, we are now in a position to reduce them to the same set of syntactic primitives. Zagona (1990) and Stowell (1993) capture the idea that Tense relates two times syntactically by breaking it downstructurallyinto its semantic components. Once we assume that both Tense and Aspect relate two times, then *both Tense and Aspect can be broken downstructurallyinto the same set of semantic components*.

Tense has been analyzed as a head that projects a maximal projection (TP) and takes two time-denoting phrases as arguments. We propose that Aspectlike Tenseis a head that projects a maximal projection (AspP) and takes two time-denoting phrases as arguments. For Stowell, the head of TP is a *spatiotemporal* predicate that establishes a temporal ordering relation between its two time-denoting arguments. We propose that the head of AspP is also a *spatiotemporal* predicate that establishes a temporal ordering relation between its two time-denoting arguments. (See sections 4.7 and 4.8 for an extensive discussion of the typology of predicates that can head AspP.)

Under this proposal, summarized in (7) and illustrated in (8), the syntax and semantics of Tense mirror the syntax and semantics of Aspect.

(7) *The phrase structure of Tense and Aspect*
Both Tense and Aspect are dyadic spatiotemporal ordering predicates taking time-denoting phrases as arguments. The external argument of Aspect (Asp0) is a reference time (the AST-T); its internal argument is the time of the event denoted by the VP (EV-T). The external argument of Tense (T0) is a reference time (UT-T); its internal argument is the AST-T.

(8)



Recall that the role of Aspect is to focus a time interval in the temporal contour of the event described by a sentence. The time focused by Aspect is the Assertion Time, following Klein (1995). Why does Aspect pick out (focus) an interval in the time of the event denoted by the VP? Because Asp0 is a spatiotemporal head that establishes a topological relation between the AST-T and the EV-T (see section 4.7). This relation can be one of ordering (e.g., the AST-T is ordered AFTER or BEFORE the EV-T) or inclusion (e.g., the AST-T is ordered WITHIN the EV-T). Tense then orders the time interval focused by Aspect (the AST-T) with respect to the UT-T. 2

If the strict parallel we establish between Tense and Aspect is correct, then temporal and aspectual relations can be reduced to a simple schema: predicates establishing a spatiotemporal relation between two time-denoting arguments.

4.4

On the Dual Role of Some Functional Heads

Under the syntactic theory of temporal relations presented here, the external temporal argument of T0 is base-generated in the specifier of T0, as proposed by Zagona (1990) and Stowell (1993). In this regard, we depart from the standard theory where the specifier of T0 is taken to be the position where the Extended Projection Principle (the nominal features of Tense) and/or nominative Case is checked (Chomsky 1995 and references therein). Likewise, under our proposal, the specifier of Asp0 (which under certain analyses is taken to be the position where some nominal features are checked) hosts the external argument of Asp0. We

assume here that Tense and Aspect are heads that have dual syntactic functions, as has been argued for n and P (see Chomsky 1995). As lexical heads, they have argument structure that they project into the syntax. As functional heads, they enter (at least in some cases) into feature-checking relations with arguments that have been displaced to their checking domain. If this approach is correct, it suggests that syntactic projections make use of multiple specifiers to fulfill the multiple functions of their heads (Koizumi 1994, Ura 1994, Chomsky 1995). Thus, the inner specifier of TP would be the canonical position for the external (temporal) argument of Tense whereas the outer specifier of TP would be the landing site of the pied-piped phrase whose features are checked by Tense. The same line of reasoning would apply to Asp0 if this head is shown to have a role in checking nominal features.

4.5
The Grammar of Temporal Relations

We have proposed that the syntax and semantics of both Tense and Aspect can be derived from a uniform theory of the grammar of temporal relations. In the following two subsections we illustrate our proposal with an analysis of two Aspects: the Progressive and the Perfect. Our analysis is summarized in (9).

(9) Both T0 and Asp0 are spatiotemporal ordering predicates. The head of TP is a temporal ordering predicate with the meaning of AFTER for Past Tense and WITHIN for Present Tense (Stowell 1993). The head of AspP is a spatiotemporal ordering predicate with the meaning of AFTER for Perfect Aspect and WITHIN for Progressive Aspect.

We argue that the proposal in (9) has two related conceptual advantages over alternative analyses of the Progressive and the Perfect. First, it derives the compositional interaction between Tense and Aspect without resorting to ad hoc diacritics or features (e.g., [+/ completed], [+/ perfective]). In particular, within a Reichenbachian system of tenses, diacritics must be introduced in order to distinguish between the simple past and the past progressive (for instance), as shown in (10). The arrow over E in (10b) indicates that the designated event is not punctual but spreads over time.

(10) a. Past: E, R__S
b. Past progressive: $\vec{E}$, R__S

Any model based on Reichenbach's three times can easily integrate the Perfect into a system of Tense representation by assuming that the Perfect acts like a Past Tense: it locates the Event Time in the past with respect to the Reference Time. Our proposal that both Past and Perfect are spatiotemporal predicates with the meaning of AFTER also captures this idea. The tricky question for all approaches is how to integrate the Progressive into a system of Tense representation. The thesis we defend here is that the Progressive acts like a Present Tense: both Present Tense and Progressive Aspect are spatiotemporal predicates with the meaning of WITHIN. The theory we propose thus uniformly reduces Past, Present, Progressive, and Perfect to the same set of primitives: they are spatiotemporal predicates ordering two times. A further advantage of our proposal is that it derives, without additional stipulations, the syntax and semantics of recursive Aspect; further, it can explain why certain combinations of Aspects are unattested (section 4.6.2).

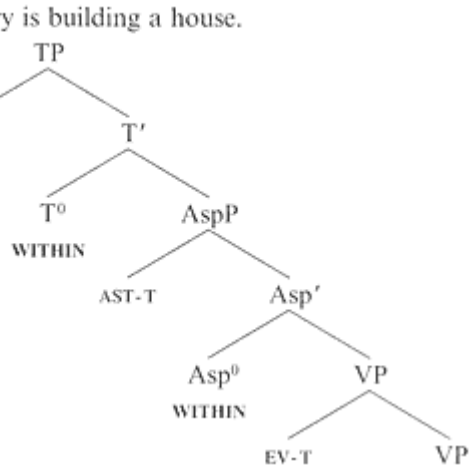
4.5.1
The Phrase Structure of the Progressive Aspect

Consider the progressive sentences in (11a) and (12a). They describe a process: Henry is/was *in* the process of building a house. Their phrase structure is given in (11b) and (12b).

(11) *Present Progressive*

a. Henry is building a house.

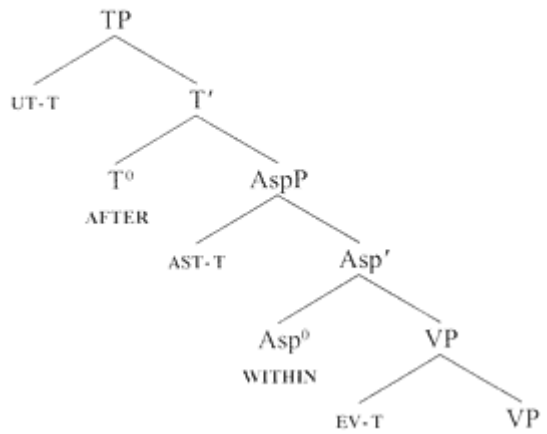
b.



(12) *Past Progressive*

a. Henry was building a house.

b.



Progressive Aspect is a spatiotemporal predicate with the meaning of WITHIN: it orders the AST-T WITHIN the EV-T. It thus picks out a time contained WITHIN the interval defined by the EV-T. Present Tense in (11b) is also a spatiotemporal predicate with the meaning of WITHIN: it orders the UT-T WITHIN the AST-T. (11b) thus focuses a subinterval of the time that defines the event of building. This subinterval is located in the present because it contains the UT-T. The single difference between (11b) and (12b) is the Tense of the sentence: present versus past. The head of TP in (12b) is the spatiotemporal predicate AFTER: it orders the UT-T AFTER the AST-T. (12b) thus focuses a subinterval WITHIN the interval defined by the event of building. This subinterval is itself located in the past because the UT-T is ordered AFTER this subinterval (that is, AFTER the AST-T).

Following Smith (1991), we assume that only the time interval within the time of the event *focused* by Aspect is visible to semantic interpretation (see section 4.2). The time focused by Progressive Aspect in (11b) and (12b)which we have called the Assertion Time AST-T, following Klein (1995)must be a subinterval of the event time (i.e., it must be properly contained WITHIN the EV-T). Since the AST-T does not include the end-points of the event, no assertion is made about whether the event of building culminated in (12b)even though the sentence is in the past. (12a) can thus be followed by . . . *but he never finished it/ . . . and he is still building it/ . . . and he finished it this summer*. 3

To conclude, although our model of temporal relations is based on three times (EV-T, AST-T, UT-T), it cannot be reduced to the classical Reichenbachian system (that is, E/EV-T, R/REF-T, S/UT-T). Within a Reichenbachian system, a Reference Time (R) cannot be ordered WITHIN another time; that is, a REF-T *cannot refer to a subpart of the event time*. Reichenbach is thus forced to resort to an ad hoc system of diacritics to derive the interpretation of the Progressiveas illustrated in (10b), repeated here, where an arrow over E was used to indicate that the event is not punctual but spreads over time. Note that this representation cannot explain why no assertion is made about whether or not the described event culminated since the whole event is ordered in the past (i.e., before S) in (10b).

In contrast, we have integrated the Progressive Aspect into a system of temporal representationand uniformly derived the compositional interaction of Tense and Progressive Aspectwithout resorting either to features such as [+/ completed, +/ perfective] or to ad hoc diacritics. In a progressive sentence the event denoted by the VP is presented without initial or final bounds because Asp0 establishes a topological relation between the time to which the assertion of the sentence is confined and the event time: it orders the AST-T WITHIN the EV-T. Thus, the AST-T in (11b) and (12b) refers to a subpart of the EV-T that includes neither the initial nor the final bound of the event. Under this analysis the Progressive acts like a Present Tense: both Present Tense and Progressive Aspect are spatiotemporal predicates with the meaning of WITHIN.

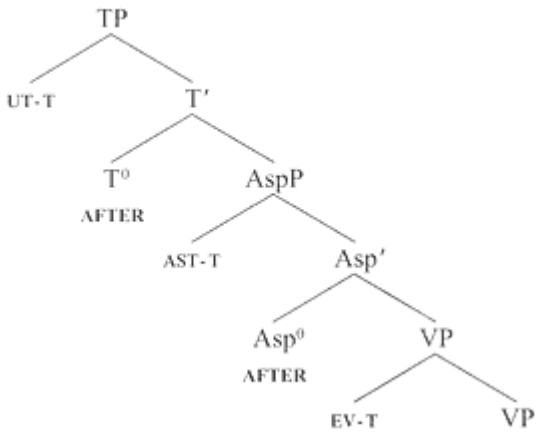
4.5.2
The Phrase Structure of the Perfect

We now integrate the English Perfect (*have* V *-ed*) into our system of Tense representation by analyzing the Perfect as a spatiotemporal predicate with the meaning of AFTER. Under this analysis the Perfect Aspect acts like a Past Tense: both Past and Perfect are spatiotemporal predicates with the meaning of AFTER. The phrase structure we propose for the Perfect is illustrated in (13b) and (14b).

(13) *Past Perfect*

a. Henry had built a house.

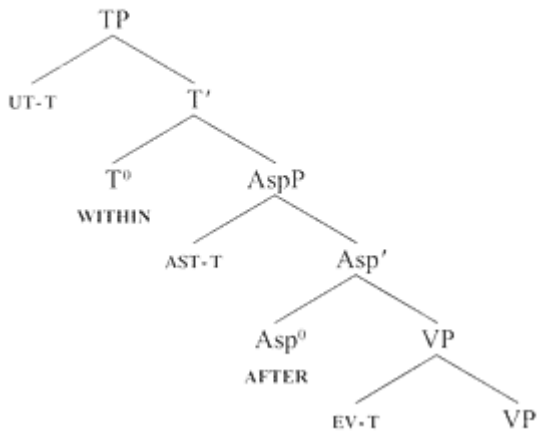
b.



(14) *Present Perfect*

a. Henry has built a house.

b.



Perfect Aspect is a spatiotemporal predicate with the meaning of AFTER. It orders the AST-T AFTER the EV-T: it thus picks out a time AFTER the interval defined by the EV-T. (13b) and (14b) differ in one single respect: (13b) has Past Tense whereas (14b) has Present Tense. Past in (13b) orders the UT-T AFTER the AST-T. Since the UT-T follows the AST-T, which itself follows the EV-T, the event of building in (13a) is understood as completed prior to some past reference time (our AST-T). Present in (14b) orders the UT-T WITHIN the AST-T. The event of building in (14a) is thus understood as

completed prior to a present reference time (i.e., prior to the AST-T, which contains the UT-T). This analysis captures the idea that the Perfect acts like a Past Tense: it locates the EV-T in the past with respect to a reference time (our AST-T) by ordering the AST-T AFTER the EV-T.

As was the case with the Progressive, we have derived the viewpoint of the Perfect (i.e., the event is presented as completed) without resorting to ad hoc features such as, [+/ completed, +/ perfective, +/ terminated]. The event is viewed as completed because the Perfect orders a reference time (our AST-T) AFTER the EV-T and, hence, AFTER the time that defines the final bound of the event. 4

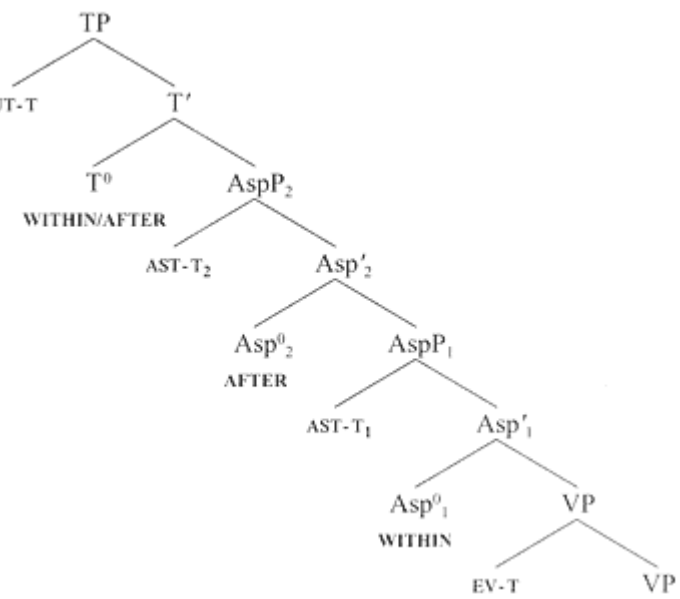
4.6
Aspect Recursion

We now show how our theory of temporal relations passes a crucial test: it derives, without any additional stipulation, the syntax and semantics of recursive aspect.

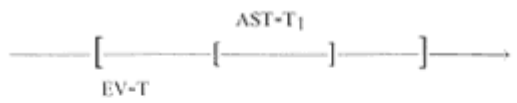
4.6.1
The Perfect of a Progressive

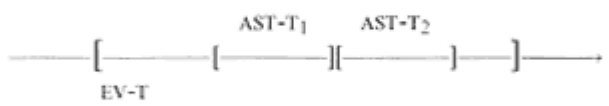
A perfect progressive sentence such as *Rosa has/had been reading Move* a will have the phrase structure in (15). There are three temporal heads in (15), each ordering two temporal arguments.

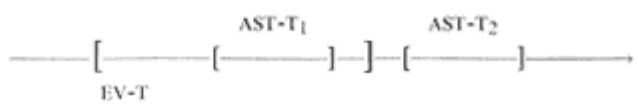
(15)



The lowest head (Asp_1^0) is the spatiotemporal predicate WITHIN; it orders its external argument (AST-T1) WITHIN the EV-T. It thus picks out a time contained WITHIN the EV-T, as illustrated by the schema in (16). "Rosa PROG read *Move* a" thus focuses a subinterval of the event of reading *Move* a by Rosa. Moving up the tree, the next (Asp_2^0) head is a spatiotemporal predicate with the meaning of AFTER. It orders its external argument (AST-T2) AFTER its internal argument (AST-T1). Thus, Asp_2^0 picks out a time AFTER the interval defined by AST-T1, as illustrated by the schema in (17). Note that AST-T2 denotes a time interval AFTER a *sub*-interval of the event time (i.e., AFTER AST-T1). Hence, it can denote a time that falls before its culmination as in (17a), falls after the culmination of the event as in (17b), or contains its culmination. 5

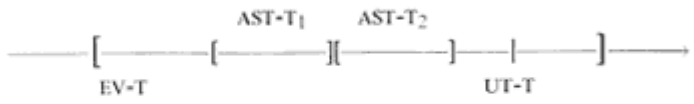
(16)


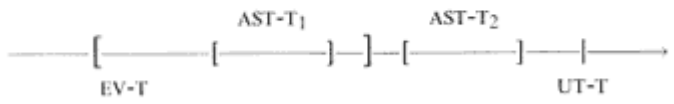
(17) a.


b.


Once we add Tense, AST-T2 will itself get located with respect to the UT-T. Tense orders its external argument (UT-T) AFTER or WITHIN its internal argument (AST-T2). Note that although the UT-T is ordered by Tense with respect to AST-T2, it is unordered with respect to the final bound of the event. Thus, when the UT-T is ordered by Past AFTER AST-T2, it can fall either before the culmination of the event as in (18a) or after the culmination of the event as in (18b). Likewise, when the UT-T is ordered by Present WITHIN AST-T2, it can fall either before the culmination of the event as in (19a) or after the culmination of the event as in (19b).

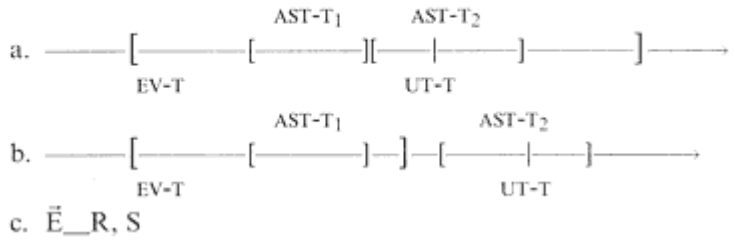
(18) *Past Perfect Progressive*
Rosa had been reading *Move* α.

a.


b.


c. Ē__R__S

(19) *Present Perfect Progressive*
Rosa had been reading *Move* α.



For comparison, in (18c) and (19c) we have given the representations of the Past Perfect Progressive and the Present Perfect Progressive, respectively, within a Reichenbachian system. Note first that our AST-T2 in (18) and (19) plays the role of Reichenbach's reference time (R) in (18c) and (19c). Just as Reichenbach's reference time (R) mediates between two times (the event and speech times E and S), so our AST-T2 mediates between two times: a subinterval of the EV-T (i.e., AST-T1) and the UT-T. Thus, in (18a c) the speech time follows a reference time (our AST-T2) whereas in (19a c) the speech time overlaps with a reference time (our AST-T2). However, whereas Reichenbach's reference time is ordered AFTER the Event Time itself and a diacritic over E is used to indicate that the designated event (E) is not punctual but spreads over time, our AST-T2 is ordered only after a subinterval of the event time (i.e., AFTER AST-T1). The fact that AST-T2 is thus unordered with respect to the final bound of the event explains why no assertion is made in a perfect progressive sentence about whether or not the event described by the sentence culminated before UT-T. Thus, in the past perfect progressive sentence *Rosa had been reading Move* a *(when Lou arrived)*, no assertion is made about whether Rosa completed her reading. Either Rosa never completed it, or she completed it after UT-T as in (18a) or before UT-T as in (18b). The same applies to the present perfect progressive sentence *Rosa has been reading Move* a. If this sentence has the temporal schema in (19a), where the UT-T is ordered WITHIN AST-T2 but before the final bound of the event, then Rosa is still reading *Move* a at UT-T. In contrast, under the schema in (19b) Rosa finished her reading before the UT-T (e.g., Rosa has been reading *Move* a all year; she only finished it a week ago/she still is reading it).

4.6.2
Constraining Aspect Recursion

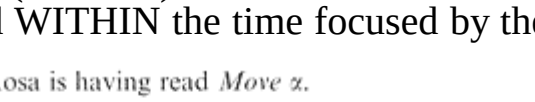
We have demonstrated how our theory of the grammar of temporal relations derives the semantic interaction of two Aspectsnamely, the

Progressive and the Perfect. Note, however, that not all the logical combinations of these two Aspects are possible. In particular, although the Perfect of the Progressive is grammatical, both the Progressive of the Perfect (e.g., **Rosa is having read Move* a) and the Progressive of the Progressive (e.g., **Rosa is being reading Move* a) are ungrammatical. 6 We now go through the respective derivations of these illicit combinations of aspects, in order to see whether we canwithin a minimalist approachexplain when and why aspect recursion is ungrammatical.

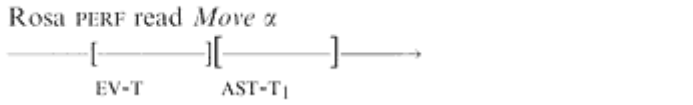
We start with the derivation of the Progressive of the Progressive, again proceeding bottom up. We have two aspectual heads, both of which have the meaning of WITHIN. The lower head $(\text{Asp}_1^{\emptyset})$ has two arguments: AST-T1 and EV-T. $\text{Asp}_1^{\emptyset}$ orders the former WITHIN the latter and thus picks out a time contained WITHIN EV-T, as illustrated in (20a). The higher head ($\text{Asp}_2^{\emptyset}$) takes this time (AST-T1) as its internal argument and orders its external argument (AST-T2) WITHIN AST-T1, as illustrated in (20b). The time focused by the higher Progressive Aspect (i.e., AST-T2) is thus itself a time contained WITHIN the time focused by the lower Progressive Aspect (i.e., AST-T1).

(20) *Rosa is being reading *Move* α.

a. Rosa PROG read *Move* α



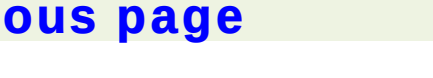
b. *Rosa PROG PROG read *Move* α



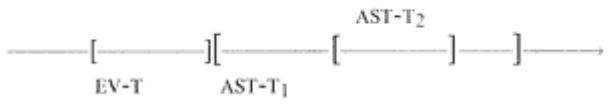
Now consider the derivation of the Progressive of the Perfect in (21). The lower head $(\text{Asp}_1^{\emptyset})$ orders its external temporal argument (AST-T1) AFTER its internal argument (the EV-T). $(\text{Asp}_1^{\emptyset})$ thus picks up a time AFTER the EV-T, as illustrated in (21a). The higher head ($\text{Asp}_2^{\emptyset}$) takes this time (AST-T1) as its internal argument and orders its external argument (AST-T2) WITHIN AST-T1. The time focused by he Progressive Aspect (AST-T2) is thus itself a time contained WITHIN the time focused by the Perfect Aspect (AST-T1), as illustrated in (21b).

(21) *Rosa is having read *Move* α.

a. Rosa PERF read *Move* α



b. *Rosa PROG PERF read *Move α*

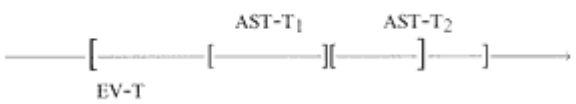


Why does Aspect recursion in (20) and (21) yield ungrammatical outputs? Crucially, the topological relation between the two Assertion Times in (20b) and (21b) is the same: AST-T2 is fully contained within AST-T1. This generalization is not a trivial result: since the Perfect of a Progressive and the Progressive of a Progressive yield (formally) the same output within our system of temporal representation, it is now possible to uniformly explain their ungrammaticality. The question then is why the containment relation between the two Assertion Times in (20b) and (21b) is illicit. Recall that the role of Aspect is to convey a viewpoint on the event denoted by VP. The first Aspect in either (20a) or (21a) provides a viewpoint by focusing a time interval either WITHIN or AFTER the event time. However, the addition of an Aspect to either (20a) or (21a) does not provide a new (distinct) viewpoint on the situation since the time interval picked out by the additional Aspect (AST-T2) is itself properly contained withinand, as such, is nondistinct fromthe time interval already picked out by the lower Asp0 (AST-T1). 7 In sum, the additional Aspect in either (20) or (21) is semantically vacuous: it does not shift the viewpointthat is, it does not focus a time interval that is itself distinct from the time focused by the first Aspect. The following descriptive generalization thus emerges: Aspect recursion is illicit if the addition of an Aspect does not shift the viewpoint or, alternatively, vacuously shifts the viewpoint, as stated in (22).

(22) *Constraint on Aspect Recursion*
No vacuous viewpoint shifting is allowed.

Crucially, the Perfect of the Progressive does not violate this constraint. The addition of the Perfect Aspect to the Progressive Aspect is not semantically vacuous: it provides a new viewpoint on the situation since the time interval picked out by the additional Aspect (AST-T2) is distinct from the time interval picked out by the lower Asp0 (AST-T1), as illustrated in (23).

(23) Rosa PERF PROG read *Move α*



The restriction in (22) has an economy flavor and is reminiscent of other conditions such as Grimshaw's (1993) "Minimize Structure," which states that no superfluous projection can form part of a derivation (see also Boskovic * 1997 and Fox 1995 for related discussion). The aspectual restriction on possible combinations of aspects given in (21) seems to fall under the same type of economy conditions.

4.6.3
Conclusion

We have analyzed Tenses and Aspects as predicates expressing topological relations between two times: inclusion (progressive/present), subsequence (perfect/past), and precedence (prospective/future) (see Leder 1991, Klein 1995). If Tenses and Aspects are semantically predicates taking times as arguments, then they should, as any predicate would, project their (temporal) arguments onto given argument positions in the syntax.8 The proposal that the time intervals ordered by Tense and Aspect are base-generated in the specifier positions of temporal heads is thus the null hypothesis.

To conclude this section, we would like to point out some advantages of our model over some alternative syntactic models that represent Reichenbach's three time pointsspeech time (S), reference time (R), and event time (E)as directly linked to the heads T0 and Asp0.9 Within our model the UT-T, AST-T, and EV-T are not associated with syntactic heads. Rather, they are temporal arguments, base-generated in specifier positions and ordered with respect to each other by spatiotemporal predicates. This proposal allows us to formally represent *inclusion* relations between two times. It is precisely because our model can represent inclusionalongside subsequence and precedencethat it can derive the semantics of the Progressive and of recursive aspects such as the Perfect Progressive. In particular, recall that Progressive Aspect picks up a *subinterval* of the event time by ordering the AST-T within the EV-T. It is this subintervaland crucially not the whole eventthat is ordered by Tense with respect to the UT-T. In other words, in order to explain why in a (perfective) progressive sentence, no assertion is made about the culmination of the eventeven when the sentence is in the pastwe must be able to pick out a subinterval of the event time. We achieve this straightforwardly by analyzing the UT-T, AST-T/REF-T, and EV-T as time-denoting arguments, projected onto argument (specifier) positions and ordered with respect to each other by the spatiotemporal heads (Tense or Aspect) of which they are (respectively) arguments.

A further advantage of our proposal over other analyses is that it elegantly explains the typology of temporal systems found across languages, to which we turn in the following sections.

4.7
On the Nature of Tense and Aspect as Ordering Predicates

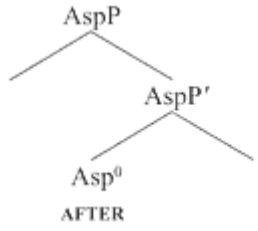
We have integrated Aspect into a system of temporal representation and uniformly derived the compositional interaction of Tense and Aspect. Progressive Aspect was analyzed as AST-T WITHIN EV-T; Perfect Aspect as AST-T AFTER EV-T. Conversely, Prospective Aspect, which we discuss in sections 4.7.2 and 4.8.2, would be analyzed as ordering the AST-T BEFORE the EV-T.

The question arises, however, of why we define Perfect as a spatiotemporal predicate ordering one time AFTER another, rather than ordering one time BEFORE another. Conversely, why would we define Prospective Aspect as ordering one time BEFORE another, rather than ordering one time AFTER another? This question does not appear to be relevant in a Reichenbachian system of Tense representation. Thus, the Reichenbachian Present Perfect in (24) can be read either as ordering the EV-T BEFORE the REF-T or as ordering the REF-T AFTER the EV-T.

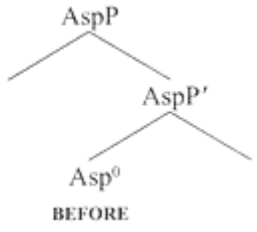
(24) Perfect: EV-T__REF-T, UT-T

However, if Aspects and Tenses are spatiotemporal predicates that project their (temporal) argument structure in the syntax, then the question of why Perfect means AFTER and not BEFORE (and conversely why Prospective Aspect means BEFORE and not AFTER) has theoretical and empirical consequences. How do we decide in a principled way whether Perfect has the phrase structure in (25a) or (25b)?

(25) a.



b.



Moreover, why is the EV-T the internal argument of Asp0 and not its external argument? Likewise, why is the AST-T the external and not the internal argument of Asp0? The same questions arise with respect to the definition of the three basic Tenses. We will derive the proposal that Perfect means AFTER and not BEFORE and, conversely, that Prospective

Aspect means BEFORE and not AFTER from a universal principle governing the mapping of spatiotemporal relations (defined in terms of Hale 1984) onto phrase structure.

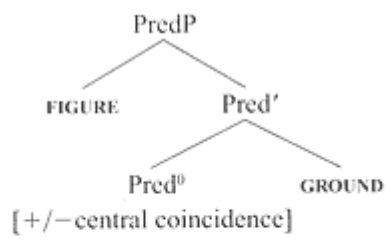
4.7.1
Hale 1984: Central versus Noncentral Coincidence

Hale (1984) argues that spatial and temporal relations can be understood in terms of the location of an entity (the *figure*) with respect to a place (the *ground*). Spatiotemporal relations are then defined in terms of a single basic semantic opposition: *central* versus *noncentral coincidence* in the location of the figure with respect to the ground. Central coincidence indicates that the location of the figure coincides with the ground. In English this relation is expressed by prepositions such as *on*, *in*, *at*, *along*, *over*, and *through*. Conversely, noncentral coincidence (expressed in English by prepositions such as *from*, *out of*, *(up) to*, *onto*, and *into*) indicates that the location of the figure either ends or begins at the ground.

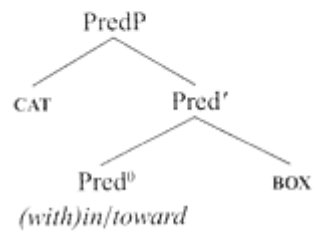
4.7.2
Why AFTER and Not BEFORE?

We propose that spatiotemporal relations are mapped onto phrase structure according to the universal phrase structure schema in (26). This schema breaks down topological relations syntactically into their primitive components: the figure (mapped onto the specifier position), the ground (mapped onto the complement position), and the spatiotemporal predicate (defining the topological relation between the figure and the ground), as illustrated in (27).

(26)



(27)



In (27) the head WITHIN orders its external argument (*cat*) WITHIN its internal argument (*box*). Thus, the figure is always the external argument of the ordering predicate and the ground, its internal argument.

We can now recast our initial question as follows: why is the time picked up by Aspectthat is, the AST-Tthe figure and not the ground? Conversely, why is the EV-T the ground and not the figure? To answer this question, consider (28)

(28) *Central coincidence*
a. [I //////// F]
Noncentral coincidence
b. ////////[I F]
c. [I F]////////

Recall that the role of Aspect is to focus (locate) an interval in the internal temporal constituency of the event. Thus, in (28a) Progressive Aspect focuses an interval WITHIN the time interval defined by the initial and final bounds of the event. The containment relation in (28a) exactly parallelsin the temporal realmthe containment relation in the spatial realm illustrated in (27). The figure (the AST-T) is contained WITHIN the ground (the EV-T) just as the figure (the cat) is contained WITHIN the ground (the box). This relation of containment is one of central coincidence. By analogy, in (28b,c) the EV-T will be defined as the ground, the interval selected by Aspect (i.e., the AST-T) as the figure, and the ordering relations in (28b,c) as relations of noncentral coincidence between the figure and the ground. In sum, once we define spatiotemporal relations in terms of the location of a figure with respect to a ground, the AST-T is the figure located by Aspect with respect to the groundthat is, the EV-T. This in turn entails that Perfect in (28c) means AFTER and not BEFORE: it picks out an interval (the figure) AFTER the final bound of the event (the ground). Likewise, Prospective Aspect in (28b) means BEFORE and not AFTER: it picks out an interval (the figure) BEFORE the initial bound of the event (the ground). By the same line of reasoning, we can explain why Past means AFTER and not BEFORE and, conversely, why Future means BEFORE and not AFTER. Note that if this proposal is correct, then the grammar of temporal relations can itself constrain the logically possible hierarchical relations between TP, AspP, and VP: VP (which introduces the EV-T) must be the complement (the ground) and not the specifier (the figure) of Asp0. Similarly, AspP will have to be the complement of TP. 10

4.8
On the Typology of Aspectual Systems

We now demonstrate that the proposal that Aspects are spatiotemporal predicates of [+/ central coincidence] is empirically verified: the spatiotemporal predicates that we have postulated as the abstract head of AspP surface overtly across languages. Our analysis further explains (a) why verbs of stance, posture, or location can express Progressive Aspect; (b) why verbs of motion toward the ground are used to express Prospective Aspect; and (c) why verbs of direction or motion from the ground are used to express Perfect Aspect.

4.8.1
Central Coincidence

4.8.1.1
Prepositions of Central Coincidence

We have defined the Progressive as a spatiotemporal predicate expressing central coincidence between the location of the figure and the ground. In English this relation is expressed by prepositions such as *on*, *in*, *at*, *along*, *over*, and *through*. *On*, for instance, specifies that the center of the figure coincides with the center of the ground; *along*, in turn, specifies that the center of the figure's trajectory or linear arrangement coincides with the ground (cf. Hale 1984). Prepositions of central coincidence are very commonly used to form progressive sentences crosslinguistically, as illustrated in (29) (31). For instance, in Basque the progressive is formed by combining the verb *ari* 'engage' with the nominalized form of the verb suffixed with the inessive (locative) postposition *-n* '(with)in/on/at'. In French the progressive is formed by combining an infinitive verb with a complex locative expression (translatable as 'engaged in') that can be decomposed as 'in'+ 'along', both of which are prepositions of central coincidence. In Dutch, in turn, a locative preposition ('at') combines with an infinitive verb.

(29) Maritere euskara ikas-te-*n* ari da. [V + t(z)e + LOC ARI]
Maritere Basque learn-NOM-*LOCin* engage AUX(are)
'Maritere is (engaged in) learning Basque.'

(30) Zazie est *en train de* jouer. [BE + en train de + INFINITIVE]
Zazie is *in along of* play
'Zazie is (engaged in) playing.'

(31) Ik ben het huis *aan* het bouwen. [BE + AT + INFINITIVE]
I am the house *at* the build
'I am (at the) building (of) the house.'

According to Bybee, Perkins, and Pagliuca (1994), constructions of the type "the subject is AT verbing" are crosslinguistically a very common source for progressives. These authors suggest that at first the locative preposition was used to indicate that the subject was involved in an activity at a certain location, where spatial location implies temporal location. The progressive meaning evolves when the spatial meaning is lost and only the temporal meaning remains.

Historical evidence for our analysis of the progressive is provided by Vlach (1981) and Bybee, Perkins, and Pagliuca (1994). Vlach argues that the historical antecedent of the progressive in English is a construction in which a prepositioneither *on* or *at*, later shortened to *a-* as in *asleep*combined with a gerund(-ive) noun phrase, as illustrated in (32), quoted from Vlach 1981 and Bybee, Perkins, and Pagliuca 1994.

(32) a. John is *on/at* hunting/building a house.
b. He was a-coming home.

The use of prepositions of central coincidence to form progressives is not absent from Modern English either, as pointed out to us by Eloise Jelinek. For example:

(33) a. I am in the middle of washing the dishes.
b. She is at rest.

Our proposal elegantly explains why Progressive Aspect surfaces pervasively as a locative preposition of central coincidence.

4.8.1.2
Progressive and Locative 'Be'

There are two copular verbs in Spanish: *ser* and *estar*. Roughly, *ser* is used either in equational sentences or with individual-level predicates whereas *estar* is used with either locative or stage-level predicates. The copula used in progressive sentences is locative 'be', *estar*, which derives from Latin *stare* 'to stand' (Bybee, Perkins, and Pagliuca 1994).

(34) a. *Progressive*
Oihana *está* estudiando.
Oihana *is* studying
b. *Simple locative*
Oihana *está* en Oba.
Oihana *is* at/in Oba

This correlation between the copula used in progressive sentences and the copula used in locative sentences is crosslinguistically pervasive. 11 Indeed,

according to Bybee, Perkins, and Pagliuca (1994, 132), there are "no clear cases of progressives formed with a copula without a locative element." Our proposal that the Progressive is a locative immediately derives this generalization.

4.8.1.3
Verbs of Central Coincidence

We can straightforwardly explain why not only locative prepositions, but also verbs of stance, posture, or location, surface crosslinguistically to express Progressive Aspect. In particular, Bybee, Perkins, and Pagliuca (1994, 129) report that "[t]he verbal auxiliary [in progressive sentences] may derive from a specific postural verb such as 'sit', 'stand' or 'lie', or it may express being in a location without reference to a specific posture but meaning only 'be at', 'stay', or more specifically, 'live', 'reside'."

Verbs of stance or posture express relations of central coincidence. For instance, the verbs *stand* and *lie* (in, say, *Max is standing/lying on the table*) specify that the location (in this case the linear arrangement) of the figure (Max) coincides centrally with the ground (the table). If stance and posture verbs express central coincidence, then it comes as no surprise that these verbs are used to form progressives crosslinguistically, since the meaning of the progressive is precisely that of central coincidence.

To conclude this section: We have argued that the Progressive is a predicate of central coincidence. This proposal explains why *location is a necessary semantic element* in progressive sentences (Bybee, Perkins, and Pagliuca 1994). 12

4.8.2
Noncentral Coincidence

4.8.2.1
Prepositions of Noncentral Coincidence

Crosslinguistic evidence for the proposal that Perfect Aspect is the spatiotemporal predicate of noncentral coincidence with the meaning of AFTER, and not BEFORE, is provided by Bull (1960), who argues that in languages like Gaelic, the translation of *He had already sung* is literally 'he was AFTER his singing' (*Rabh sé ndiaidh seinnm*). As Bull (1960, 26) himself concludes, "These languages demonstrate that both the concept of Aspect and the concept of order produce identical results."

4.8.2.2
Directional/Motion Verbs as Aspects

Finally, we integrate verbs of motion and direction into our grammar, explaining why precisely these classes of verbs play a pervasive role in the expression of both Prospective and Perfect Aspect.

Centripetal Motion

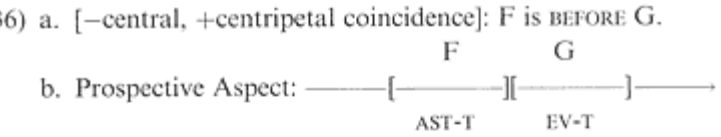
Bybee, Perkins, and Pagliuca (1994) report that motion verbs such as *go* are the most prominent lexical source for primary futures. French, for instance, has a periphrastic future formed by combining the verb *aller* 'go' in the present tense with an infinitive verb, as illustrated in (35).

(35) Zazie va partir en voyage (la semaine prochaine).
Zazie go leave on trip (the week next)
'Zazie will leave for a trip (next week).'

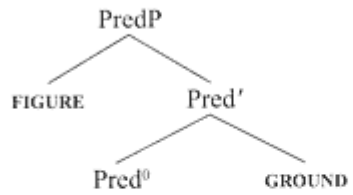
With verbs of motion the location of the figure corresponds to its trajectory and is viewed as either *ending* or *beginning* at the ground. *Go/Aller* are verbs of *centripetal* movement: they indicate movement of the figure toward the ground. Thus, in *Max is going to Paris* the trajectory of the figure (Max) is viewed as moving toward/ending at the ground (Paris). Verbs of "motion toward" are used crosslinguistically to express Prospective Aspect because they express noncentral centripetal coincidence, be it in the spatial or the temporal realm. When used in the temporal realm, these verbs specify that the location of the figure (the AST-T) ends before/at the ground (the EV-T)yielding the interpretation of Prospective Aspect, as illustrated in (36) and (37) (see also (28b)).

(36) a. [−central, +centripetal coincidence]: F is BEFORE G.

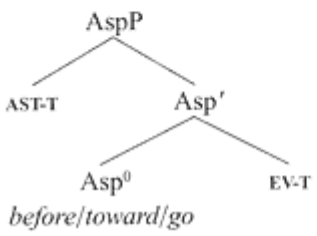
b. Prospective Aspect:



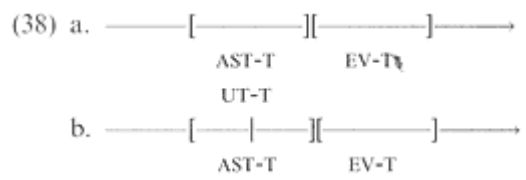
(37) a.



b.



Thus, consider the *aller* future in (35). The aspectual head ALLER (noncentral centripetal coincidence) locates the AST-T BEFORE the EV-T, as in (38a). Present Tense (central coincidence) locates the UT-T WITHIN the AST-T, as in (38b). This straightforwardly derives the interpretation of (35). It has a future event interpretation because the EV-T follows the UT-T, as shown in (38b). It has a present time relevance (Smith 1991) because the time for which an assertion is made is a time that itself contains the UT-T.



To conclude, note that the future interpretation of motion verbs arises only because these verbs express *centripetal* motion; that is, the trajectory of the figure is viewed as going toward/ending at the ground. This point is emphasized by Bybee, Perkins, and Pagliuca (1994, 268): "[S]imple movement does not evolve into future. To derive future there must be an allative component, 'movement toward', either inherent in the semantics of the verb or explicit in the construction . . . . [T]he source meaning for movement future is that 'the agent is on a path moving towards a goal'."

Centrifugal Motion

By a parallel line of reasoning, we can explain why Perfect (/Anterior) can be expressed by verbs of "motion of the figure from the ground," such as 'come from' (in Margi) or 'throw out' (in Palaung). Motion from the ground is defined as noncentral *centrifugal* coincidence. The trajectory of the figure in 'Max COME FROM home' is viewed as beginning at the ground (at home). Transposed into the temporal realm, this means that the AST-T (the figure) will start AFTER the EV-T (the ground), yielding the interpretation of a Perfect (see (28c)).

4.9 Conclusion

We have proposed a theory of temporal relations that reduces Tense and Aspect to the same set of substantive and structural primitives. Although this theory makes use of three times (EV-T, AST-T, UT-T), it is not reducible to the classic Reichenbachian system (EV-T, REF-T, UT-T). Reichenbach's REF-T cannot refer to a subpart of the EV-T. In contrast, our AST-T can be ordered WITHIN the EV-T and hence refer to a subpart of the EV-T. The thesis that Tenses and Aspects are predicates establishing topological

relations between time-denoting arguments thus derives the compositional interaction of Tense and Aspect without resorting to a stipulative system of diacritics or/and features. The syntactic theory of temporal relations proposed here elegantly integrates recursive aspect into a system of Tense representation. We have further argued that constraints on aspect recursion can be derived from economy conditions. Our theory restricts the relations that Tenses and Aspects establish between their temporal arguments to three basic topological relations, while, at the same time, explaining the wide diversity of temporal systems instantiated in natural languages.

Notes

We wish to express our deepest gratitude to Howard Lasnik. It has been our privilege to know him, to have him as a teacher, and to work with him.

We thank Gerhard Brugger, Elizabeth Currie, Henry Davis, Rose-Marie Déchaine, Ken Hale, Roland Hinterholtz, Eloise Jelinek. Alazne Landa, Toshi Ogihara, Mary Olsen, Javi Ormazabal, Phil Resnick, Tim Stowell, Juan Uriagereka, Karen Zagona, and an anonymous reviewer for comments and discussion. This project has been partly funded by SSHRCC grant 410-95-1519, by a grant from the University of the Basque Country (UPV 033.130-HA036/98), and by a grant from the Basque government (PI-1998-127).

1. Note that in Reichenbach's (1947) system of tenses. Tense does not directly order the Event Time (E) with respect to the Speech Time (S); rather, Tense orders a Reference Time (R) with respect to S.

2. For an analysis of simple tenses (without morphological aspect), see Demirdache and Uribe-Etxebarria 1998b.

3. For discussion of the imperfective paradox, see Demirdache and Uribe-Etxebarria 1996.

4. For an analysis that derives the existential, resultative, and continuative readings of the English Present Perfect uniformly from the proposal that the Perfect is the spatiotemporal predicate *after*, see Demirdache and Uribe-Etxebarria 1996, 1997.

5. For clarity of exposition, in (16) through (19) we have represented all the boundaries of the time intervals involvedincluding the Event Time. However, the schemas (18) and (19) do not exhaust the possible ordering relations between the four times represented in the schemas. Ordering UT-T after AST-T2 in (18) and (19) yields other possible ordering relations between these four intervals that we have not illustrated in the text. Keep in mind that the final bound of the event remains crucially unordered since (a) the Progressive picks up a subinterval of the EV-T that does not include its bounds and (b) it is this subinterval (AST-T1) that is ordered by the Perfect with respect to AST-T2. That is, no assertion is made about the final bound of the event. However, the final bound can be specified by supplying additional contextual information (e.g., Rosa has been reading *Move* a all summer, and she just finished it). See Demirdache and Uribe-Etxebarria 1998b

for a more exhaustive discussion of the interpretations of the perfect progressive. For discussion of the role that bare adverbs and adverbs introduced by temporal connectives play in either specifying the boundaries of the AST-T/EV-T or measuring the duration of these intervals, as well as a discussion of restrictions on the possible combinations of tenses, aspects, and adverbs, see Demirdache and Uribe-Etxebarria 1998a.

6. We set aside here the Perfect of a Perfect because it is attested in certain languages (e.g., certain dialects of French). Thanks to Gerhard Brugger for bringing this to our attention.

7. In both (20b) and (21b) AST-T1 contains AST-T2. Note that AST-T1 is distinct from AST-T2 since there is a subinterval of AST-T1 that is not properly contained within AST-T2 (i.e., AST-T1 is bigger than AST-T2). Crucially, however, the addition of an aspect does not shift the viewpoint because the time picked out by the second aspect (AST-T2) is itself nondistinct from AST-T1, since there is no subinterval of AST-T2 that is not itself already properly contained within AST-T1.

8. See also Zagona 1990, Stowell 1993, and Currie 1997 for further arguments for projecting temporal arguments onto argument (specifier) positions in the syntax.

9. In particular, Giorgi and Pianesi (1991) propose a phrase structure based on two temporal heads: the higher temporal head instantiates the relation between Reichenbach's S and R time points, and the lower head instantiates the relation between the R and E time points. In Thompson 1995 Reichenbach's three time points are associated with syntactic heads: S is associated with the head of TP, E with the head of VP, and R with the head of AspP.

10. The mapping proposed in (26) concurs with proposals made by Gleitman (1997) and Hale (1995), among others. Uriagereka (1998) explores an alternative proposal, reversing the mapping proposed in (26). In section 4.8 we present crosslinguistic support for our proposal.

11. For instance, in St'át'imcets (Lillooet Salish), progressives are formed with the verb *wa7* 'to be (busy with, involved in)' used as an auxiliary. Davis (1996, 3) argues, "The primary function of the main verb *wa7* is locative. Unlike English *be*, it is compatible only with locative complements . . . . It appears to mean not just *-be* but *be-somewhere*; in other words, just as auxiliary *wa7* supplies temporal coordinates, main verb *wa7* supplies spatial coordinates which may be fixed by locative deictics or prepositional phrases." See also Bar-el 1998 for discussion of *wa7* in SqwXwú7mish (Squamish Salish).

12. According to Bybee, Perkins, and Pagliuca (1994), Heine, Claudi, and Hünnemeyer (1991) found more than one hundred African languages with locative sources for the Progressive.

References

Bar-el, Leora. 1998. The stage vs. individual level distinction in SqwXwu7mish: A look at adverbs of quantification. Master's thesis, University of British Columbia, Vancouver.

Boskovic, Zeljko. * 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge, Mass.: MIT Press.

Bull, William E. 1960. *Time, tense and the verb*. Berkeley and Los Angeles: University of California Press.

Bybee, Joan, Revere Perkins, and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Currie, Elizabeth. 1997. Topic time: The syntax and semantics of SqwXwu7mish temporal adverbials. Master's thesis, University of British Columbia, Vancouver.

Davis, Henry. 1996. Wa7 t'u7 Wa7'. Paper presented at the Fourth Annual Salish Morphosyntax Workshop, University of Victoria.

Demirdache, Hamida, and Myriam Uribe-Etxebarria. 1996. Temporal relations. Ms., University of British Columbia, Vancouver, and University of California, Irvine.

Demirdache, Hamida, and Myriam Uribe-Etxebarria. 1997. The syntactic primitives of temporal relations. Paper presented at the Troisième Colloque International "Langues et Grammaire," University of Paris III.

Demirdache, Hamida, and Myriam Uribe-Etxebarria. 1998a. Adverbs and time boundaries. Ms., University of British Columbia, Vancouver, and University of the Basque Country.

Demirdache, Hamida, and Myriam Uribe-Etxebarria. 1998b. On the projection of temporal structure. Paper presented at the 21st GLOW Colloquium, University of Tilburg.

Enç, Mürvet. 1987. Anchoring conditions for tense. *Linguistic Inquiry* 18, 633 658.

Fox, Danny. 1995. Economy and scope. *Natural Language Semantics* 3, 283 341.

Giorgi, Alessandra, and Fabio Pianesi. 1991. Toward a syntax of temporal representation. *Probus* 3, 187 213.

Gleitman, Lila. 1997. Talk presented at the Linguistic Colloquium Series, University of Maryland, College Park.

Grimshaw, Jane. 1993. Minimal projection, heads and optimality. Ms., Rutgers University, New Brunswick, N.J.

Hale, Ken. 1984. Notes on world view and semantic categories: Some Warlpiri examples. In Pieter Muysken and Henk van Riemsdijk, eds., *Features and projections*. Dordrecht: Foris.

Hale, Ken. 1995. *Have* and *be*: Linguistic diversity in the expression of simple relations. Ms., MIT, Cambridge, Mass.

Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer. 1991. From cognition to grammar: Evidence from African languages. In Elizabeth Traugott and Heine, eds., *Approaches to grammaticalization*, vol. 1. Amsterdam: John Benjamins.

Hornstein, Norbert. 1990. *As time goes by: Tense and Universal Grammar*. Cambridge, Mass.: MIT Press.

Klein, Wolfgang. 1995. A time-relational analysis of Russian aspect. *Language* 71, 669 695.

Koizumi, Masatoshi. 1994. Layered specifiers. In Mercè Gonzàlez, ed., *NELS 24*. GLSA, University of Massachusetts, Amherst.

Kratzer, Angelika. 1995. Stage-level and individual-level predicates. In Gregory N. Carlson and Francis Jeffry Pelletier, eds., *The generic book*. Chicago: University of Chicago Press.

Leder, Harry. 1991. Tense and temporal order. Doctoral dissertation, MIT, Cambridge, Mass.

Reichenbach, Hans. 1947. *Elements of symbolic logic*. New York: The Free Press.

Smith, Carlota. 1991. *The parameter of aspect*. Dordrecht: Kluwer.

Stowell, Tim. 1993. The syntax of tense. Ms., UCLA, Los Angeles, Calif.

Thompson, Ellen. 1995. The structure of tense and the syntax of temporal adverbs. In Raul Aranovich, William Byrne, Susanne Preuss, and Martha Senturia, eds., *Proceedings of the 13th West Coast Conference on Formal Linguistics*. Stanford, Calif.: CSLI Publications. [Distributed by Cambridge University Press.]

Ura, Hiroyuki. 1994. Varieties of raising and the feature-based phrase structure theory. (MIT Occasional Papers in Linguistics 7.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Uriagereka, Juan. 1998. From being to having. In Armin Schwegler, Bernard Tranel, and Myriam Uribe-Etxebarria, eds., *Romance linguistics: Theoretical perspectives*. Amsterdam: John Benjamins.

Uribe-Etxebarria, Myriam. 1994. Interface licensing conditions on negative polarity items: A theory of polarity and tense interaction. Doctoral dissertation, University of Connecticut, Storrs.

Vlach, Frank. 1981. The semantics of the progressive. In Philip J. Tedeschi and Annie Zaenen, eds., *Syntax and semantics 14: Tense and aspect*. New York: Academic Press.

Zagona, Karen. 1990. Times as temporal argument structure. Paper presented at the conference "Time in Language," MIT, Cambridge, Mass.

Chapter 5
How Far Will Quantifiers Go?

Kyle Johnson

A method now popular for fixing the scopes of arguments involves a covert movement operation, named QR (for Quantifier Raising) by Robert May. May envisioned QR as a kind of adjunction operation, attaching an argument so affected to a phrase dominating that argument. From the surface representation in (1a), for instance, QR can fashion the representations in (1b) and (1c) by adjoining the object and/or subject argument to IP.

(1) a. [IP Someone [VP loves everyone]].
b. [IP everyone1 [IP someone [VP loves *t*1]]]
c. [IP someone2 [IP everyone1 [IP *t*2 [VP loves *t*1]]]]

As the representations in (1a,b) suggest, QR has syntactic consequences rather like those displayed by topicalization, the process that derives (2b) from (2a).

(2) a. [IP Someone [VP loves Mary]].
b. [IP Mary1, [IP someone [VP loves *t*1]]].

Setting aside the question whether in both cases the moved item adjoins to IP, it is clear that in neither (1a,b) nor (2b) is the moved item assuming a new grammatical function. In the language of Chomsky 1981, the terms moved in (1) and (2) are relocating to $\bar{A}$ positions. A common conception of the typology of movement operations has it that movement to $\bar{A}$ positions obeys a single cluster of locality conditions, typically those lumped together under Chomsky's Subjacency. As a consequence, one expectation raised by the thesis that the scope of quantifiers is determined through QR is that this cluster of locality conditions should be reflected in the size of quantifiers' scopes. Indeed, this is one of the more straight-forwardly empirical means of confirming the QR thesis.

So it is interesting that it does not meet with unqualified success. Although it is true that quantifiers may not assume a scope that reaches out of some of the islands that topicalized phrases are trapped in, this is not always the case. The scope of quantifiers is, as expected, unable to extend out of a complex noun phrase, an indirect question, or an adverbial clause, as the (a) examples in (3) (5) indicate; and as the (b) examples show, this matches the behavior of topicalization.

(3) a. Someone met the child that talked to everyone.
b. *It's Mary1 that someone met the child that talked to *t*1.

(4) a. Someone wondered whether I talked to everyone.
b. *It's Mary1 that someone wondered whether I talked to *t*1.

(5) a. Someone left the meeting before I talked to everyone.
b. ?*It's Mary1 that someone left the meeting before I talked to *t*1.

(I use the cleft construction in (3) (5) to represent the locality condition holding of topicalized phrases because of the difficulty in forming judgments about sentences created by topicalization when they have been stripped from their licensing discourses. Clefts are the closest thing to a discourse-independent topicalization, and because the locality conditions on topicalizations and clefts are identical, they lend themselves to my purposes.) But topicalized phrases are able to escape from finite declaratives with greater ease than are quantifiers. Thus, (6a) is perfectly acceptable, but (6b) does not let the embedded object gain scope over the root object.

(6) a. It's Mary1 that I told someone you would visit *t*1.
b. I told someone you would visit everyone.

It is not that the scopes of quantifiers are necessarily clause bound; the embedded quantifier in (7), for example, is able to have the root quantifier in its scope.

(7) Someone wanted to visit everyone.

Rather, it seems that finite clauses, generally, act as islands for quantifiers. In this respect, then, the scope of quantifiers is subject to a tighter locality constraint than is topicalization. 1

There are also cases where a quantifier's scope seems able to reach out of environments where topicalization cannot. In (8a), for example, the universal quantifier escapes the subject it is embedded within, as its ability to bind the pronoun indicates.

(8) a. A resident of almost every California city curses its traffic.
b. *It's Sacramento that a resident of *t* curses its traffic.

But topicalization, like other forms of Ā-movement, is unable to escape subjects in this way, as (8b) indicates. In these contexts, then, the scope of quantifiers seems to be freer than the locality conditions on topicalization would allow.

If we hold constant the thesis that Ā-movement coheres as a class with respect to locality constraints, these facts push us toward one of only a handful of conclusions. We might exploit the fact that QR is covert movement, and topicalization overt movement, and seek a difference in the way these two types of movement interact with locality conditions. Or we might abandon a treatment of quantifier scope that involves movement and seek an alternative from which the locality conditions emerge. Or, more conservatively, we might abandon the view that QR is a member of the Ā-movement family and find another class of movement operations to include it in.

This last strategy is one that Kitahara (1996) and Hornstein (1994, 1995) have championed in recent years, and it promises to offer a solution to the contrast in (6). Their proposals are spurred by the pressure Chomsky's recent writings have brought to bear on movement operations. Chomsky (1995) entertains the thesis that all movement operations are licensed by feature checking of the sort that is typical of A-movement. Kitahara and Hornstein seek a method for assimilating QR to A-movement; moreover, they argue that there are empirical advantages to doing so. One is simply that it might take us closer to identifying the locality condition on QR. I will suggest in section 5.1 that this proposal be expressed in somewhat different terms; specifically, I will propose that we equate QR with scrambling, much as Diesing (1992), Beck (1996), and many others have done, instead of with A-movement proper. Then, in section 5.2, I will explore the consequences this treatment has for the case in (8a).

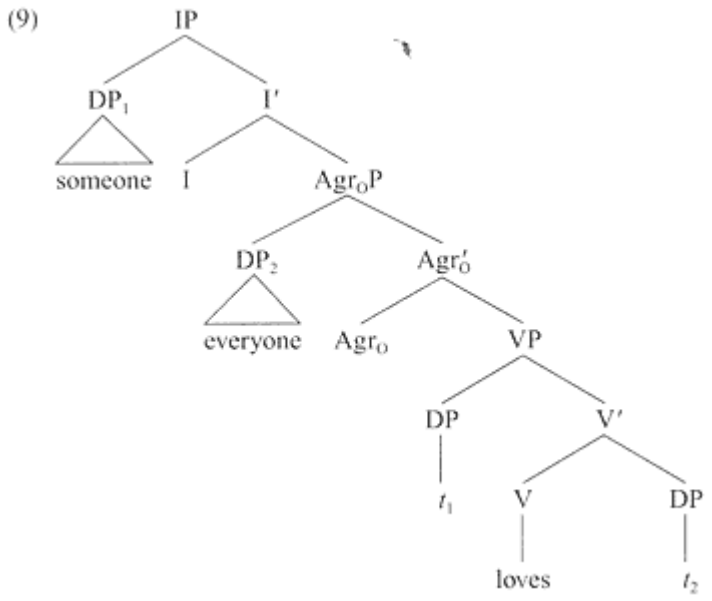
5.1
What Kind of Movement Is QR?

5.1.1
Quantifier Ambiguities through A-Movement

My discussion is based on Hornstein 1995, which has the most complete examination of the proposal. Hornstein argues that QR should be considered an instance of A-movement. More particularly, he suggests that

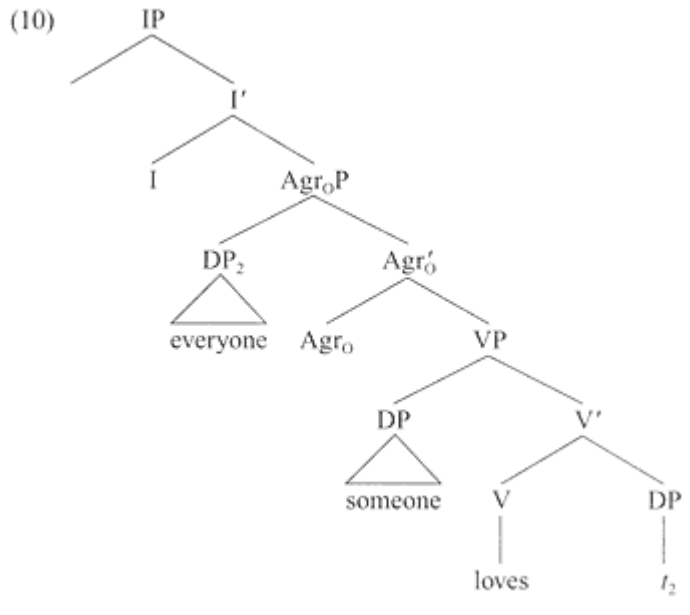
an object quantifier gains wider scope than a local subject, as in (1b), through a combination of the object's raising and the subject's lowering. 2 He adopts, first, Chomsky's (1991, 436) suggestion that structural accusative Case is assigned to DPs that covertly move into AgrOPa functional projection located above VP but lower than the functional projections that hold finiteness information. This would give to (1a) the LF representation in (9).

(9)



(I ignore throughout the effects of verb movement and the results of separating I0 into its parts.) Note that in (9) I have adopted the version of the derived-subject hypothesis that places subjects underlyingly within VP; movement of the subject into its surface Case-marked position leaves the trace shown in specifier of VP. Hornstein relies on this feature of the derived-subject hypothesis to allow for the subject quantifier to "lower" back into its underlying position. This would produce, finally, the LF representation in (10), which matches the result of QR in (1b).

(10)



That subjects may lower back into the position from which they get their q-role is known from the ambiguity of cases like (11) and has been put to use in cases parallel to (10) by Diesing (1992) and Kratzer (1995).

(11) A boy seems to be here.

As May (1985, 97ff.) shows, the reading of (11) in which *a boy* is construed as part of the clausal complement to *seem* (its most natural interpretation) arises by putting *a boy* back into that embedded clause. 3

The wide-scope reading of objects, then, comes about from two sorts of A-movements: one up (*raising*, let's call it) and one down (*lowering*). The range of situations where objects may have wider scope than local subjects should therefore be defined by the conditions that allow each of these operations to interact in the way (10) diagrams. This means that we should see the effects of constraints on both lowering and raising; objects should not be able to gain scope over subjects if either lowering doesn't allow the subject to move far enough down, or raising doesn't allow the object to move far enough up.

Hornstein suggests that this mixture of constraints is what yields the locality conditions we have witnessed above. Subjects never lower to a position below the one from which they get their q-role, so subjects will not be able to lower into embedded clauses unless they have raised from those clauses. As a consequence, quantifiers in embedded clauses will only

be able to gain wider scope than terms in higher clauses if they are able to raise into those clauses. If raising is A-movement of the sort that Hornstein envisages, then this will not be possible of objects quite generally. Hence, objects embedded within complement clauses should not be able to take scope outside those clauses, whether or not those embedded clauses are islands for Ā-movement. In this way, the puzzling islandhood of finite declarative clauses is explained. This leaves the ability of an object quantifier to take scope outside the embedded nonfinite clause in (7) mysterious, however. For this Hornstein suggests that a process of restructuring, like that which is found in Romance between 'want' and its infinitival complement, applies. This process presumably allows the object of the infinitival clause embedded under 'want' to A-move into the higher clause.

A kind of minimal pair that Hornstein discusses, which, like the contrast between (6b) and (7), involves a difference between finite and nonfinite clauses, is (12).

(12) a. Somebody believes that everyone is kind.
b. Somebody believes everyone to be kind.

In (12b), but not (12a), the universal quantifier is able to take scope over the indefinite. Hornstein's proposal relates this to the fact that the subject of the embedded clause in (12b) is related to a Case-marked position in the higher clause. This means that this subject has raised into the specifier of AgrOP of the higher clause, bringing it within reach of the indefinite. In (12a), by contrast, the subject receives Case from the embedded clause and therefore does not raise into the range of the higher indefinite. The different scope possibilities of the embedded subjects in (12a) and (12b) are connected to the ways in which they get their Case, and this is just as Hornstein's proposal would have it.

These facts, then, speak on behalf of the raising component to Hornstein's scheme. We should now look for evidence of the lowering component. Are there reasons for thinking that subject lowering is a necessary component to giving it narrower scope than a local object? Hornstein provides a series of arguments on behalf of this thesis as well, but his arguments are built on judgments that I have not been able to reproduce. 4 Nonetheless, I think there are phenomena that speak on behalf of this conjecture. But to see this, we must first confront a problem that arises with Hornstein's particular execution of his idea.

As (9) indicates, the position into which objects move on Hornstein's account is the first available specifier position above VP. As a conse-

quence, objects should have narrower scope than almost everything else in the sentence that is not within VP. In particular, objects should have scope narrower than sentence negation, which, since Pollock 1989, is thought to reside between the functional head holding tense information and the functional head mediating subject agreement. This, however, is not always true; for example, (13) is ambiguous.

(13) Gary didn't answer many of the questions on the exam.

This sentence can report that the number of questions Gary didn't answer is great (*many of the questions* has scope wider than *not*) as well as deny that Gary answered many questions (*many of the questions* has scope narrower than *not*). If objects can have scope no higher than the position given in (9), then only the second of these readings should be available.

One solution to this problem, which would preserve the hypothesis that (9) correctly shows how far objects may raise, would be to exploit the fact that the object in (13) is an indefinite. It is known that indefinites are frequently able to gain wider scope than otherwise available to quantifiers (see Abusch 1993 94 for one such case). Some authors have attempted to explain this by adopting a treatment of indefinites in which their quantificational part is quite separate from their lexical part, much along the lines of suggestions made by Heim (1982) and Kamp (1981). Kratzer (1997) (and see also Reinhart 1997), for instance, suggests that indefinites introduce a variable over choice functions, which is caught by an invisible, higher, existential quantifier. Sato-Zhu (1996) suggests that all such cases of wide-scope readings of indefinites come about in this way.

But this solution will not be general enough. Judgments become trickier, but there are examples similar to (13) that nonetheless do not involve indefinites and are therefore not susceptible to this kind of analysis. Consider, for example, (14).

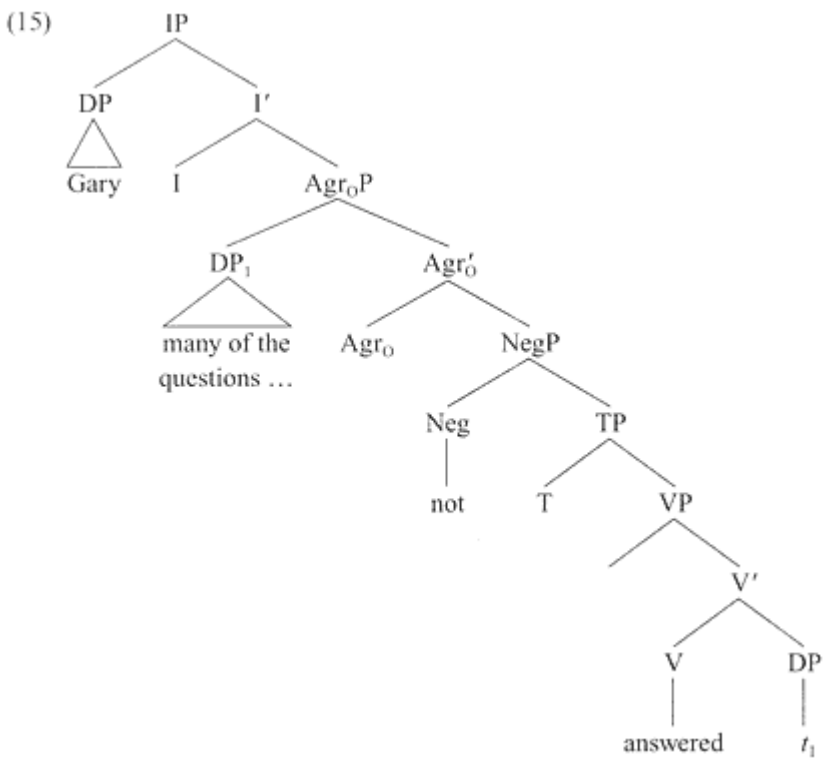
(14) a. I haven't read almost everything.
b. I can believe every one of Will's claims.

On its object wide-scope interpretation, (14a) asserts that there is almost nothing that I have read. On its object narrow-scope interpretation, it denies that I have read almost everything. That these interpretations are distinct can be seen by considering the situation in which I have read everything. In that (remarkable) situation, (14a) is false on the first (object wide-scope) interpretation, but true under the second (object narrow-scope) interpretation. In (14a), then, it must be possible for the object to take scope beyond *not*.

In (14b) we find a similar ambiguity, involving the modal *can* instead of negation. When the object has wider scope than *can*, (14b) asserts that for every one of Will's claims, there is a possibility of my believing it. When the object has narrower scope than *can*, (14b) reports that there is a possibility that I will find all of Will's claims believable. These, of course, are distinct interpretations. It could be, for instance, that I find each of Will's claims individually plausible, but know that together they are inconsistent. Inconsistency is a belief buster for me, so in this situation (14b) will be true on its object wide-scope interpretation, but not its object narrow-scope interpretation. This ambiguity indicates that *every one of Will's claims* is capable of taking scope beyond *can*.

In both cases in (14), then, we see that it is not just indefinites that are able to gain a wider scope than indicated in (9). It therefore seems necessary to let QR bring objects past the position that *not* and modals occupy. We could incorporate this fact into Hornstein's framework by placing the specifier of AgroP higher than negation, as in (15).

(15)



Then, as this representation indicates, it would be possible for objects to gain wider scope than negation by way of their movement into the specifier of AgrOP. And if T0 is the position in which modals are interpreted (perhaps because this is where they are base-generated), then this will also give objects wider scope than modals. Movement to the specifier of AgrOP must, however, now be seen as optional, or subject to lowering, if the narrower-scope reading is to be achieved.

There are other ways of wrapping Hornstein's suggestions around this fact, but for my purposes this method can act as proxy for the others. We are now ready to consider the lowering component to scope assignment. 5 Consider (16), where the subject of (13) has been replaced by an existential quantifier.

(16) Some student or other answered many of the questions on the exam.

This sentence allows *many of the questions* to have scope wider than *some student*; on this interpretation it says that many of the questions on the exam got answered (perhaps by different students). If Hornstein is correct, this reading requires *some student* to lower. A useful fact about *some student* is that it cannot remain within the scope of negation when it is in the same clause as that negation. In (17), for instance, *some student* must have scope wider than *not*.

(17) I have not met some student.

(This sentence cannot report that I haven't met anybody.) Now, this property of *some* should prevent *some student* from lowering in (18).

(18) Some student or other didn't answer many of the questions on the exam.

And indeed, it cannot be within the scope of *not*. Further, if (my modification of) Hornstein's proposal is correct, this should also mean that *many of the questions* cannot have scope wider than *some student* in (18). So this sentence should not be able to report that many of the exam questions didn't get answered (by different students). And so far as I have been able to detect, this is correct. This is not what would be expected if object quantifiers were able to raise to a position higher than the surface position of the subject, as, say, in May's original proposal. We have seen from (13) and (16) that *many of the questions* is able to take scope over negation and the subject quantifier independently; so a classical May-like

interpretation of QR should allow for their combination. Requiring subject lowering, however, will explain their failure to combine in (18).

So there is evidence for both parts of Hornstein's proposal: subject lowering is a necessary ingredient to the scope ambiguities under consideration, and object raising better mimics A-movement than it does $\bar{A}$-movement.

But there are problems too.

5.1.2
Adjuncts Can Undergo QR

What makes A-movement different from $\bar{A}$-movement is not understood, but a useful generalization about A-movement is that it is typically done in order to give arguments a "grammatical function." In Indo-European languages this usually means that arguments are brought into positions where their associated Case or agreement morphology is triggered. One might therefore extend the scenario we have just reviewed for direct objects to indirect objects without departing centrally from the proposal under consideration. In (19), for example, we could imagine that the indirect object A-moves into a phantom functional projection that licenses indirect objects. 6

(19) A different student talked to every visitor.

This will put it into a position under which the subject can lower, and the availability of an object wide-scope reading is accounted for.

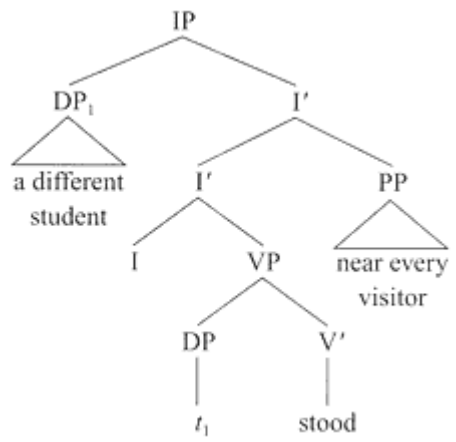
However, importing a similar sort of story for cases where adjuncts are involved, as in (20), would depart from the central proposal.

(20) A different student stood near every visitor. (can mean: for every visitor, some (possibly different) student stood near him or her)

Unless we abandon the hypothesis that grammatical functions are only assumed by arguments (or widen our criteria for argumenthood), there is no sense in which *every visitor* can gain wider scope than the subject by A-moving past it.

In order for the thesis to be preserved, it is necessary to find another way for the subject to lower into the adjunct's scope. Hornstein suggests that this can be achieved by letting adjuncts be base-generated in a position higher than that which subjects lower to. In (20), for example, *near every visitor* might be positioned as in (21).

(21)



When *a different student* lowers in (21), it will fall within the scope of *near every visitor*; and this representation could well represent the wider-scope reading for *every visitor*, if the PP headed by *near* can be construed as a quantifier of the right sort.

A feature of this approach is that the scope an adjunct has will correspond to the position that determines what it modifies. In (21), for instance, *near every visitor* takes I' as its scope, and also modifies I'. On Hornstein's account, then, adjuncts should not be able to assume a scope position that differs from the position at which their meaning is composed with the meaning of the rest of the sentence. This ability is reserved to arguments.

However, examples such as (22) show that this is incorrect. 7

(22) A different student tried to stand near every visitor.

In (22) *near every visitor* modifies the embedded clause, and not the root clause; that is, this sentence is not synonymous with *A different student tried near every visitor to stand*. And yet, *near every visitor* may have *a different student* in its scope. The correspondence expected on Hornstein's proposal between scope and modification breaks down.

It may appear that we can avoid the conclusion I have just drawn from (22) by denying that there are two predicates in this example. There is some cause to suspect this. In German and Dutch an infinitival clause following 'try' collapses in a way that is sometimes analyzed as invoking complex predicate formation (rather like the restructuring process entertained in connection with (7)).8 If English hosts a similar process, we might imagine that *tried to stand* in (22) forms a single predicate that *near*

every visitor modifies. The difference between taking *near every visitor* to modify *stand* and taking it to modify *tried to stand* will be so slight that I doubt mere reflection on the meaning of (22) will allow us to choose. So it's possible that I am mistaken in taking *near every visitor* to modify *stand* in (22); if so, (22) could very conceivably allow us to preserve the correspondence between scope and modification that Hornstein's proposal requires. A complex predicate analysis would give (22) a monoclausal representation entirely parallel to (20).

But this evasion does not seem probable, given the interpretation that (23) allows.

(23) A different boy tried to stand near every visitor only after a different girl had tried to ▲.

The elided VP in this example can be *stand near every visitor*. So far as I know, there are no grammatical instances of VP-ellipsis removing part of a predicate, 9 so in (23) we can be pretty certain that there is no complex predicate in the second conjunct made up of *tried to stand*. Nonetheless, *near every visitor* can modify *stand* in the second conjunct and still have scope over *a different girl*. So these examples do seem to pose a genuine counterexample to Hornstein's solution. Adjuncts may have a scope that differs from what they modify.

What we learn from these cases, then, is that adjuncts must be able to undergo QR. But, as noted at the outset of this section, this is a QR that cannot be construed as A-movement. The raising component of Hornstein's scheme for fixing scope interactions therefore should not be thought of as necessarily involving A-movement.

5.1.3
The Non-Clause-Bound Cases

As noted above, A-movement does not typically move the complements of verbs into a higher clause. It is this feature of A-movement that Hornstein's account relies on to capture the fact that *every book* does not have *a different student* in its scope in (24).

(24) A different student said that I had read every book.

But, as we have seen, there are situations where a quantificational complement is able to gain scope outside the clause it originates in. Hornstein suggests that restructuring is responsible for these cases. He recommends that we understand cases like (25) to involve a process that allows the object of *read* to move into a Case-marked position in the higher clause.10

(25) A different student wanted to read every book.

There is reason for thinking that restructuring is not responsible for this, however, Kennedy (1997) notes that the range of verb+infinitival constructions that allow for the wide-scope reading of the object is considerably broader than it is in the restructuring contexts of Romance (see Rizzi 1982). Thus, although verbs such as *expect*, *hope*, and *require* do not typically license restructuring in Romance, these verbs in English do permit the objects of their infinitival complements to gain widest scope.

(26) a. At least one American tourist expects to visit every European country this year.
b. At least one American tourist hopes to visit every European country this year.
c. Some government official is required to attend every state dinner.

Note also that restructuring is possible across an infinitival complementizer in Italian, but the presence of a complementizer in (25) makes the wide-scope interpretation of the embedded object unavailable.

(27) A different student wanted for you to read every book.

In (27) *a different student* does not vary as a function of the books; it cannot fall within the scope of *every book*. If restructuring were responsible for the exceptional wide scope of the object in (25), there would be no reason to expect (27) to be different.

Restructuring is not what is responsible for the wide-scope interpretation of the object in (25). If restructuring is the only process that allows A-movement to bring objects into higher clauses, then QR is behaving in these contexts unlike A-movement.

5.1.4
It's Scrambling

We have seen two difficulties with Hornstein's characterization of the raising component of scope interactions. It differs from A-movement in being able to move adjuncts, and it differs from A-movement with respect to the contexts in which it is released from its otherwise clause-boundedness.

If raising isn't A-movement, what is it? All of the attributes we have reviewed for QR can be found in the sort of movement operation that in studies of German and Dutch goes under the name of *scrambling*. I suggest therefore that we collapse the raising component of QR to scrambling.

Like QR, scrambling can move object arguments out of embedded clauses when those clauses are infinitives. This is possible even outside of

restructuring contexts; see Den Besten and Rutten 1989. (I illustrate with Dutch examples.)

(28) . . . dat Jan *Marie*1 heeft geprobeerd [*t*1 te kussen].
that Jan Marie has tried to kiss
' . . . that Jan has tried to kiss Marie.'

Like QR, this operation is blocked when the infinitival clause has a complementizer, as in (29).

(29) * . . . dat Jan *Marie*1 heeft geprobeerd [om *t*1 te kussen].
that Jan Marie has tried C0 to kiss
' . . . that Jan has tried to kiss Marie.'

And, like QR, it is prevented when the embedded clause is not an infinitival, as in (30).

(30) * . . . dat Jan *boken*1 heeft besloten [dat er *t*1 gelezen heeft].
that Jan books has decided that he read has
' . . . that Jan has decided that he has read books.'

Finally, scrambling is not restricted to arguments. Like QR, it can relocate adjuncts, as in (31). 11

(31) . . . dat Jan naast Jamie heeft geprobeerd zijn soep to eten.
that Jan near Jamie has tried his soup to eat
' . . . that Jan has tried to eat his soup near Jamie.'

So scrambling has all the characteristics we have identified in the raising component of QR. When we know what kind of movement scrambling is, we will know what QR is.

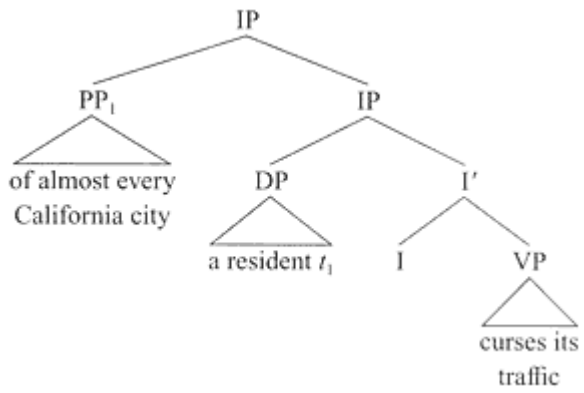
5.2 Inverse Linking

We are ready, now, to return to the example in (8a), repeated here, which poses yet another problem for Hornstein's interpretation of the raising component to his QR.

(32) A resident of almost every California city curses its traffic.

In this example of so-called inverse linking, *almost every California city* has scope over *a resident* and is able to bind *its*. This suggests an LF representation in which *of almost every California city* has moved out of the subject argument it is a part ofsomething like (33), perhaps.12

(33)



A similar situation arises in (34).

(34) I recommended a representative from almost every department to its oversight committee.

Here too, an LF representation that places the PP outside the argument it originates in is called forsomething like (35), perhaps.

(35) I recommended [from almost every department]1 [a representative *t*1] to its oversight committee.

Now, as Kennedy (1997) notes, these are not representations that can be easily viewed as manufacturable by A-movement. But they are operations that fit into the scrambling mold. German, for example, has cases rather like (33) and (35).

(36) a. . . . weil Hans über Ellipse ein neues Buch gekauft hat.
because Hans about ellipsis a new book bought has
' . . . because Hans bought a new book about ellipsis.'
b. . . . weil über Ellipse ein neues Buch erschienen ist.
because about ellipsis a new book published is
' . . . because a new book about ellipsis is published.'
c. . . . weil aus Konstanz eine Fußballmannschaft da war.
because from Konstanz a soccer team there was
' . . . because a soccer team from Konstanz was there.'

The analogy between scrambling and QR we have built in section 5.1 therefore recommends this process as the source for cases such as (32) and (34).

It is difficult to find grammatical instances of this phenomenon in German that are precisely parallel to (32); in particular, it is hard to find cases where a PP has scrambled out of the subject of a transitive verb.

Müller (1993, chap. 3) claims that this is generally blocked and offers (37) by way of illustration.

(37) a. * . . . daß darüber ein Buch den Fritz beeindruckt hat.
that about-that a book the Fritz impressed has
' . . . that a book about that impressed Fritz.'
b. * . . . daß darüber den Fritz ein Buch beeindruckt hat.
(Müller 1993, 135 136, (54))

With some grilling, however, I have managed to squeeze from speakers approval of examples such as (38).

(38) . . . weil aus Konstanz gestern ein Student ein Papier über
because from Konstanz yesterday a student a paper about
Bindung vorgetragen hat.
binding presented has
' . . . because yesterday a student from Konstanz presented a paper
about binding'

But, apparently, there are severe restrictions on scrambling from the subjects of transitive verbsrestrictions that are relaxed for subjects of intransitives (as in (36b,c)) and direct objects (as in (36a)). So if the analogy I am offering between QR and scrambling is to be maintained, we must come to understand why this dislike for scrambling from transitive subjects is not shared by QR. Let's proceed as though this understanding was at hand.

I am not equipped to make a thorough comparison of QR from argument DPs and scrambling from argument DPs, but there is at least one attribute they share. Müller (1993) shows that scrambling from DPs is subject to a condition that prevents it from affecting DPs with a genitive: the grammatical examples in (36) contrast with (39). 13

(39) * . . . daß ich über Benjamin gestern Antjes Buch gelesen habe.
that I about Benjamin yesterday Antje's book read have
' . . . that I read Antje's book about Benjamin yesterday.'
(Müller 1993, 136, (58a))

There is a parallel constraint on QR. Grammatical examples of inverse linking, such as (32) and (34), contrast with (40).

(40) a. My report about almost every California city curses its traffic.
b. I recommended my report about almost every department to its
oversight committee.

In (40) *its* cannot be bound by the (almost) universal quantifiers, as could the parallel pronouns in (32) and (34). This, of course, is an indication that QR cannot bring the quantifier out of the argument it begins within in these examples. To the extent that this shared trait is representative, it suggests that QR and scrambling are the same.

English has an overt construction that I suggest is related to the instances of scrambling relevant for bringing about inverse linking. This construction, named *extraposition from NP* by Ross (1967), associates the examples in (41) to their companions in (42). 14

(41) a. A report about almost every California city appeared today.
b. I interviewed a representative from almost every department today.

(42) a. A report appeared today about almost every California city.
b. I interviewed a friend today from almost every department.

Of course, a salient difference between extraposition from NP and the scrambling in (36) is the linear position of PP and the DP it is related to. In the scrambling cases the PP precedes the DP, whereas in extraposition the PP follows the DP. Topicalized DPs in German may strand a PP they are construed with, as in (43). (See Den Besten and Webelhuth 1990, Fanselow 1987, Müller and Sternefeld 1993, and Müller 1993, among others, for discussion.)

(43) Ein Buch hat sich Hans über Syntax ausgeliehen.
a book has self Hans about syntax borrowed
'A book, Hans borrowed about syntax.'
(De Kuthy 1998, (2))

But scrambling cannot arrange the terms in this way.15

(44) * . . . daß niemand ein Buch gestern über die Liebe gelesen hat.
that no one a book yesterday about the love read has
' . . . that no one read a book yesterday about love.'
(Müller 1993, 410, (3))

If I am right, this contrast should not be attributed to a difference in these constructions, but instead to other differences in the languages' word order.

One indication that extraposition from NP is the overt manifestation of scrambling/QR might be that it too is blocked by the presence of a genitive, something we have seen to hold of these other processes.

(45) a. *My report appeared today about almost every California city.
b. *I interviewed my friend today from almost every department.

I will set aside a direct comparison of Extraposition from NP with scrambling and turn instead to examine the many correspondences between the constraints on extraposition from NP and inverse linking. These correspondences suggest we equate them.

Ross (1967) and Akmajian (1975) show that a PP may not extrapose out of a DP that is buried within another DP. Examples such as (46), in which *on quarks* should be construed as modifying *a book*, aren't acceptable.

(46) *I bought [a report [about a book]] yesterday *on quarks*.

Similarly, a quantifier buried within the complement of a DP is prevented from gaining scope outside that DP. In (47) *it* cannot be understood as bound to *almost every*.

(47) I showed [a report [about a book on almost every department]] to
its oversight committee.

Selkirk (1977) argues that a parallel constraint is at play in preventing extraposition from the nominal complement to a partitive, as in (48).

(48) *I bought [many of the reports] yesterday *about quarks*.

Here too, we find that the wide-scope interpretation of embedded quantifiers is blocked.

(49) I showed [many of the reports about almost every department] to
its oversight committee.

As in (47), it is very difficult to understand *its* as bound to the universal quantifier in (49).

Finally Rochemont and Culicover (1990) note that extraposition from NP has an effect on the relationship that holds between adjectives in the host DP and the PP that has extraposed. This effect can be reproduced in inverse linking contexts as well. The effect can be illustrated by considering the contrast in (50).

(50) a. I dissected an alleged insect from Saturn yesterday.
b. I dissected an alleged insect yesterday *from Saturn*.

In (50a) *from Saturn* is ambiguously related to *alleged*. On one of its meanings, this sentence can assert that there was a creature from Saturn, alleged to be an insect, that I dissected. It can also assert that there was a

creature, alleged to be an insect from Saturn, that I dissected. Let's call the first interpretationthe one in which creatures from Saturn are taken to existthe *X-files* reading. 16 Interestingly, (50b) has only the X-files reading. In (50b) *from Saturn* cannot be understood to be within the scope of *alleged*. Extraposition from NP seems to fix the scope of the extraposed item to be wider than the DP it is related to.

Now consider the inverse linking reading for (51).

(51) I dissected an alleged insect from every planet in our solar system.

If we make the reasonable assumption that a particular insectlike creature may originate from only one planet (i.e., that creatures cannot be so large that they could reside on many planets at once), then (51) permits just the reading where *every planet in our solar system* takes scope over *an alleged insect*. (That is, (51) speaks about different alleged insects from each of the planets.) And, as expected, (51) has only the X-files reading found in (50b).

This last similarity between extraposition from NP and QR distinguishes these operations from $\bar{A}$-movement. As is well known, $\bar{A}$-movement allows for the moved item to be construed with the scope it would have had if it hadn't moved. This can be illustrated for the cases at hand by the topicalization example in (52).

(52) ??It's Saturn1 that I dissected [an alleged insect from *t*1].

This example is marginal because of the bounding constraints on $\bar{A}$-movement (It is generally difficult to move a term out of an adjunct PP.) But to the extent that it is grammatical, it has either of the two readings that (50a) has. In particular, *Saturn* may be construed within the scope of *alleged*, yielding the anti-X-files reading. Topicalization, then, does not disambiguate these examples in the same way that both extraposition from NP and QR seem to.

5.3 Conclusion

If we wish to reduce QR to one of the movement operations we independently need, I have argued that we should reduce it to scrambling rather than A-movement. This will still achieve the attractive goal of deriving the locality conditions on QR, without stumbling on the problems reviewed here for the A-movement thesis. Further, I have argued that the extraposition-from-NP phenomena be equated with scrambling and that they be seen as the source for the inverse linking cases.

Notes

My thanks to Sigrid Beck, Norbert Hornstein, Howard Lasnik, Winnie Lechner, Bonnie Schwartz, Satoshi Tomioka, an anonymous reviewer, and the very helpful participants of my fall 1996 syntax seminar at the University of Massachusetts. My debt to Howard, of course, extends far beyond this small paper: it is his example of clarity and rigor to which I still aspire.

1. I am ignoring here important counterexamples to this discussed by Fox and Sauerland (1996), Farkas (1995), and Farkas and Giannakidou (1996). These counterexamples involve situations where universal quantifiers gain exceptionally wide scope over indefinites; one that violates the finite-clause restriction I have just reviewed is *Yesterday, a guide made sure that every tour to the Louvre was fun*. I do not know of similar counterexamples involving overt scrambling, which I will argue QR should be assimilated to, so for the thesis of this chapter to be maintained, it must turn out that these cases do not involve QR. Fox and Sauerland make a proposal that would have this desired outcome, but it may not be general enough to handle the cases noted by Farkas and by Farkas and Giannakidou.

2. I use the term *lowering* here loosely: I mean by this any method of construing a moved element as if it were in its unmoved position. This could be done by moving the element back into its original position, by using the copy-and-delete interpretation of movement, or by giving traces a special interpretation. Kitahara's (1996) proposal, for instance, uses Aoun and Li's (1989) Scope Principle, which allows a moved element to be interpreted as if it had the scope of its trace.

3. In unpublished work Howard Lasnik has observed that some of the standard cases that are offered as evidence for lowering could instead be seen as involving an ambiguity in the subject. He notes, more particularly, that many of these cases involve situations where an indefinite is claimed to lower, and in fact indefinites might have different kinds of interpretations that alone could give the appearance of lowering. Perhaps, for instance, the supposedly lowered and unlowered readings are actually nothing more than the difference between specific and nonspecific readings of the indefinite.

I do not think all such cases of lowering can be so reanalyzed, however. Consider, for example:

(i) A boy seems not to be here.

This has an interpretation in which *a boy* falls within the scope of *not*; it can be used to report that it seems that there is no boy here. This reading demands that *a boy* fall within the scope of *not*; it cannot arise solely by virtue of an ambiguity in *a boy*. Further, because this interpretation does not place *seem* within the scope of *not*, we cannot try to get this interpretation by letting *not* have root-wide scope. Only lowering *a boy* can achieve what is needed here. That lowering might also be responsible for the ambiguity in (11) is suggested by the interaction that this ambiguity has with anaphor binding in cases such as *A boy seems to himself to be here*. See Hornstein 1995, 38 39, 160.

4. The arguments are found in Hornstein 1995, chap. 8. One difficulty with them is that they all rest on the unsupported assumption that a lowered subject returns to the VP from which its q-role comes. I don't think this is a secure assumption, however; it seems almost certain that Germanic clauses hold enough positions for subjects to be lowered to some intermediary position. Further, all but one of the arguments rests on the assumption that in VP-ellipsis contexts, the subject of the elided VP is prevented from lowering back into that VP. (This follows from Hornstein's assumption, just mentioned, that lowered subjects are necessarily restored to their VPs, and certain beliefs about the identity conditions on antecedent and elided VPs.) But, as Angelika Kratzer has pointed out to me, this will founder on Hirschbühler's (1982) example:

(i) A Canadian flag is in front of most buildings and an American flag is ▲ too.

In this case *a Canadian flag* and *an American flag* may have narrower scope than *most buildings*. On Hornstein's scheme this means that *an American flag* must have lowered. But this will be made impossible on the assumptions he adopts concerning VP-ellipsis. Thus, either his arguments that lowering is a necessary component in the object-over-subject scope reading are wrong (because they rest on incorrect presuppositions about VP-ellipsis), or (i) is a counterexample to this thesis. In fact, I think there are good reasons for rejecting Hornstein's presuppositions about VP-ellipsis.

5. The argument I report here is taken from Johnson and Tomioka 1998.

6. This is precisely what Hornstein suggests.

7. Hornstein's (1994) exploration of QR includes evidence from antecedent-contained deletion, whose resolution is widely thought to involve QR. The same problem with adjuncts emerges in these contexts as well, and Hornstein recommends the same remedy sketched in the text.

8. See Evers 1975but the phenomenon is explored in a great many more recent works as well.

9. It is not possible for VP-ellipsis to leave behind one portion of a particle construction, for example.

(i) a. *While Jerry wouldn't call the numbers out, Sally would ▲.

b. *While Jerry won't call the numbers OUT, he will ▲ IN.
(compare: While Jerry won't call the numbers OUT, he will call them IN.)

c. *While Jerry won't CALL the numbers out, he will HAND ▲.
compare: While Jerry won't CALL the numbers out, he will HAND them out.)

10. Kennedy (1997) makes the same point with respect to cases of antecedent-contained deletion.

11. My thanks to Bart Hollebrandse for help with the Dutch.

12. As before, we must understand the contribution of the preposition to not disturb the relationship between the quantifier and the material it has within its scope.

13. De Kuthy (1998) offers the nearly parallel (i) as grammatical.

(i) Über Syntax hat Karl nur Margas Buch gelesen.
about syntax has Karl only Marga's book read
'Karl read only Marga's book about syntax.'
(De Kuthy 1998, (10))

Presumably, the improvement with this example is related to the fact that the PP has topicalized here, rather than scrambled. The constraints on topicalization in German are known to be different, and more liberal, than those on scrambling.

14. What, precisely, the relationship is between an extraposed phrase and the DP it modifies has been a point of contention, as has been the similar relationship between a "scrambled" PP and the DP it is construed with in German. On some accounts there is no movement relationship between these constituentsfor extraposition from NP, see especially Culicover and Rochemont 1990 and Rochemont and Culicover 1990; and for the cases in German, see De Kuthy 1998 and references cited therein, and Müller 1993, sec. 2.3.3.4 for something of a hybrid suggestion. The matter is undecided because of lexical and other influences whose source has not been completely identified. I will continue to speak as if the relationship is movement.

15. This example is ungrammatical, of course, only on the reading where *über die Liebe* is construed as a modifier of *ein Buch*.

16. After the mysteriously popular television show.

References

Abusch, Dorit. 1993 94. The scope of indefinites. *Natural Language Semantics* 2, 83 136.

Akmajian, Adrian. 1975. More evidence for an NP cycle. *Linguistic Inquiry* 6, 115 130.

Aoun, Joseph, and Yen-hui Audrey Li. 1989. Constituency and scope. *Linguistic Inquiry* 20, 141 172.

Beck, Sigrid. 1996. *Wh*-constructions and transparent logical form. Doctoral dissertation, Universität Tübingen.

Besten, Hans den, and Jean Rutten. 1989. On verb raising, extraposition and free word order in Dutch. In Dany Jaspers, Wim Klooster, Yvan Putseys, and Pieter Seuren, eds., *Sentential complementation and the lexicon: Studies in honour of Wim de Geest*. Dordrecht: Foris.

Besten, Hans den, and Gert Webelhuth. 1990. Stranding. In Günther Grewendorf and Wolfgang Sternefeld, eds., *Scrambling and barriers*. Amsterdam: John Benjamins.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freidin, ed., *Principles and parameters in comparative grammar*. Cam-

bridge, Mass.: MIT Press. [Also in Noam Chomsky, *The Minimalist Program*. Cambridge, Mass.: MIT Press (1995).]

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Culicover, Peter W., and Michael S. Rochemont. 1990. Extraposition and the Complement Principle. *Linguistic Inquiry* 21, 23 47.

De Kuthy, Kordula. 1998. Splitting PP arguments from NPs. In Tibor Imre Kiss and Walt Detmar Meurers, eds., *Proceedings of the ESSLLI Workshop: Topics in Constraint-Based Germanic Syntax*. Saarbrücken.

Diesing, Molly. 1992. *Indefinites*. Cambridge, Mass.: MIT Press.

Evers, Arnold. 1975. The transformational cycle in Dutch and German. Indiana University Linguistics Club, Bloomington.

Fanselow, Gisbert. 1987. *Konfigurationalität: Untersuchungen zur Universalgrammatik am Beispiel des Deutschen*. Tübingen: Gunter Narr.

Farkas, Donka. 1995. Evaluation indices and scope. In Anna Szabolcsi, ed., *Ways of scope taking*. Dordrecht: Kluwer.

Farkas, Donka, and Anastasia Giannakidou. 1996. How clause-bounded is the scope of universals? Report 96-03, Linguistics Research Center, University of California, Santa Cruz.

Fox, Danny, and Uli Sauerland. 1996. Illusive scope of universal quantifiers. In Kiyomi Kusumoto, ed., *NELS 26*. GLSA, University of Massachusetts, Amherst.

Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Doctoral dissertation, University of Massachusetts, Amherst.

Hirschbühler, Paul. 1982. VP-deletion and across-the-board quantifier scope. In James Pustejovsky and Peter Sells, eds., *Proceedings of NELS 12*. GLSA, University of Massachusetts, Amherst.

Hornstein, Norbert. 1994. An argument for minimalism: The case of antecedent-contained deletion. *Linguistic Inquiry* 25, 455 480.

Hornstein, Norbert. 1995. *Logical Form: From GB to minimalism*. Cambridge, Mass.: Blackwell.

Johnson, Kyle, and Satoshi Tomioka. 1998. Lowering and mid-size clauses. In Graham Katz, Shin-Sook Kim, and Winhart Haike, eds., *Proceedings of the 1997 Tübingen Workshop on Reconstruction*. Sprachtheoretische Grund lagen für die Computer Linguistik, Tübingen.

Kamp, Hans. 1981. A theory of truth and semantic representation. In Jeroen Groenendijk, ed., *Formal methods in the study of language*. Amsterdam: Mathematical Center.

Kennedy, Christopher. 1997. Antecedent-contained deletion and the syntax of quantification. *Linguistic Inquiry* 28, 662 688.

Kitahara, Hisatsugu. 1996. Raising Q without QR. In Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson, and C. Jan-Wouter Zwart, eds., *Minimal ideas*. Amsterdam: John Benjamins.

Kratzer, Angelika. 1995. Stage-level and individual-level predicates. In Gregory N. Carlson and Francis Jeffry Pelletier. eds., *The generic book*. Chicago: University of Chicago Press.

Kratzer, Angelika. 1997. Scope or pseudoscope? Are there wide-scope indefinites? In Susan Rothstein, ed., *Events in grammar*. Dordrecht: Kluwer.

May, Robert. 1985. *Logical Form: Its structure and derivation*. Cambridge, Mass.: MIT Press.

Müller, Gereon. 1993. On deriving movement type asymmetries. Doctoral dissertation, Eberhard-Karls-Universität Tübingen.

Müller, Gereon, and Wolfgang Sternefeld. 1993. Improper movement and unambiguous binding. *Linguistic Inquiry* 24, 461 507.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20, 365 424.

Reinhart, Tanya. 1997. *Wh*-in-situ in the framework of the Minimalist Program. *Natural Language Semantics* 6, 3 28.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.

Rochemont, Michael, and Peter W. Culicover. 1990. *English focus constructions and the theory of grammar*. Cambridge: Cambridge University Press.

Ross, John R. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Sato-Zhu, Eriko. 1996. The logical interpretation of English and Japanese sentences. Doctoral dissertation, State University of New York, Stony Brook.

Selkirk, Elisabeth. 1977. Some remarks on noun phrase structure. In Peter Culicover, Thomas Wasow, and Adrian Akmajian, eds., *Formal syntax*. New York: Academic Press.

Chapter 6
Japanese Complex Noun Phrases and the Antisymmetry Theory

Keiko S. Murasugi

In the principles-and-parameters framework grammatical constructions are just artifacts: they arise from the universal principles and the parameters with values set for the particular language. As there are no universal rules or principles specific to a construction, it is quite possible that intuitively equivalent constructions in two languages have very different syntactic properties. For example, it is possible that the Japanese relative clause in (1) should be analyzed not like its intuitive English counterpart in (2a), but like the pure sentential modifier (in a pure complex NP) in (2b).

(1) [DP[Taroo-ga *e*i yonda] honi]
Taroo-NOM read book

(2) a. [DP the booki [Taroo read *e*i]]
b. [DP the evidence [that Taroo read the book]]

Ultimately, it is the theory that decides whether (1) should be analyzed like the relative construction, the pure complex NP construction, or even some other construction in English.

The purpose of this chapter is to explore the analysis of Japanese relatives within Kayne's (1994) antisymmetry theory, where a relative clause is taken to be a complex NP with a specific D-CP structure. I will argue that within this theory Japanese relatives should not be assigned the D-CP structure; instead, they have the same structure as pure complex NPs in the language, shown in (3).

(3) [DP[IP . . .] [D' D [NP . . . [N' N . . .]]]]

This means that Japanese "relatives" are more similar in syntactic properties to English pure sentential modifiers than to English relatives. In

turn, this conclusion implies that Japanese does not have relative clauses, which by definition have the D-CP structure, but only pure complex NPs.

In section 6.1 I will outline the history of the analysis of Japanese "relatives" and show that the question of whether Japanese "relatives" are indeed relatives can be raised in an informal way even from this historical perspective. In section 6.2, assuming Kayne's theory, I will discuss one of the notable syntactic properties of Japanese "relatives": the nonmovement property. I will argue that given this property, it is quite reasonable to analyze Japanese "relatives" in the same way as pure sentential modifiers. In section 6.3 I will examine some acquisition data and present an argument that Japanese "relatives" in fact should be analyzed as pure sentential modifiers with the structure in (3). Finally, in section 6.4 I will speculate on why Japanese lacks relative clauses with the D-CP structure altogether.

6.1
Does Japanese Have Relatives? A Historical Perspective

It has been widely accepted that Japanese has gapless relatives. The following example from Kuno 1973 is often cited as a typical case: 1

(4) [DP[syuusyoku-ga muzukasii] [buturigaku]]
getting job-NOM hard-is physics
'physics, which is hard to get a job in'

Accordingly, it is commonly assumed that Japanese has a relative construction with peculiar properties. However, when we consider Kuno's (1973) actual analysis of "gapless relatives," we can raise the question whether Japanese has relative clauses at all.

It should be noted first that Kuno himself does not consider (4) a gapless relative. His main observation is that there is a striking parallelism between the relative construction and the topic construction. Thus, corresponding to (4) is example (5), with a topic.

(5) [Buturigaku-wa] [syuusyoku-ga muzukasii].
physics-TOP getting job-NOM hard-is
'As for physics, it is hard to get a job (in it).'

On the basis of this observation, Kuno proposes that Japanese relatives are in general derived by deleting the topic under identity with the relative head. The derivation of (4) is illustrated in (6).

(6) [DP[[buturigaku-wa]$_i$ [syuusyoku-ga muzukasii]] [buturigaku]$_i$]
∅

This analysis implies that (4) is not gapless: the relative clause has a gap in the topic position.

Kuno in fact seems to distinguish relative clauses and pure complex NPs in terms of whether they can be derived by topic deletion. (1), repeated in (7), is a relative construction because the sentential modifier can be derived from the topic sentence in (8) through deletion of the topic.

(7) [DP[Taroo-ga *e*i yonda] honi]
Taroo-NOM read book
'the book that Taroo read'

(8) [(Sono) hon-wa]i [Taroo-ga *e*i yonda].
that book-TOP Taroo-NOM read
'As for that book, Taroo read it.'

On the other hand, the examples in (9) are pure complex NPs, since the "corresponding topic sentences" are judged ungrammatical, as shown in (10).

(9) a. [DP[sakana-ga yakeru] [nioi]]
fish-NOM burn smell
'the smell of a fish burning'
b. [DP[doa-ga simaru] [oto]]
door-NOM shut sound
'the sound of a door shutting'

(10) a. *[(Sono) nioi-wa] [sakana-ga yakeru]].
that smell-TOP fish-NOM burn
'As for that smell, a fish burns.'
b. *[(Sono) oto-wa] [doa-ga simaru]].
that sound-TOP door-NOM shut
'As for that sound, a door shuts.'

Thus, for Kuno, pure complex NPs are gapless, whereas relative clauses always contain a gap in the topic position. It appears that his analysis is based on the hypothesis or intuition that relative clauses, by definition, contain a gap.

Then how did it become widely assumed that Japanese has gapless relatives? Later authorsfor example, Hale (1980) and Saito (1985)

pointed out that Kuno's analysis exhibits a redundancy. It assumes that the Japanese topic construction can be truly gapless as in (5). In order to account for the gapless examples, Kuno proposes that the topic can be licensed by a sort of "aboutness" relation; that is, a topic sentence is allowed as long as the topic is "about" the rest of the sentence. But once it is assumed that the "aboutness" relation licenses the topic construction, there is no reason not to suppose that the same relation also licenses the relative construction directly. The most straightforward analysis for (4) would be to say that the relative clause relative head relation is legitimate because the "aboutness" relation holds between the two. And this analysis, as opposed to the topic deletion analysis, implies that (4) is gapless exactly like (5).

This analysis seems to be an improvement on Kuno's. Even if lexical anaphors and NP-traces have similar distributions, this by itself does not mean that NP-traces are created by the deletion of lexical anaphors. At the same time, however, this analysis raises a fundamental question. As noted above, Kuno had a clear criterion for distinguishing relative clauses and pure complex NPs: the former contain a gap (in the topic position), and the latter do not. But if examples like (4) are gapless, it is at least less clear whether they should be considered relative clauses or pure complex NPs. Hale and Saito followed Kuno in assuming that (4) and (7), as opposed to (9a,b), are relative clauses. From this, they concluded that the same "aboutness" relation licenses the topic construction and the relative construction. But there seems to be no clear reason at this point to suppose that (4) is a relative clause whereas (9a,b) are pure complex NPs. And if (4) is not a relative clause but a pure complex NP, it may even be possible to say the same for (7).

The situation is further complicated by the fact that Japanese has other kinds of complex NPs that are not found in English. Consider again the examples in (10). The literal word-by-word translation of (10a), for example, is '*(the) smell (that) a fish burns' with the verb *yakeru* in present tense. If this is a pure complex NP, then the licensing condition on pure complex NPs is more generous in Japanese than in English. We even find examples like those in (11)

(11) a. [DP[tuti-o koneta] [katamari]]
soil-ACC softened and mixed clod
'the clod that was produced by softening and mixing soil'
(*Lit*. (the) clod (that) one softened and mixed soil)

b. [DP[zyagaimo-o yudeta] [mono]]
potato-ACC boiled thing
'the thing that was produced by boiling potatoes'
(*Lit*. (the) thing (that) one boiled potatoes)

According to Kuno's criterion, these are pure complex NPs. The topic sentences corresponding to them are unacceptable, as shown in (12).

(12) a. *[(Sono) katamari-wa] [tuti-o koneta].
that clod-TOP soil-ACC softened and mixed
'As for that clod, one softened and mixed soil.'
b. *[(Sono) mono-wa] [zyagaimo-o yudeta].
that thing-TOP potato-ACC boiled
'As for that thing, one boiled potatoes.'

And the examples in (11), like those in (10), do not have counterparts in English. 2

If the licensing condition for pure complex NPs is quite loose in Japanese, as (9) and (11) indicate, then the possibility arises that even (4) and (7) are pure complex NPs. Since constructions are artifacts and do not have a place in syntactic theory, there is no theory-independent way to decide which of the above complex NPs contain relative clauses and which are pure complex NPs. It is ultimately the theory that settles the matter. In the remainder of this chapter I will investigate this issue within Kayne's (1994) antisymmetry theory. I will first show that his analysis of N-final relatives leads to the conclusion that Japanese does not have relative clauses but only pure complex NPs. I will then suggest an explanation within his theory for the absence of relative clauses in Japanese.

6.2
The Nonmovement Property of Japanese Relative Clauses

Kayne (1994) proposes that within his antisymmetry theory N-final relative clauses have the structure in (13).3

(13) [DP[IP . . . *t*i . . .]j [D' D [CP NPi [C' C *t*j]]]]

This structure is derived as in (14).

(14) a. [DP[D' D [CP[C' C [IP . . . NP . . .]]]]
b. [DP[D' D [CP NPi [C' C [IP . . . *t*i . . .]]]]

From (14a), first the relative head moves to [Spec, CP] as in (14b). Then

the IP moves to [Spec, DP] to yield (13). On the other hand, within this theory N-final pure complex NPs can be analyzed as in (15a) or (15b).

(15) a. [DP[CP . . .]i [D' D [NP . . . [N' N *t*i]]]]
b. [DP[CP . . .] [D' D [NP . . . [N' N . . .]]]]

(15a) is the structure that results if the pure sentential modifier originates as the complement of N and moves to [Spec, DP]. In (15b), on the other hand, the pure sentential modifier is base-generated directly in [Spec, DP]. In this section I will discuss one of the notable properties of Japanese "relatives," the nonmovement property, and argue that given this property, (15) is at least as plausible as (13) for the structure of Japanese "relatives." 4

The evidence that Japanese "relatives" need not involve movement is found in Kuno 1973. The relevant observation is that Japanese "relatives" do not exhibit island effects. The example in (16) clearly demonstrates this generalization.

(16) [DP[[DP[*e*i *e*j kiteiru] [yoohukuj]]-ga yogoreteiru] [sinsii]]
wearing-is suit-NOM dirty-is gentleman
'the gentleman who [the suit that he is wearing] is dirty'

In this example *sinsi* 'gentleman' is "relativized" out of a complex NP. If Japanese "relatives" necessarily involve movement, we expect this example to be ruled out by Subjacency. Examples of this kind, then, indicate that Japanese "relatives" need not be derived by movement.

Perlmutter (1972) provides an explicit account for Kuno's observation. Japanese allows empty pronouns in any argument position, as illustrated in (17b) and (18b).

(17) a. Dare-ga sono hon-o mottekita no?
who-NOM that book-ACC brought
'Who brought that book?'
b. Taroo-ga *pro* mottekita.
Taroo-NOM brought
'Taroo brought it.'

(18) a. Taroo-wa nani-o mottekita no?
Taroo-TOP what-ACC brought
'What did Taroo bring?'
b. *pro* hon-o mottekita.
book-ACC brought
'He brought a book.'

From this it can be concluded that the gap *e*i in (16) need not be a trace of movement, but can be base-generated as *pro*. Thus, the absence of Subjacency effects in Japanese "relatives," Perlmutter argues, follows from the fact that Japanese is a pro-drop language.

Perlmutter claims that Japanese "relatives" can be base-generated with *pro* and hence need not be derived by movement. Hoji (1985) makes an even stronger claim, namely, that they can never involve movement. His argument is based on the absence of the connectivity or reconstruction effect in Japanese "relatives." Consider first the English example in (19).

(19) [DP[the picture of himselfi]j [that Johni likes *t*j best]]

The reflexive *himself* is contained within the relative head and is not bound by its antecedent *John*. Yet the example is fine. This situation obtains when the required binding is satisfied with the relative head construed in the position of the trace. This kind of effect, called the connectivity or reconstruction effect, is observed with movement, but not with a base-generated DP-pronoun structure, as shown in (20).

(20) a. [That picture of himselfi]j, Johni liked *t*j.
b. *[That picture of himselfi]j, Johni liked itj.

Hoji (1985) points out that the Japanese counterpart of (19) is impossible. A relevant example is shown in (21).

(21) *[DP[Johni-ga *e*j taipusita] [zibuni-no ronbun]j]
John-NOM typed self-GEN paper
(*Lit*. selfi's paper (that) Johni typed)

As Hoji notes, this absence of the connectivity effect indicates that Japanese relative clauses *cannot* involve movement, but must be base-generated with *pro* (if there is a gap).

Further evidence for Hoji's generalization comes from the "relativization" of adjuncts. First, (22) apparently shows that the relativization of reason/manner adjuncts, unlike that of arguments, is constrained by Subjacency.

(22) a. *[DP[IP[DP[IP *e*i *e*j kubi-ni natta] hitoj]-ga minna okotteiru]
was fired person-NOM all angry-is
riyuui]
reason
'the reason that [[all the people who were fired for it] are angry]'

b. *[DP[IP[DP[IP *e*i *e*j mondai-o toita] hitoj]-ga minna
problem-ACC solved person-NOM all
siken ni otiru] hoohooi]
exam in fail method
'the method that [[all the people who solve problems by it] fail
the exam]'

The grammatical status of these examples parallels that of the English (23a,b).

(23) a. *[the reasoni [that [all of the students who were fired *t*i] are
angry]]
b. *[the manneri [that [all of the students who solved the problem
*t*i] failed the exam]]

This parallelism can be accounted for straightforwardly if *pro* can occur only in argument positions in Japanese. Then (22a,b), as opposed to (16), must be derived by movement.

But the restriction on the "relativization" of reason/manner phrases in Japanese is much tighter. As noted in Saito 1985, they are clause-bound. Some relevant examples are provided in (24) (25).

(24) a. [[Mary-ga *e*i kaetta] riyuui]
Mary-NOM left reason
'the reasoni [that Mary left *t*i]'
b. *[[Mary-ga [John-ga *e*i kaetta to] omotteiru] riyuui]
Mary-NOM John-NOM left C think reason
'the reasoni [that Mary thinks [that John left *t*i]]'

(25) a. [[Mary-ga *e*i mondai-o toita] hoohooi]
Mary-NOM problem-ACC solved method
'the methodi [that Mary solved the problem *t*i]'
b. *[[Mary-ga [John-ga *e*i mondai-o toita to] omotteiru]
Mary-NOM John-NOM problem-ACC solved C think
hoohooi]
method
'the methodi [that Mary thinks [that John solved the problem
*t*i]]'

If (24b) and (25b) can be derived by movement, we expect them to be grammatical exactly like their English counterparts in (26a,b).

(26) a. [the reasoni [(for which) John thinks [that Mary was fired *t*i]]]
b. [the methodi [(by which) John thinks [that Mary solved the problem *t*i]]]

On the basis of the examples in (24) (25) and others, I argued in Murasugi 1991 that the relativization of manner/reason adjuncts is simply impossible in Japanese. According to this analysis, (24a) and (25a) do not contain any gap, and they are pure complex NPs exactly like (9a,b), repeated here.

(27) a. [DP[sakana-ga yakeru] [nioi]]
fish-NOM burn smell
'the smell of a fish burning'
b. [DP[doa-ga simaru] [oto]]
door-NOM shut sound
'the sound of a door shutting'

Then, the interpretation of (24a), for example, parallels that of the English (28).

(28) the reason for John's leaving

This analysis, if correct, provides support for Hoji's generalization. The examples in (24) and (25) cannot be base-generated with *pro*, since *pro* can appear only in argument positions. They cannot be derived by movement, either, if Japanese relative clauses can never involve movement as Hoji proposes. Hence, it follows that they cannot contain a gap at all.

The nonmovement property of Japanese "relatives" implies that the gap in the "relative," if it exists at all, is always *pro*. Thus, if the Japanese "relatives" are indeed relative clauses, they should have the structure in (29b), as opposed to the one in (29a) proposed by Kayne.

(29) a. [DP[IP . . . *t*i . . .]j [D' D [CP NPi [C' C *t*j]]]]
b. [DP[IP . . . (*pro*i) . . .]j [D' D [CP NPi [C' C *t*j]]]]

The relative head in (29b) is base-generated in [Spec, CP] and can be coindexed with a *pro* within the relative clause IP preposed to [Spec, DP].

But (29b) is "strange" in several respects. The structure in (29a) is based on the assumption that relative clauses are derived by movement of the relative head. The relative head is in [Spec, CP] because this position is the target for this type of movement. Further, the relative clause IP originates as the complement of C so that the relative head can move to [Spec, CP] from within the IP. But in (29b) the relative head NP is base-

generated in [Spec, CP]. One may wonder if the base generation of an NP in [Spec, CP] is legitimate. Whatever the answer to this question may be, it is clear that if a relative clause is not derived by the movement of the relative head, there are no reasons to suppose that the relative head is in [Spec, CP]. Also, given that the relative head is base-generated, there is no specific reason that the relative clause IP must originate in the complement position of C.

Overall, then, there is no reason to suppose that Japanese "relative clauses" have the D-CP structure in (29). Nothing seems to prevent them from having the structure of pure complex NPs, for example, the structure in (15b), repeated here.

(30) [DP[CP . . .] [D' D [NP . . . [N' N . . .]]]]

In the following section I will present evidence from acquisition that favors (30) over (29b) for Japanese "relatives." The evidence also suggests that what is in [Spec, DP] of (30) is not a CP but an IP.

6.3
The Overgeneration of *No* in Child Japanese

As discussed in detail in Murasugi 1991, around the age of 2 to 4 some Japanese-speaking children produce ungrammatical "relative clauses" like those in (31).

(31) a. *buta san-ga tataiteru no taiko (M: 2;11)
piggy-NOM hitting-is NO drum
'the drum that the piggy is playing'
b. *ohana motteru no wanwa (T: 2;6)
flower holding-is NO doggie
'the doggie that is holding a flower'

Here, the problem is the particle *no* following the "relative," which is not allowed in adult grammar. Without *no* the examples are grammatical. As shown in (32), the same phenomenon is observed with clear cases of pure complex NPs.

(32) *syuukuriimu tukutteru no nioi (E: 2;11)
cream puffs making-is NO smell
'the smell of someone making cream puffs'

Going over the analysis of these data in Murasugi 1991, I will argue that they constitute evidence that Japanese "relative clauses" are not relative

clauses (with the D-CP structure), but pure sentential modifiers with the structure in (3), repeated here.

(3) [DP[IP . . .] [D' D [NP . . . [N' N . . .]]]]

6.3.1
The Analysis Proposed in Murasugi 1991

It had been assumed in the literature that the *no* observed in (31) is the genitive Case marker (see Harada 1980, Clancy 1985). However, I argued in Murasugi 1991 that it is instead a complementizer. First, in many dialects of Japanese, including the Tokyo dialect, *no* is three-ways ambiguous. It can be the genitive Case marker, a pronominal, or a complementizer, as shown in (33a c), respectively. 5

(33) a. Taroo-no hon
Taroo-GEN book
'Taroo's book'
b. akai no
red-is one
'the one that is red'
c. [CP[IP Doroboo-ga kane-o nusunda] no]-wa koko-kara da.
thief-NOM money-ACC stole C-TOP here-from is
'It is from here that the thief stole money.'

One clear piece of evidence that the *no* in (31) is not the genitive Case marker comes from dialects where different morphemes are used for genitive Case, on one hand, and the pronominal and the complementizer, on the other. In the Toyama dialect, for example, the genitive Case marker is *no*, but *ga* is used for the pronominal and the complementizer, as illustrated in (34).

(34) a. Taroo-no hon
Taroo-GEN book
'Taroo's book'
b. akai ga
red-is one
'the one that is red'
c. [CP[IP Doroboo-ga kane-o nusunda] ga]-wa koko-kara da.
thief-NOM money-ACC stole C-TOP here-from is
'It is from here that the thief stole money.'

And interestingly, children acquiring the Toyama dialect produce *ga*, and not *no*, in "relative clauses." Some examples are shown in (35).

(35) a. *akai ga boosi (K: 2;11)
red-is GA cap
'the cap that is red'
b. *anpanman tuitoru ga koppu (K: 2;11)
(a character) attaching-is GA cup
'the cup that is pictured with "anpanman"'
(*Lit*. the cup that "anpanman" is attaching)

If the overgenerations observed in (31) and (35) are of the same kind, as seems plausible, then the *no* in (31) cannot be the genitive Case marker.

There is also evidence that the *no* in (31) is not a pronominal. Suppose it is. Then, in (31b), for example, *ohana motteru no* would be a nominal phrase that is headed by the pronominal *no* and modifies the head of the larger nominal phrase, *wanwa* 'dog'. But in Japanese, whenever a nominal phrase modifies a nominal head, the genitive Case marker is obligatory, as shown in (36).

(36) a. Taroo-no midori-no hon
Taroo-GEN green-GEN book
'Taroo's green book'
b. akai no-no motinusi
red-is one-GEN owner
'the owner of the one that is red'
(the owner of the red one)

And in my experimental study, the children who overgenerated *no* consistently used the genitive Case marker appropriately. Hence, if overgenerated *no* in (31) were a pronominal, we would expect the forms in (37) instead of the actual forms in (31).

(37) a. *buta san-ga tataiteru no-no taiko
piggy-NOM hitting-is NO-GEN drum
'the drum that the piggy is playing'
b. *ohana motteru no-no wanwa
flower holding-is NO-GEN doggie
'the doggie that is holding a flower'

Since the children never produced these forms, it is safe to conclude that the *no* in question is not a pronominal.

These considerations lead to the conclusion that the overgenerated *no* is of the category C. What the children do, then, is produce an overt com-

plementizer in the prenominal CP. 6 The remaining question is how the children retreat from the overgeneration of *no*.

In Murasugi 1991 I argued that prenominal sentential modifiers in adult Japanese are of the category IP (instead of CP) and that children can attain this adult grammar on the basis of positive evidence. Note first that an overt complementizer is obligatory in English pure complex NPs.

(38) the fact [CP *(that) John fell asleep]

Kayne (1981) and Stowell (1981) propose an Empty Category Principle (ECP) account for this fact. Suppose that the complementizer *that* is missing in (38). Then, there must be an empty C heading the CP. But this empty category is not properly licensed by a "governing" head and hence is excluded by the ECP. It follows that the complementizer *that* must be present in this example.

If this analysis for (38) is on the right track, then the sentential modifiers in Japanese pure complex NPs must be of the category IP. In Japanese complex NPs a complementizer is never allowed. (39a) is an example of a pure complex NP, and (39b) contains a "relative clause."7

(39) a. [[Taroo-ga inemuri-o sita (*no)] zizitu]
Taroo-NOM falling asleep-ACC did C fact
'the fact that Taroo fell asleep'
b. [[Taroo-ga *e*i yonda (*no)] honi]
Taroo-NOM read C book
'the book that Taroo read'

The fact that (39a) is grammatical without *no* indicates that the modifying clause cannot be a CP. If it were a CP, then it would have to be headed by an empty C, which should be excluded by the ECP exactly as in (38). Hence, (39a), without *no*, should lack the C projection, and the modifying clause should be of the category IP. Further, if the clause is a bare IP, it is correctly predicted that a complementizer is never allowed.

Here, the children acquiring adult Japanese can go through the same reasoning on the basis of examples like (39a). If they are equipped with the ECP, or whatever universal principle explains (38), then they can infer on the basis of positive evidence (i.e., grammatical pure complex NPs like (39a) without *no*) that pure sentential modifiers cannot be of the category CP but must be of the category IP. Further, if they generalize this to all prenominal sentential modifiers, they will conclude that "relative clauses"

are IPs as well. At this point they should cease to produce *no*, the complementizer, in "relative clauses."

This analysis implies that CP is the unmarked category for the sentential modifiers in complex NPs. Children learning Japanese hypothesize initially that those sentential modifiers are CPs and hence produce *no* in the head C position. Then, on the basis of evidence from pure complex NPs, they eventually discover that those sentential modifiers are IPs in Japanese.

6.3.2
A Revised Analysis within the Antisymmetry Theory

Let us now examine how the analysis I have just summarized fares with Kayne's antisymmetry analysis of N-final relatives. Suppose Japanese "relatives" have the D-CP structure as in (29b), repeated here.

(40) [DP[IP . . . (*pro*i) . . .]j [D' D [CP NPi [C' C *t*j]]]]

Given this structure, the overgenerated *no* in (31) cannot be a complementizer. If it were, it should follow the relative head, contrary to fact. Thus, this structure is clearly incompatible with the analysis of the acquisition data presented above. Given this structure, the overgenerated *no* apparently must be a D. We have already seen clear evidence from the Toyama dialect that the *no* in question is not the genitive Case marker. So, the only possibility left seems to be that *no* is a pronominal with the category D.

Initially this hypothesis may seem plausible. First, the pronominal *no* is a pronoun; hence, it would not be surprising if it were a D. Second, the argument presented above against the hypothesis that the overgenerated *no* is a pronominal does not hold given the structure in (40). Let us again consider (31b), repeated here.

(41) *ohana motteru no wanwa (T: 2;6)
flower holding-is NO doggie
'the doggie that is holding a flower'

I stated above that if the *no* is a pronominal, then it should be the head of the nominal phrase *ohana motteru no*, which modifies *wanwa*. Hence, the children should produce the genitive *no* between the modifying nominal phrase and the modified nominal head. Since this never happens, I concluded that the *no* is not a pronominal. However, if the *no* is the D in the structure (40), there is no modification relation of the kind stated above. That is, if (41) has the structure in (40) with *no* in the D position, *ohana*

motteru no does not modify the NP *wanwa*. The former is not even a constituent. Thus, it is not a problem that the genitive *no* does not appear in (41).

But other problems arise with the hypothesis that the *no* in question is a pronominal D. As shown in (33b), the pronominal *no* can be modified by a "relative clause." More examples of modified *no* are shown in (42). 8

(42) a. sono karai no
that spicy-is one
'that spicy one'
(the one that is spicy)
b. Taroo-ga katta no
Taroo-NOM bought one
'the one that Taroo bought'
c. Taroo no
Taroo one
'Taroo's one'

In fact, the pronominal *no* has a peculiar property: it cannot appear in isolation, but must be accompanied by a modifier. Thus, (43) is ungrammatical.

(43) *No-o mottekite kudasai.
one-ACC bring please
'Please bring one.'

These facts show that the pronominal *no* has properties quite different from those of English pronouns such as *she, he, it, they*. They suggest in fact that the pronominal *no* is an N rather than a D.

Examples such as (42a) and (42b) are particularly important in the present context. These examples are "relative clauses" with the pronominal *no* in the "relative" head position. If *no* is of the category D, then these "relative clauses" cannot have the structure in (40). And if they do have the structure in (40), then *no* must be an N heading the NP in [Spec, CP] rather than a D. Thus, the antisymmetry analysis of N-final relatives itself implies that the pronominal *no* is an N.

Further, the children who overgenerate *no* do produce forms such as (42) quite productively. Actual examples are shown in (44).9

(44) a. Kore, osaru-san no. (E: 2;11)
this monkey-Mr. one
'This is Mr. Monkey's one.'
(This is Mr. Monkey's.)

b. akai no (E: 2;11)
red-is one
'the red one'
(the one that is red)

This indicates that these children consider the pronominal *no* an N rather than a D, exactly as in adult Japanese. It is thus difficult to maintain that the overgenerated *no* is a pronominal D.

As noted earlier, the acquisition facts in (31) pose a problem for the D-CP analysis of Japanese "relatives." If (40) is the structure of Japanese "relatives," there is no plausible way to account for them. On the other hand, if Japanese "relatives" are not relative clauses but pure sentential modifiers, the analysis proposed in Murasugi 1991 can be maintained almost intact. According to this hypothesis, the examples in (31) have the structure in (30), repeated in (45).

(45) [DP[CP . . .] [D' D [NP . . . [N' N . . .]]]]

Here, the overgenerated *no* can be of the category C heading the prenominal CP, as proposed above. The children eventually discover that the sentential modifier in (45) is an IP in adult Japanese and consequently cease to overgenerate *no*.

The hypothesis that Japanese "relatives" are pure sentential modifiers in fact leads to an improvement on the analysis in Murasugi 1991. Recall that Japanese pure complex NPs such as (39a), repeated in (46), provide the children with positive evidence that pure sentential modifiers are IPs in Japanese.

(46) [[Taroo-ga inemuri-o sita (*no)] zizitu]
Taroo-NOM falling asleep-ACC did C fact
'the fact that Taroo fell asleep'

From such cases, I proposed, children make a generalization that all prenominal sentential modifiers, including "relative clauses," are of the category IP. This was a crucial step to account for children's recovery from overgenerating *no* in "relative clauses." But this step is no longer necessary if Japanese "relatives" are pure sentential modifiers. If (46) and a simple "relative" like the one in (47) have the same structure, either can serve as direct positive evidence that Japanese "relative clauses" are IPs.

(47) [[Taroo-ga *e*i yonda] honi]
Taroo-NOM read book
'the book that Taroo read'

The acquisition data in (31), then, constitute evidence for the analysis of Japanese "relatives" as pure sentential modifiers. (45) is the structure of "relatives" in child Japanese. In adult Japanese, on the other hand, pure sentential modifiers are IPs. Thus, Japanese "relatives" have the structure in (3), repeated in (48).

(48) [DP[IP . . .] [D' D [NP . . . [N' N . . .]]]]

In the following section I will speculate on why Japanese lacks true relatives with the D-CP structure altogether. I will show that a plausible explanation is possible within the antisymmetry theory.

6.4
The Absence of Relative Clauses in Japanese: A Proper Binding Analysis

As noted above, Kayne (1994) proposes the following structure (repeated from (29a)) for N-final relatives:

(49) [DP[IP . . . *t*i . . .]j [D' D [CP NPi [C' C *t*j]]]]

I entertained (and eventually rejected) the possibility that Japanese "relatives," which never involve movement, instead have the structure in (50) (repeated from (29b)).

(50) [DP[IP . . . (proi) . . .]j [D' D [CP NPi [C' C *t*j]]]]

But as also discussed above, the D-CP structure is motivated to the extent that the relative head moves to [Spec, CP]. Hence, (50) lacks independent motivation. The question of why Japanese does not have true relatives with the D-CP structure, then, reduces to the question of why (49) is not possible in this language. In other words, the explanation for the nonmovement property also explains why Japanese does not have true relatives with the D-CP structure.

As Kayne notes himself, the structure in (49) contains an unbound trace, *t*i. It thus apparently violates the Proper Binding Condition, shown in (51).

(51) Traces must be bound. (Fiengo 1977)

However, Kayne also notes that this is not necessarily a problem, since there are cases where unbound traces are allowed. The case he cites is remnant topicalization in German.

At this point let us examine more closely the contexts where unbound traces are allowed. Saito (1986) argues that there is a clear asymmetry

between A- and Ā-traces with respect to the application of the Proper Binding Condition. 10 For example, (52a) is perfectly acceptable, but (52b) is not.

(52) a. [How likely [*t*i to win]]j is Johni *t*j?
b. *[Which picture of *t*i]j does John wonder whoi Mary likes *t*j?

In (52a) *John* raises to the matrix subject position and then the *wh*-phrase *how likely ti to win*, which contains the trace of *John*, moves to [Spec, CP]. The trace of raising *t*i is not bound, and yet the example is grammatical. In (52b), on the other hand, the *wh*-phrase *who* first moves to the embedded [Spec, CP], and then the larger *wh*-phrase *which picture of t*i, which contains the trace of *who*, moves to the matrix [Spec, CP]. The example is simply uninterpretable. It seems then that traces of A-movement such as *t*i in (52a) can be licensed through reconstruction or connectivity, but traces of Ā-movement such as *t*i in (52b) have to be bound in the strict sense.

Let us apply this generalization to the structure in (49). If movement to [Spec, CP] is in general Ā-movement, the trace *t*i should be an Ā-trace. Since Ā-traces must be bound in the strict sense, the structure should be excluded. As long as *t*i is a trace of Ā-movement, there does not seem to be any way to save the structure. Thus, the nonmovement property of Japanese "relative clauses" is derived. And if the D-CP structure in (49) (50) is motivated only to the extent that the relative head moves to [Spec, CP], as assumed here, then it also follows that Japanese cannot have true relatives with this structure. Thus, the absence of relative clauses in Japanese is explained by Kayne's analysis of N-final relatives and the Proper Binding Condition.

Since the structure in (49) always seems to violate the Proper Binding Condition, the explanation offered above appears to imply that no language can have N-final relatives. And if this is the case, one can question Kayne's antisymmetry analysis of N-final relatives itself. If the structure in (49) is not realized in any language, it would be difficult to argue for it. I speculate at this point that there are contexts in which the structure in (49) is allowed, and hence there are languages (other than Japanese) that have N-final relatives with this structure. In the remainder of this section I will present one possible way to substantiate this speculation.

Discussing the typology of relatives, Keenan (1985) singles out Japanese "relatives" as being unique among the N-final relatives.

> A more regular difference between prenominal and postnominal RCs [relative clauses] concerns the form of the main verb of Srel, which we shall denote by Vrel. In prenominal RCs, Vrel is almost always in some sort of non-finite form, that is a form different from the one it would have as the main verb of a simple declarative sentence. Typically Vrel exhibits a reduction in tense-aspect marking and in verb agreement morphology compared with main clause declarative verbs . . . . We may note that the prenominal RCs in Japanese do not put Vrel in a non-finite or specifically relative form, but the Japanese case appears to be the exception among prenominal RCs here. (Keenan 1985, 160 161)

Simply put, Japanese "relatives" are unique among N-final relatives in that their main verbs are in the regular finite form.

Mahajan (1990) observes that the nonfinite/finite distinction relates to the A/$\bar{A}$ distinction in Hindi scrambling in an interesting way. As it is shown in Nemoto 1993 that Mahajan's generalization holds in Japanese as well, I will use Japanese examples to illustrate his observation here. Let us first consider the examples in (53) and (54). 11

(53) ?Karera-oi [[otagaii-no sensei]-ga *t*i hihansita].
they-ACC each other-GEN teacher-NOM criticized
'*Each other's teachers criticized them.'
(*Lit*. Themi, each other'si teachers criticized *t*i.)

(54) a. Karera-oi [John-ga [Mary-ga *t*i hihansita to] itta].
they-ACC John-NOM Mary-NOM criticized C said
'John said that Mary criticized them.'
(*Lit*. Themi, John said that Mary criticized *t*i.)

b. *Karera-oi [[otagaii-no sensei]-ga [Mary-ga *t*i hihansita
they-ACC each other-GEN teacher-NOM Mary-NOM criticized
to] itta].
C said
'*Each other's teachers said that Mary criticized them.'
(*Lit*. Themi, each other'si teachers said that Mary criticized *t*i.)

(53) shows that a phrase preposed by clause-internal scrambling can serve as the antecedent for a lexical anaphor. This implies that clause-internal scrambling can be A-movement. (54a) shows that long scrambling out of a finite clause is possible in Japanese. On the other hand, (54b) shows that a phrase preposed by this kind of scrambling cannot be the antecedent of a lexical anaphor. This means that long scrambling out of a finite clause is necessarily $\bar{A}$-movement.

Here, interestingly, long scrambling out of a nonfinite clause patterns with clause-internal scrambling, and not with long scrambling out of

a finite clause. The Japanese example in (55) is adapted from Nemoto 1993.

(55) ?Karera-oi [Mary-ga [otagaii-no sensei]j-ni [PROj *t*i homeru
they-ACC Mary-NOM each other-GEN teacher-to praise
yooni] tanonda].
to asked
'*Mary asked each other's teachers to praise them.'
(*Lit*. Themi, Mary asked each other'si teachers to praise *t*i.)

This shows that long scrambling out of a nonfinite clause can be A-movement. Mahajan (1990) thus arrives at the generalization in (56).

(56) Long scrambling
a. out of a finite clause . . . must be $\bar{A}$-movement.
b. out of a nonfinite clause . . . can be A-movement.

Since it is shown convincingly in Webelhuth 1989 and Saito 1989 that the landing site of scrambling is not [Spec. IP] and yet it is nonoperator movement, it is reasonable to suppose that the generalization in (56) holds for this kind of movement in general. 12

Let us now return to the discussion of relative clauses and consider again the structure in (49), repeated in (57).

(57) [DP[IP . . . *t*i . . .]j [D' D [CP NPi [C' C *t*j]]]]

Suppose now that [Spec, CP] in a relative clause, at least in some languages, can be a nonoperator position. This is not implausible, since relativization does not establish an operator-variable relation in the way that *wh*-question movement does. And suppose also that *clause* in the generalization in (56) means IP. Then, the generalization implies that relativization, that is, the movement of NPi to [Spec, CP] in (57), can be an A-movement as long as the relative clause is nonfinite. In this case *t*i is an A-trace, and there is nothing wrong with the structure in (57). As shown in (52), an A-trace need not be bound in the strict sense and can be licensed through reconstruction or connectivity.

The discussion here is quite speculative. However, I believe it shows that Kayne's analysis need not imply that N-final relatives are never possible. There seems to be a way to allow the structure in (57) in those languages where relative clauses have nonfinite main verbs. Note that even if this speculation is correct, Japanese still cannot have relative clauses with the structure in (57). Since Japanese "relatives" have finite main verbs, the

movement to [Spec, CP] will involve extraction out of a finite IP. Hence, the movement is necessarily Ā-movement, and *t*i in (57) is an Ā-trace. The structure is ruled out in this language exactly as (52b) is. Therefore, the absence of relative clauses in Japanese still follows.

6.5 Conclusion

As noted at the outset, whether Japanese has relative clauses or not is a theoretical question. It is the theory that ultimately provides an answer. In this chapter I argued that given Kayne's (1994) proposal that relative clauses have the D-CP structure, we are led to conclude that Japanese lacks relative clauses altogether. Further, I argued that this absence of relative clauses in Japanese follows from the Proper Binding Condition and Kayne's analysis itself.

It has been widely assumed for some time that one of the remarkable properties of Japanese is that it has relative clauses with peculiar properties. The "relatives" in this language need not contain a gap and can never be derived by movement. If the conclusions in this chapter are correct, then a very different picture emerges. Japanese does not have relative clauses, but uses pure complex NPs quite extensively. I noted above that there are independent reasons to believe that the licensing condition on pure complex NPs is quite generous in Japanese. Thus, we find examples such as (9a,b), repeated in (58a,b).

(58) a. [DP[sakana-ga yakeru] [nioi]]
fish-NOM burn smell
'the smell of a fish burning'
(*Lit*. the smell that a fish burns)
b. [DP[doa-ga simaru] [oto]]
door-NOM shut sound
'the sound of a door shutting'
(*Lit*. the sound that a door shuts)

As far as complex NPs are concerned, the most distinctive property of Japanese then is that pure sentential modifiers can modify the nominal heads in ways that are not allowed in other languages. Because of this, the language can employ pure sentential modifiers to express the "relative clause meaning." It is also because of this that Japanese appears to have a relative construction with very peculiar properties.

Notes

Part of the material in this chapter was presented at the Berlin Workshop on Relative Clauses held in November 1996. I would like to thank the audience there for helpful comments and suggestions. Thanks are also due to Mamoru Saito and an anonymous reviewer for detailed comments on an earlier version of the chapter. Preparation of this revised version was made possible by the Nanzan University Pache Research Grant 1A.

I would like to take this opportunity to express my indebtedness to Howard Lasnik on more general grounds. I would not have been able to write this chapter if I had not learned syntax from him. I hope this chapter reflects his excellent teaching.

1. Kuno (1973) does not assume the DP hypothesis (nor do Hale (1980) and Saito (1985), cited below). But I will use the label DP for (4) and other examples they discuss to make the presentation of examples in this chapter consistent.

2. For more detailed discussion on examples like (11), see Murasugi 1997 and the references cited there.

3. According to Kayne, specifier is an adjoined position. But since this particular proposal is not relevant to the present discussion. I will ignore it here.

4. In section 6.3 I will present a supporting argument for (15), as opposed to (13), and at the same time propose a modification of (15), arguing for the structure in (3), where the sentential modifier is an IP.

5. See Murasugi 1991 for a detailed discussion on the three *nos*. See also Hoji 1990 for the analysis of Japanese cleft sentences such as (33c).

6. This implies that those children already know that *no* is a complementizer. In Murasugi 1991 I cited the fact that they frequently use *no* as the sentence-final particle in simple declarative sentences. If this *no* is a C, this fact shows that they have already acquired *no* of the category C.

   It should be mentioned in this context that there are complementizers other than *no* in Japanese. Thus, the conclusion in the text also implies that those children who overgenerate consider *no* the appropriate complementizer to be used in prenominal CPs. This does not seem surprising. The complementizer *ka* (Q-morpheme) is limited to question CPs and is clearly inappropriate. The other complementizer *to* appears typically when the CP subcategorizes the verb. This fact, I assume, makes the children overgenerate *no* and not *to* in the position in question.

7. Recall that children produce pure complex NPs and "relative clauses" with *no*, which is not allowed in adult Japanese. Thus, the children's utterances in (31) (32) correspond to (39) with *no*.

8. As the *no* in (42c) is a pronoun, the genitive *no* should appear between *Taroo* and this *no*, as in (i)

   (i) *Taroo-no no
   Taroo-GEN one
   `Taroo's one'

But when the pronoun *no* is preceded by the genitive *no*, the latter is deleted by a low-level process called *no*-reduction. See Okutsu 1974 for detailed discussion of this phenomenon.

9. The *no*-reduction mentioned in note 8 applies in (44a). Interestingly, the children never overgenerate *no* when the "relative head" is the pronominal *no*, as indicated in (44b). If the overgenerated *no* is a C, as argued in Murasugi 1991, then this means that *no*-reduction applies not only to the genitive *no* but also to the complementizer *no*.

10. See also Saito 1989, fn. 14, for relevant discussion. Saito (1989) and Lasnik and Saito (1992) adopt a different generalization.

11. These examples are discussed in detail in Tada 1990 and Saito 1992.

12. (56) clearly does not hold for movement to [Spec, IP] or operator movement. Movement to [Spec, IP] out of a finite clause is simply impossible; and operator movement is necessarily Ā-movement.

References

Clancy, Patricia. 1985. The acquisition of Japanese. In Daniel I. Slobin, ed., *The crosslinguistic study of language acquisition: Volume 1, 373 524*. Hillsdale, N.J.: Lawrence Erlbaum.

Fiengo, Robert. 1977. On trace theory. *Linguistic Inquiry* 8, 35 61.

Hale, Ken. 1980. Remarks on Japanese phrase structure: Comments on the papers on Japanese syntax. In Yukio Otsu and Ann Farmer, eds., *Theoretical issues in Japanese linguistics*, 185 203. (MIT Working Papers in Linguistics 2.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Harada, Kazuko I. 1980. Notes on the acquisition of the genitive Case particle *no*. Ms., Kinjo Gakuin University, Nagoya.

Hoji, Hajime. 1985. Logical form constraints and configurational structures in Japanese. Doctoral dissertation, University of Washington, Seattle.

Hoji, Hajime. 1990. Theories of anaphora and aspects of Japanese syntax. Ms., University of Southern California, Los Angeles.

Kayne, Richard. 1981. ECP extensions. *Linguistic Inquiry* 12, 93 133.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Keenan, Edward. 1985. Relative clauses. In Timothy Shopen, ed., *Language typology and syntactic description II: Complex constructions*, 141 170. Cambridge: Cambridge University Press.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, Mass.: MIT Press.

Lasnik, Howard, and Mamoru Saito. 1992. *Move* a. Cambridge, Mass.: MIT Press.

Mahajan, Anoop. 1990. The A/A' distinction and movement theory. Doctoral dissertation, MIT, Cambridge, Mass.

Murasugi, Keiko. 1991. Noun phrases in Japanese and English: A study in syntax, learnability and acquisition. Doctoral dissertation, University of Connecticut, Storrs.

Murasugi, Keiko. 1997. The syntax-discourse interface. In Michiko Temma, ed., *Studies on inference in reading comprehension and its application to second language reading*, 57 84. Tsuda College, Tokyo.

Nemoto, Naoko. 1993. Chains and Case positions: A study from scrambling in Japanese. Doctoral dissertation, University of Connecticut, Storrs.

Okutsu, Keiichiro. 1974. *Seisei nihon bunpooron* (A generative analysis of Japanese). Tokyo: Taishukan.

Perlmutter, David. 1972. Evidence for shadow pronouns in French relativization. In Paul M. Peranteau, et al., eds., *The Chicago which hunt*, 73 105. Chicago Linguistic Society, University of Chicago, Chicago. Ill.

Saito, Mamoru. 1985. Some asymmetries in Japanease and their theoretical implications. Doctoral dissertation, MIT, Cambridge, Mass.

Saito, Mamoru. 1986. LF effects of scrambling. Ms., University of Southern California, Los Angeles.

Saito, Mamoru. 1989. Scrambling as semantically vacuous A'-movement. In Mark Baltin and Anthony Kroch, eds., *Alternative conceptions of phrase structure*, 182 200. Chicago: University of Chicago Press.

Saito, Mamoru. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1, 69 118.

Stowell, Tim. 1981. Origins of phrase structure. Doctoral dissertation, MIT, Cambridge, Mass.

Tada, Hiroaki. 1990. Scrambling(s). Ms., MIT, Cambridge, Mass.

Webelhuth, Gert. 1989. Syntactic saturation phenomena and the modern Germanic languages. Doctoral dissertation, University of Massachusetts, Amherst.

Chapter 7
A Conspiracy Theory of Case and Agreement

Javier Ormazabal

7.1 Case Checking and Movement

Case theory plays a central and ubiquitous role in the Minimalist Program. However, in a very fundamental sense it constitutes an unreduced oddity in a system that overall tries to motivate its primitives in terms of properties of the interface levels. Case features are uninterpretable in both the argument and the attracting head, and their only motivation seems to be that they force the general displacement of arguments to nonbase positions. On the other hand, Case theory appears to be empirically very well motivated; this fact makes it more difficult to derive its effects from other principles of syntactic movement.

This is by no means a problem confined to the Minimalist Program; it has been present since the introduction of Case theory in the late 1970s. There have been many attempts to give some independent conceptual basis to the theory of Case. From Rouveret and Vergnaud's (1980) Case Filter to Aoun's (1982) and Chomsky's (1986) Visibility Condition and its continuations in the Minimalist Program (see, e.g., Lasnik 1995, Boskovic * 1995), most theories have tried to find some interface motivation for Case, either at PF or at LF. More innovative approaches such as those of Uriagereka (1994) and Martin (1996) explore a more functional approach where Case, although itself uninterpretable, survives in the mapping to the semantic component as a means to distinguish sets of agreement features in the V-T (ense) complex.

My goal in this chapter is to explore an approach to A-movement that derives the displacement effects generally attributed to Case from the conspiracy of different factors. In particular, I will argue that syntactic accusative/absolutive Case does not exist and that whenever complements move from VP to some higher checking position, this movement is

motivated by some property other than Case. I show that this proposal is empirically supported by the behavior of rich clitic and agreement systems across languages; in particular, I argue that many restrictions on clitic and agreement clusters traditionally discussed in the literature (see Perlmutter 1971, Bonet 1991, 1994, Albizu 1997, and references therein) can be syntactically accounted for if we assume a restrictive version of the Minimal Link Condition (MLC) where at most two arguments can move from the VP shell to check the same type of feature. I then conclude that whenever ditransitive predicates realize their three arguments VP-internally, at least one of these arguments does not move out. This conclusion leads me to propose a major revision of standard assumptions about A-movement and to argue that Case theory cannot play the central role that the standard minimalist approach attributes to it. I then argue that what triggers object movement is an animacy feature: only arguments that are inherently [+animate] must move out of VP. Finally, I also show that other asymmetries sensitive to the type of the complementin particular, the behavior of strong personal and demonstrative pronounsalso follow from the inherent [± animate] properties of the arguments involved.

7.2 The Minimal Link Condition and Multiple Agreement

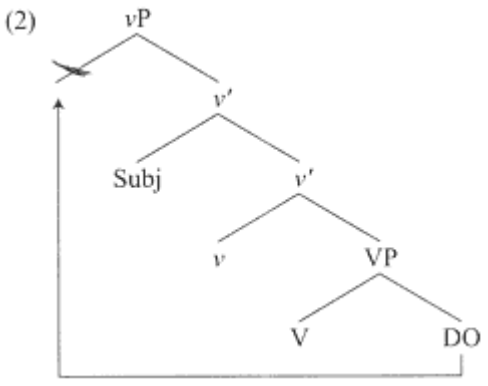
The Minimal Link Condition in (1) (MLC; see Chomsky 1995, especially the discussion in sections 4.5.5 and 4.10) has been designed to derive some of the Relativized Minimality effects that were subsumed under different versions of the Empty Category Principle (ECP) in the Government-Binding framework. The MLC, incorporated into the definition of Attract F in Chomsky (1995), determines that when two elements compete to be attracted, the closer one moves.

(1) b is closer to K than a unless b is in the same minimal domain as (a) K or (b) a.

This condition also ensures that subject and object end up in the right checking configuration; once we assume the MLC, there is no need to stipulate that in a typical transitive construction the complement of the verb checks its Case in a lower position than the subject. The MLC blocks direct movement of the object to check the Case features of T, because the c-commanding subject in [Spec, n] is neither in the minimal domain of the object nor in the minimal domain of the target. On the other hand, the subject does not block movement of the object to the outer specifier of

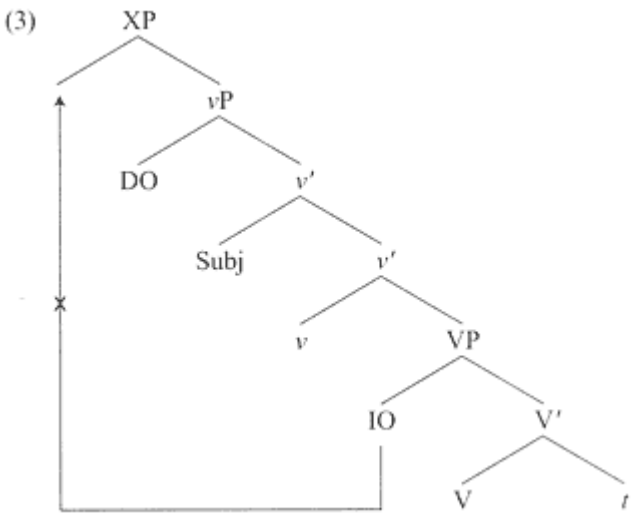
nP, because it is in the same minimal domain as the target, as shown in (2).

(2)



However, given the specific way in which it is formulated, the MLC, as a by-product, also has a very interesting consequence: it rules out any possibility of moving out of VP the three arguments (nominative/ergative, accusative/absolutive, and dative) of a ditransitive predicate for Case reasons. Little is known about the mechanisms of dative Case/agreement checking, but it is natural to assume that it takes place in some projection higher than nP. If that is correct, consider the point of the derivation where the object has moved to the outer (Spec, nP) and there is a feature that potentially attracts the Case feature of the indirect object, as in (3).

(3)



Given that the specifiers of nP are neither in the same minimal domain as the target nor in the same one as the indirect object, the subject is closer to XP and should be attracted. Consequently, the indirect object can never move out of the nP shell whenever three arguments are realized. 1 This effect has been mostly considered an inherent limitation of the Minimal Link Condition as standardly defined: it is an issue particularly relevant in languages with rich agreement where the auxiliary system overtly agrees with the three arguments of the verb, but it has empirical consequences elsewhere.

Different strategies have been pursued to adapt the system to a three-way agreement model, either by modifying the MLC (or Shortest Movement) or by making different assumptions about clausal structure that are compatible with the MLC.2 In this chapter I want to pursue a very different line of research: I will propose that there is no need to weaken the MLC in any of the directions proposed in the literature and that, as currently analyzed, it yields the right type of empirical results. In support of this proposal I will discuss some properties of rich auxiliary systems and multiple clitic clusters; I will show that the MLC straightforwardly accounts for well-known clitic/agreement cluster restrictions that show up pervasively across languages.

7.3
Restrictions on Clitic/Agreement Clusters

Perlmutter (1971) was the first to discuss the following contrast involving combinations of clitics in Romance:3

(4) a. Pedro me lo envía.
Pedro CL.1DAT CL.3ACC send.3NOM
'Pedro sends him to me.'
b. *Pedro le me envía.
Pedro CL.3DAT CL.1ACC send.3NOM
'Pedro sends me to him.'

Both (4a) and (4b) combine two preverbal clitics, corresponding to the accusative and dative arguments of the sentence. However, whereas the combination of dative first person and accusative third person clitics in (4a) is possible, the combination of dative third person and accusative first person clitics in (4b) is not. The paradigm in (4) illustrates a more general contrast between third person and first/second person accusative clitics in ditransitive structures, known in the literature as the *me-lui constraint*.

Similar restrictions are found in languages with multiple agreement; these languages systematically show certain gaps in the paradigm of ditransitive verbs where only third personbut not first or second personaccusative/absolutive markers cooccur with dative agreement. The contrast in (5) illustrates this point in Basque, but the same phenomenon can be observed in other rich agreement languages as well. 4

(5) a. Azpisapoek etsaiari misilak saldu d -f -izki -o
traitors-ERG enemy-DAT missiles-ABS sell -3ABS -3DAT
-te.
-3ERG
'The traitors sold missiles to the enemy.'
b. *Azpisapoek etsaiari ni saldu na -i -o -te.
traitors-ERG enemy-DAT me-ABS sell 1ABS -3DAT -3ERG
'The traitors sold me to the enemy.' (Laka 1993)

Bonet (1994) convincingly argues that the cooccurrence restrictions in (4) (5) form part of the very same pattern; as a first approximation, she proposes the descriptive generalization in (6), according to which the presence of a dative agreement or a dative clitic blocks accusative/absolutive agreement other than third person.

(6) If DAT, then ACC/ABS = 3rd person.

Although (6) only makes reference to the internal arguments of the verb, there is a hidden assumption that will be of great relevance later in the discussion: in addition to the dative and accusative/absolutive elements, the examples in (4) (5) contain a subject agreement in the morphological structure of the verb. In accusative languages the accusative Case is never active without the presence of nominative Casethat is precisely the main insight of Burzio's Generalizationand no issue arises. However, in ergative languages the ergative agreement is not required for absolutive and dative arguments to occur. This can be seen in (7), where absolutive first and second person agreement is allowed to appear with third person dative in the absence of an ergative agreement. Thus, as (7b) shows, combinations of first person absolutive and third person dative agreements are possible in the context of unaccusative verbs in Basque, contra Bonet's descriptive generalization.

(7) a. Jon etorri f- zai -t.
Jon-ABS come 3ABS -1DAT
'Jon came to me.'

b. Joni joan na -tzai -o.
Jon-DAT go 1ABS -3DAT
'I went to Jon.'

Therefore, all cases of person restrictions on the accusative/absolutive clitic and agreement systems involve the presence of a subject agreement, in addition to the dative one. I thus reformulate Bonet's descriptive generalization as in (8).

(8) If DAT *and NOM/ERG*, then ACC/ABS = 3rd person.

The syntactic account in terms of the Minimal Link Condition that I will pursue in this chapter differs from the standard morphological approaches with respect to the status assigned to the part added in (8). Contrary to what most morphological approaches imply, I will argue that the presence or absence of the subject plays a fundamental role in accounting for the morphological cooccurrence restrictions introduced in this section.

7.4
Morphological or Syntactic Locality?

Traditionally, most analyses that consider cluster restrictions like (4) (7) propose a morphological approach to the problem. At the same time most proposals exhibit a strong connection with general syntactic principles; in fact, on the basis of these restrictions, many authors have argued that the morphological component has access to syntactic information of some sort.

7.4.1
Locality in Morphology

Albizu (1997) proposes the condition in (9) (his (24)) to account for the cooccurrence restrictions affecting clitic/agreement marker clusters presented in the previous section. 5

(9) *Person-Case Constraint (PCC)*
"A person-morphosyntactic feature P1 must be less referential than,
or as equally referential as, a person-morphosyntactic feature P2 that
locally c-commands it at the Morphological component."

The notion of local c-command that Albizu introduces applies in the morphological component and is parasitic on the syntactic notion of minimal link introduced in section 7.2. Albizu also assumes a hierarchical structure where the dative element locally c-commands the absolutive or

accusative head and the ergative element is higher in the structure. Given that, let us see how Albizu's system will work for the cases presented in (4b) and (5b) and repeated in (10).

(10) a. *Pedro le me envía.
Pedro CL.3DAT CL.1ACC send.3NOM
'Pedro sends me to him.'
b. *Azpisapoek etsaiari ni saldu na -i -o -te.
traitors-ERG enemy-DAT me-ABS sell 1ABS -3DAT -3ERG
'The traitors sold me to the enemy.'

The condition in (9) rules out the accusative or absolutive first/second person in the presence of a third person dative that locally c-commands it, since the former is more referential than the latter, according to a referentiality scale that distinguishes first/second from third person; this would account for the ungrammaticality of (10a,b).

Albizu's proposal raises interesting questions concerning the nature and role of his morphological condition. To begin with, the morphological Person-Case Constraint (PCC) in (9) raises a conceptual problem, since it invokes notions of a very diverse nature. In particular, it is not at all obvious why the morphological component should be sensitive to some scale of referentiality of the type introduced in this constraint. In addition, the introduction of a locality condition in the morphological component results in a certain degree of redundancy between morphology and syntax. If possible, this redundancy should be eliminated by reducing one of the conditions to the other, and it is reasonable to assume that the MLC is to be favored, as it is more grounded than its morphological counterpart; as I will show, whereas all the cases covered by the PCC have an immediate explanation in terms of the MLC, the reverse is not true. Moreover, I will show that a syntactic account is also favored over the morphological condition on empirical grounds. If this is correct, our goal should be to account for these clitic restrictions in terms of the Minimal Link Condition, which correctly prevents movement of more than two arguments out of VP. 6

Interestingly, a syntactic account in terms of the MLC poses new interesting problems; given that not all three-way agreement combinations are banned in natural languages, apparent violations of the MLC must be accounted for independently. In particular, we must explain why combinations such as (11a,b) are not ruled out by the MLC, even though they involve three arguments moving to their corresponding Case-checking positions.

(11) a. Pedro me lo envía.
Pedro CL.1DAT CL.3ACC send.3NOM
'Pedro sends him to me.'
b. Azpisapoek etsaiari misilak saldu d -f -izki -o
traitors-ERG enemy-DAT missiles-ABS sell -3ABS -3DAT
-te.
-3ERG
'The traitors sold missiles to the enemy.'

Until I come back to this problem in section 7.5, let us assume there is a way to fix it. I now show that there is strong empirical evidence for a syntactic analysis in terms of the MLC.

7.4.2
Syntactic Properties of Clitic/Agreement Restrictions

7.4.2.1
Absence of the Third Argument

In section 7.3 I showed that in ergative languages no restriction is imposed on the agreement system when the ergative argument is absent. Examples such as (12a,b) are thus grammatical, even though they involve both absolutive and dative agreements and in (12b) the "more referential" absolutive is locally c-commanded by the dative.

(12) a. Jon etorri f -zai -t.
Jon-*abs* come 3ABS -1DAT
'Jon came to me.'
b. Joni joan na -tzai -o
Jon-DAT go 1ABS -3DAT
'I went to Jon.'

In his analysis of this paradigm, Albizu (1997) follows Bobaljik's (1993) proposal for unaccusative subjects in ergative languages and assumes that the absolutive subject first moves to AgrOP to check Case and then moves to [Spec, TP]. After the second movement, which is motivated by the Extended Projection Principle (EPP) feature (a nominal feature in T), the absolutive argument ends up in a position higher than the dative argument. Consequently, the absolutive is not, according to Albizu, locally c-commanded by the dative argument.

Independently of whether this analysis of unaccusative predicates is correct, the syntactic movement of the absolutive subject to [Spec, TP] over the dative does not have an overt reflex in the morphology of the

auxiliary. Thus, factually there is no change internal to the complex auxiliary head. The absolutive and the dative markers occupy the same position relative to each other in both the grammatical example in (12b) and the impossible one in (13): the first person absolutive prefix *na-* precedes the dative marker *-o* in both the impossible form *naiote* of (13) and the grammatical form *natzaio* of (12b), the only difference being the presence of an additional ergative agreement in the former but not in the latter.

(13) *Azpisapoek etsaiari ni saldu na -i -o -te.
traitors-ERG enemy-DAT me sell 1ABS -3DAT -3ERG
'The traitors sold me to the enemy.'

If there is a correlation between the linear order of the affixes and their hierarchical relationas both Albizu and I assumethe dative affix locally c-commands the absolutive marker in both morphological structures, and the PCC will not be able to explain this contrast. 7

On the other hand, an explanation of the contrast in (12) (13) in terms of the MLC is straightforward. Independently of whether further movement of the absolutive argument takes place, in the absence of a third argument the MLC will not block movement of either the dative or the absolutive to their corresponding checking positions, and (12b) will be ruled in. In contrast, in the presence of the additional subject argument in (13), the MLC blocks attraction of the indirect object, as shown in section 7.2. This accounts for the contrast between unaccusatives and ditransitive predicates illustrated in (12b) and (13), respectively.

7.4.2.2
Nonargumental Clitics/Agreement Markers

Somewhat different is the case of nonargumental clitics, which coexist with argument clitics or agreement markers in the same clause. The Catalan example in (14), from Bonet 1991, is one such case. In this example first person ethical clitics cooccur with third person dative clitics. However, the nonargumental clitic does not seem to interfere with the argumental ones, and no person restriction applies (see Bonet 1991, Laka 1993, Albizu 1997, and references there).

(14) Me li van dir que havia suspes l'exam.
CL.1ETH CL.3DAT said that had-3NOM failed the exam
'They told him (on me) that he had failed the exam.'

Similar cases can also be found in agreement systems. In addition to the three grammatical markers, the auxiliary in Basque shows an optional

extra-argumental second person agreement, the so-called allocutive agreement, which refers to the addressee in the discourse. As in the case of Romance ethical datives, absolutive first and second person agreement is compatible with the presence of an allocutive dative marker in the auxiliary complex (example from Albizu 1997).

(15) Peruk ni kalean ikusi na -i -k -f.
Peru-ERG me-ABS street-the-in see 1ABS -2ALLOC -3ERG
'Peru saw me in the street (I am addressing you-male).'

To explain why the sentences in (14) (15) do not violate the PCC, Albizu proposes that the allocutive and ethical clitics are in an F projection higher than TP/AgrSP and do not locally c-command the absolutive in AgrO (see Oyharçabal 1993). In that case according to Albizu, the PCC does not block a combination where the absolutive is "more definite" than the dative element. The problem with this explanation is that the lack of local syntactic relations between an argumental clitic or agreement marker and the argumental ones has no morphological effect, once again as in the cases discussed in section 7.4.2.1: thus, the dative allocutive marker in Basque appears in the same morphological position as argumental dative agreements do in nonallocutive forms. Therefore, given that the condition applies to morphological structures, the local c-command relations are the same within the auxiliary in the two cases, in which case the PCC should treat the two forms alike.

In contrast, a syntactic explanation in terms of the MLC is straight-forward in this case as well: whether the allocutive and ethical elements are somewhere higher than TP or outside the syntactic system altogether, it is agreed that they are not part of the argument structure of the verb; it is thus natural to assume that they are not base-generated inside the verbal projection. Consequently, the nonargumental elements do not compete to enter into a checking relation, and the MLC imposes no restriction on the movement of the only two real arguments in the VP shell, as in regular transitive constructions.

7.4.2.3
Strong/Weak Dative Pronouns

The behavior of dative strong pronouns constitutes a potential problem for the MLC approach proposed here. In support of his morphological proposal, Albizu (1997) observes that nonclitic dative pronouns do not block first and second person accusative clitics. Compare the two examples in (16) (from Albizu 1997).

(16) a. A él me recomendó Juan.
to him CL.1ACC recommend.3NOM Juan
'To him, Juan recommended me.'
b. *Me le recomendó Juan.
CL.1ACC CL.3DAT recommend.3NOM Juan

At first glance the contrast in (16) might suggest that the restriction is morphologically driven. However, the properties of clitic and strong dative pronouns are not comparable at all. There are strong reasons to believe that nonclitic dative arguments in Romance have properties of prepositional phrases and that they do not check Case with the verbal system but rather with the prepositional head to which they are attached. In this regard, the *a* element in (16a) behaves like a real preposition comparable to English *to*. Various authors have argued that the clitic/strong dative pronoun alternation illustrated in (16) corresponds to the double object/*to* construction alternation observable in languages like English; 8 the *a* complementdemoted, in some sensechecks Case in some position internal to the prepositional phrase, independently of the verbal agreement system. If this is correct, the dative argument in (16a), a prepositional phrase, is not a candidate to be attracted out of the VP shell to check Case and only two arguments compete in these constructions; the result is grammatical, as expected. In contrast, the dative clitic in (16b) competes with the subject and the direct object as a possible target for the attraction of Case features, which yields an MLC configuration.

7.4.2.4
Nonfinite Clauses

The most serious challenge to any syntactic analysis of the PCC in terms of the MLC is perhaps the case of nominalized clauses in Basque, which show no PCC effects. Tenseless constructions in Basque are formed by direct attachment of an inflectional morpheme *-tzea* to the root of the verb; there is no overt auxiliary and no agreement marking on the verbal complex. Despite this difference in the agreement pattern, the case morphology on the arguments is exactly the same as in tensed clauses: overt ergative, absolutive, and dative markings are manifested on the noun phrases. What makes these constructions particularly interesting in the context of our discussion is that these tenseless clauses, where the agreement markers are not realized, allow combinations of arguments that are not possible when the auxiliary maintains overt agreement and the PCC shows its effects. Compared with the impossible (17), which shows overt agreement on the auxiliary, the

nominalized version in (18), where no agreement marker appears on the verb, is perfectly grammatical (Laka 1993, Albizu 1997).

(17) *Zuk ni etsaiari saldu na -i -o -zu
you-ERG me-ABS enemy-DAT sell 1ABS -3DAT -2ERG
'You sold me to the enemy.'

(18) Gaizki iruditzen zait [zuk ni etsaiari saltzea].
wrong seem AUX [you-ERG me-ABS enemy-DAT sell-NOMIN]
'Your selling me to the enemy seems wrong to me.'

Albizu (1997) argues that the contrast in (17) (18) corroborates the morphological nature of the agreement restrictions; when the auxiliary system does not encode the agreement relations morphologically, the morphological restriction does not seem to apply, and the absolutive argument can coexist with any dative and ergative argument.

However, as Albizu himself shows on the basis of Basque and Chukchi (Comrie 1981), in fact no agreement relation holds between the arguments and the verb in the infinitival constructions. If correct, this observation suggests that nominal arguments in infinitival constructions do not need to check Case with the verbal complex. Consequently, whatever the relevant licensing relations are, no movement is necessary in constructions like (18). Albizu's observation raises a very general question about how Case checking takes place in these constructions. If they do not check Case at all, as Albizu's conclusion implies, standard approaches to Case theory, where all Case features must be checked, are at best problematic. This result already supports the claim that the Case-theoretic approach to movement and its connection with agreement relations must be seriously revised, as suggested in section 7.1. On the other hand, these constructions do not constitute a direct counterexample to the MLC approach I have advocated here; since no argument movement takes place from VP, no MLC effects are expected. 9

7.5
Strong/Weak Pronouns and Restrictions on Third Person Pronouns

So far I have argued that the MLC elegantly accounts for the various crosslinguistic restrictions on the distribution of first and second person clitic and agreement clusters. What remains to be explained is why these restrictions do not apply to all structures with three arguments base-generated in the VP shell and, in particular, why third person object

agreement/clitics can coexist with dative and nominative/ergative arguments, apparently violating the MLC. I address this question next.

7.5.1
The Animacy Effects

There is a very general phenomenon in Romance that suggests that the restrictions analyzed so far go beyond the clitic system; at the same time this phenomenon will give us the clue to understand the still unexplained cases of grammatical ditransitive sentences. As described in section 7.3, combinations of three arguments are possible, so long as the accusative/absolutive argument is third person. However, although combinations like (19a) are grammatical, not all such combinations are possible. In particular, when the third person accusative clitic in (19a) is replaced by its corresponding strong pronoun, as in (19b), if the dative clitic is also present, the sentence becomes ungrammatical. 10

(19) a. Te la llevamos.
CL.2DAT CL.3ACC brought.1PL.NOM
'We brought her to you.'
b. *Te (la) llevamos a ella.
CL.2DAT (CL.3ACC) brought.1PL.NOM to her
'We brought her to you.'

The contrast between the ungrammatical (19b) and grammatical cases like (19a) challenges the empirical adequacy of morphological approaches to clitic clusters, at least if we want to unify our analysis with restrictions on agreement and clitic clusters of the type discussed here. This is so because the morphological approach has nothing to say about the unexpected ungrammaticality of (19b), where no clitic cluster is involved. If instead we can show that strong pronouns have some property in common with first and second person pronouns that is relevant in all cases, then the two phenomena are amenable to a unified account. The relevant property must distinguish between the following two groups: (a) first and second person pronounsin both clitic and strong guisestogether with third person strong pronouns, and (b) third person clitics. Once again, Case cannot be the relevant property that distinguishes these two classes, since it is standardly assumed that all members in both groups share accusative Case.

A well-known observation about strong pronouns like *él/ella* 'him/her' in Romance is that they are inherently [+animate] in all syntactic positions where there is a contrast between strong and weak pronouns (see

Perlmutter and Oresnik * 1973, Jaeggli 1984, and references there). This property distinguishes them from third person clitics and also from demonstrative pronouns. Note that, as (20) (21) show, strong pronominals cannot corefer with expressions denoting inanimate objects as their antecedents, but clitic pronouns can.11

(20) a. (La) vimos a ella en esa librería (a tu amiga).
(CL.3SG.FEM) saw to her in that bookstore (to your friend)
'We saw her (your friend) in that bookstore.'
b. *(La) vimos (a) élla en esa librería (la mesa).
(CL.3SG.FEM) saw (to) her in that bookstore (the table)
'We saw it (the table) in that bookstore.'

(21) La mesa/Tu amiga, la vimos en esa librería.
the table/your friend CL.3SG.FEM saw in that bookstore
'The table/Your friend, we saw it/her in that bookstore.'

It also seems natural to assume that first and second person pronouns share the inherent [+animate] property with strong pronouns, given their deictic reference to the speaker and the addressee. In that case the animacy property makes the right cut between elements that may enter into the ditransitive constructions explored in this chapter and those that do not; third person object clitics, which do not exhibit the [+animate] feature, do not move to a Case-checking position. Consequently, only objects that are not [+animate] may appear in ditransitive contexts. Remember that the subject and the indirect object will move clearly out of the VP shell and that the MLC restricts the number of moving elements to two. If so, the inanimate object in our double object constructions must somehow be allowed to remain within VP.

Considering all the above, let us then assume that the movement of arguments from the VP shell is not driven by Case-checking considerations, but by the need to check an animacy feature: only [+animate] arguments are forced to move from their base-generated position. As a consequence, whenever three [+animate] arguments are realized, they induce MLC violations, and the derivation is canceled. This accounts for the restrictions on first and second person clitic and agreement patterns as well as on the appearance of strong pronouns.12

7.5.2
Other Objects Unmarked for the Animacy Feature

As the contrast between (22a,b) and (23) illustrates, demonstrative pronouns and referential expressions differ from strong pronouns and pattern

together with third person clitics in that they cooccur with overt dative clitics. 13

(22) a. Te llevamos (a) ésta/ésa.
CL.2DAT brought.1PL.NOM (to) this/that.FEM
'We brought this/that one to you.'
b. Te llevamos el libro/(a) la niña.
CL.2DAT brought.1PL.NOM the book/(to) the girl
'We brought you the book/the girl.'

(23) *Te (la) llevamos a ella.
CL.2DAT (CL.3ACC) brought.1PL.NOM to her
'We brought her to you.'

As expected, this pattern correlates with the animacy properties of the elements involved. Unlike strong pronouns, demonstrative pronouns are not inherently specified with regard to animacy, and they can directly refer to inanimate objects. This is illustrated in (24); as (24a) shows, both clitics and demonstrative pronouns can refer to an inanimate object (e.g., a novel), but again this is impossible in the case of strong pronouns like *ella* 'her' in (24b).

(24) a. La he leído./He leído ésta.
it (CL) have read/have read this.fem
'I read it(clitic)/this one.'
b. *He leído (a) ella.
have read (to) it(STRONG PRON)
'I read it(strong pronoun).'

In the same way, it seems natural to assume that definite expressions do not have to fulfill any inherent animacy requirement in the syntax. It is important to stress the difference in that respect between strong pronouns and definite DPs. I am assuming that strong pronouns have a [+animate] feature that restricts their syntactic distribution in certain ways, including their binding possibilities and, if I am correct, their structural position in the clause. Instead, whether definite DPs take a referent that is animate or inanimate is an issue that belongs to the semantic or the pragmatic component and has no direct effect on their syntactic behavior. It is thus not very surprising that behaviorally they should group with other objects that are not specified as [+animate] such as clitics and demonstrative pronouns.

The behavior of demonstratives and definite expressionswhich, as I have shown, pattern together with cliticssuggests that animacy, rather than any partition based on inherent or relational properties of clitics, makes the right distinction between objects that move out of the VP shell to a checking domain and those that stay in situ. The animacy feature cuts across the clitic/nonclitic distinction: demonstratives (*este/esto* 'this one, etc.') and third person clitics (*lo/la*, etc.) are not [+animate] and consequently cannot check the relevant feature with T; they must therefore remain in their base-generated position unless, of course, they are attracted to check other features (e.g., a *wh*-feature). On the other hand, strong pronouns as well as first and second person pronouns of all types must check their inherent [+animate] feature with n/T. As a consequence, only the latter compete with other syntactic arguments, inducing MLC effects. In the next section I will present independent evidence that supports the distinction proposed here.

7.6
Objects That Don't Move

The classification in terms of animacy proposed in the previous section is supported by the behavior of the elements that belong to each group. In this section I show (a) that clitics and demonstratives surface in places typically analyzed as non-Case positions and (b) that they appear syntactically lower than strong pronouns. The evidence discussed in this section points at the same time to the need for a further distinction among the inanimate objects: whereas third person clitics and demonstrative pronouns seem to stay in situ, the behavior of definite expressions suggests that they move out of their base-generated position.

7.6.1
Binding Asymmetries

One reason to assume that clitics and strong pronouns occupy different structural positions is that they behave differently with respect to binding theory. Campos (1994) observes that the relationship between an R-expression and a clitic that can act as its potential antecedent yields very different Condition C results depending on the type of adjunct the R-expression is inserted in. He notes the following contrast:

(25) a. Ana lo oyó porque Jon hizo mucho ruido.
Ana CL.3ACC heard because Jon made a lot noise
'Ana heard him because Jon made a lot of noise.'

b. Ana lo vio cuando Jon entraba en casa.
Ana CL.3ACC saw when Jon entered in house
'Ana saw him when Jon entered the house.'
c. *Ana loi visitó en la casa que Joni compró en la
Ana CL.3ACC visited in the house that Jon bought on the
playa.
beach
'Ana visited him in the house that Jon bought on the beach.'

The sentences in (25a,b), where the R-expression is inside a *because* clause and a temporal adjunct, respectively, allow coreference between the clitic and the R-expression. In contrast, in (25c), where the referential expression is embedded in a locative phrase and the clitic is in the matrix object position, coreference is impossible.

Campos interprets these facts as evidence for the existence of an object agreement projection outside VP to which nominal arguments move to check Case. He proposes that locative phrases are generated adjoined to VP, whereas temporal and causal adjuncts adjoin to a higher projection. He then argues that the clitic *lo* must be in a position higher than the locative phrase adjoined to VP but lower than causal and temporal clauses. From the agreement position, the clitic would not be high enough to c-command the temporal or causal clauses; however, the clitic would c-command the locative phrase attached to VP and therefore bind the R-expression inside, inducing Condition C effects.

If Campos is correct, third person clitics move out of their base-generated position, contrary to what I have argued here. However, there is an interpretation of the paradigm in (25) that is more compatible with the claim I made in the previous sectionnamely, that elements unmarked for the animacy feature, among them third person clitics, are not required to move from their VP-internal position in the syntax unless other features must be checked. Suppose that the locative argument in (25c) is generated as the innermost argument in the VP shell, as argued by Larson (1988) in more general terms. If this is correct, the position occupied by the clitic is still higher than the locative expression and, consequently, the clitic will c-command the R-expression inside the clause. On the other hand, other adverbial modifiers are structurally higher, as Campos (1994) himself observes, and thus the clitic in the matrix clause cannot c-command the R-expression within these adjuncts.

Assuming the previous discussion, compare the paradigm in (25) with the behavior of the third person strong pronoun *él* 'him' in (26).

(26) a. *Ana (lo) oyó a éli porque Joni hizo mucho ruido.
Ana (CL.3ACC) heard to him because Jon made a lot noise
'Ana heard him because Jon made a lot of noise.'
b. *Ana (lo) vio a éli cuando Joni entraba en casa.
Ana (CL.3ACC) saw to him when Jon entered in house
'Ana saw him when Jon entered the house.'
c. *Ana (lo) visitó a éli en la casa que Joni compró en
Ana (CL.3ACC) visited to him in the house that Jon bought in
Paris.
Paris
'Ana visited him in the house that Jon bought in Paris.'

When the object clitic *lo* 'him' is replaced by a strong pronoun of its kind, coreference with an R-expression embedded in a modifier is impossible not only if the modifier is a locative clause but also with temporal or causal clauses in general. This difference points to a structural difference in the position clitics and strong pronouns occupy; in particular, it suggests that the strong pronoun occupies a position structurally higher than the clitic, where its c-command domain is larger than that of the clitic. Thus, this contrast supports the conclusion that strong pronouns move to a higher position where they check their [+animate] feature, whereas clitics stay in their VP-internal position. 14

7.6.2
Clitics and Demonstratives as Caseless Expletives

An additional property shared by third person clitics and demonstratives indicating that they do not move up to a higher checking position is that they appear in what have traditionally been considered non-Case positions. For instance, both clitics and demonstratives may corefer with adjectival phrases in attributive and predicative contexts.15 This is illustrated by the contrast between the clitic and the strong pronoun in (27).

(27) a. Yo no soy bonita ni lo quiero ser.
I not am pretty nor CL.3ACC want to be
Lit. I am not pretty nor do I want to be it(clitic).
b. Yo no soy bonita ni quiero ser *eso*.
I not am pretty nor want to be *that*
c. *Yo no soy bonita ni quiero ser *ello*.
I not am pretty nor want to be *it*

Similarly, third person clitics and demonstratives may be introduced as sentential expletives, even in contexts where it is generally assumed that no Casein present terms, no animacy-checking relationis available. It has been well known since Grimshaw 1979 and Pesetsky 1982 that some predicates that subcategorize for clausal complements do not allow nominal complements, even if these nominals may satisfy the selectional requirements of the predicate. Thus, for instance, Grimshaw observes that some predicates that take *wh*-complements cannot take a semantically equivalent NP complement. Compare, for instance, (28) and (29).

(28) a. I asked what answer he gave.
b. I asked the answer he gave.

(29) a. I wonder what answer he gave.
b. *I wonder the answer he gave. (Grimshaw 1979, (92a), (93a))

On the basis of these differences, Grimshaw argues that only the interaction of subcategorization and selection can account for the complete distribution of the types of complements each predicate can take. Pesetsky argues that Grimshaw's cases may be accounted for independently by Case theory if we assume that the basic difference between, say, *wonder* and *ask* lies in their Case properties. Recast in Case-theoretical terms, the main difference that accounts for the contrast in (28) and (29) would be that *ask* checks accusative Case with its complement whereas *wonder* does not. As a consequence of that difference, the complement of *wonder* cannot be an NP, since, being subject to the Case Filter, it would not be able to receive Case from V.

The Spanish verb *dudar* 'to doubt' shows the same restriction as *wonder*-type verbs; it takes clausal complements, but needs the support of a prepositional element *de* 'of' in order to take a nominal complement. This is illustrated by the paradigm in (30).

(30) a. Dudo [CP que vengan].
doubt.1NOM that come.3SUBJUNCT
'I doubt that they will come.'
b. Dudo *(de) [DP tu honestidad].
doubt.1NOM (of) your honesty
'I doubt (about) your honesty.'

However, even if *dudar* does not take bare nominal complements, both the clitic *lo* 'it' and the demonstrative *esto* 'this one' can appear in the complement position of that verb. Compare in that respect the gramma-

ticality of clitics and demonstratives with the ungrammaticality of strong pronouns.

(31) a. (Eso,) lo dudo.
(that) CL.3ACC doubt.1NOM
'(That,) I doubt it.'
b. *Ello, lo dudo.
it CL.3ACC doubt.1NOM
'It(strong), I doubt.'

This contrast strongly supports the conclusion that third person clitics and demonstratives do not have the checking requirements other expressions have. At the same time it shows that there are differences among objects belonging to the group. The different behavior of definite expressions and clitics/demonstratives in the paradigm (30) (31) suggests that there are other relevant checking relations that distinguish among them. This distinction is not at all surprising, given the very different nature of the members of the class of objects that are not [+animate], but it suggests the need for a more fine-grained analysis of the checking relations responsible for the displacement of verbal arguments, which is beyond the scope of this chapter.

7.7 Concluding Remarks

What initially might have been taken as a purely technical question about the structural limitations of the Minimal Link Condition now becomes a more interesting issue: some general morphological properties of typologically very diverse languagesthe well-known person restrictions on object clitics and agreement markersseem to follow from the general architecture of the computational component and, more specifically, from the MLC. If correct, this result is an interesting piece of evidence in favor of current minimalist approaches. At the same time the consequences of the approach pursued here force us to adopt fairly important changes in the theory of Case and movement. This gives rise to new, interesting questions concerning the nature of feature checking as a uniformly asymmetric operation: unlike syntactic Case, which is a theory-internal construct, features like [±animate] point to a conception of checking operations where the feature checked is interpretable in either the target or the attracted element, and it survives to one of the interface levels. Many issues as diverse as the behavior of exceptional-Case-marking sub-

jects or the properties of referential DPs must thus be reconsidered in more detail. However, the line of research proposed here opens a promising way to derive Case theory from a system that is more uniform conceptually and more consistent with the general spirit of the Minimalist Program.

Notes

I am very pleased to participate in this volume in honor of Howard Lasnik; let this be my small contribution in acknowledgment for a whole life of dedication to those of us who have had the fortune to be his students.

Parts of the material in this chapter were presented at the Universities of the Basque Country, British Columbia, and Connecticut, at the Linguistic Seminar at Deusto, and in a graduate seminar on the syntactic structure of the Basque language that I taught at the University of Maryland. I am very grateful to the audiences, especially to my students in Maryland, for helpful comments and discussion. I am particularly indebted to Xabier Artiagoitia, Željko Bošković, Hamida Demirdache, Jon Franco, Itziar Laka, Joseba Lakarra, Alazne Landa, Juan Uriagereka, Myriam Uribe-Etxebarria, Vidal Valmala, and an anonymous reviewer for helpful comments and discussion. Obvious conceptual differences aside, my analysis of the phenomena discussed here owes much to previous work by Bonet (1994) and very especially to Albizu's (1997) proposal. It also coincides in many respects with a proposal made by Romero (1997. chap. 5), with whom I also had very interesting discussions; some extensions of common ideas are presented in Ormazabal and Romero 1998. This research was supported in part by a CICYT (Interministerial Commission of Science and Technology) grant (PB96-0272), a grant from the Basque government (PI-1998-127) to the Basque Center for Language Research (LEHIA), and by a grant from the University of the Basque Country (UPV 033.130-HAO36/98).

1. If equidistance is eliminated from the definition of the MLC in favor of a more restrictive condition and crossing is banned on more general groundsa possibility explored in Chomsky 1995, 133ff.movement of the object to the inner specifier of nP, crossing the dative argument, is impossible and the original problem arises in a more general way.

2. The closest to a solution that I know of has been proposed by Fernández (1997). However, her proposal works only under very specific assumptions about Shortest Movement, combined with a structure like that suggested by Koizumi (1993), that are not compatible with the view assumed in this chapter. See also Cheng and Demirdache 1993, Collins and Thráinsson 1993, and Uriagereka 1997 for different possible accounts of three-way agreement systems from different theoretical perspectives.

3. Notice that the ungrammaticality of (4b) is not due to a clitic-ordering restriction; thus, the opposite order in (i) is also impossible with the relevant reading (see section 7.4.2.2).

(i) *Pedro me le envía.
Pedro CL.1ACC CL.3DAT send.3NOM
'Pedro sends me to him.'

In this chapter I will focus only on restrictions related to person combinations and propose a syntactic explanation for them, leaving aside morphological conditions on clitic ordering. In contrast with the universality of the former, the latter show great dialectal variation, which suggests a more morphophonological explanation (see Bonet 1991 and references therein).

I will use the following abbreviations in the examples: CL = clitic, ACC = accusative, ABS = absolutive, DAT = dative, NOM = nominative, ERG = ergative, 1 = first person, and so on. In most Romance languages, verbal agreement with the nominative argument is overtly manifested at the rightmost position of the verbal complex. In addition, accusative and dative clitics show up generally to the left of the verb in tensed clauses. Basque, on the other hand, is an ergative language; the auxiliary shows agreement with the three main arguments of the predicate: the ergative, absolutive, and dative arguments. I leave aside details that are not relevant for the discussion. See Ortiz de Urbina 1989, Laka 1993, Gómez and Sáinz 1995, Albizu and Eguren 1995, Fernández 1997, and references there for discussion.

4. See Aissen 1987, Bonet 1994, Rosen 1990, Laka 1993, Elordui 1995, Albizu 1997, and references there for discussion of this restriction in languages with rich agreement.

5. Albizu's condition is based on Franco 1993, 46; see also Murasugi 1992.

6. But see Zwart's (1997) proposal in the opposite direction.

7. Albizu himself observes that, given the grammaticality of (12), the PCC cannot be a condition on the morphological output. This result forces him to adopt a theory of feature copying after movement, instead of checking.

8. See Uriagereka 1988, Demonte 1995, Romero 1997, and references there. There are important differences in that respect between the dative prepositional element *a* 'to' in all Romance languages and the marker that appears with animate direct objects in Spanish and Portuguese. Unlike the former, the latter does not seem to behave like a prepositional element.

9. For a more detailed analysis of these constructions in terms of Attract F, see Ormazabal and Romero 1998.

10. The optional accusative clitic in parentheses, which does not change the acceptability of the example, has been added to control for the possible effects of clitic doubling.

11. Juan Uriagereka (personal communication) observes that the animacy property of strong pronouns is not uniform across all Romance languages in all configurations. For instance, strong pronouns may appear in subject position referring to inanimate objects in Western Iberian languages; thus, the following Galician sentence is possible:

(i) Ela e forte, esta mesa.
she is strong this table
'It's strong, this table.'

It is worth noticing in this respect that Romance languages also differ with regard to the inventory of clitics and that, unlike French, Iberian Romance languages also lack subject clitics. We could guess that inanimate strong pronouns more plausibly appear in subject than in object position.

12. An anonymous reviewer observes that an additional condition must be met for the analysis to be valid: dative clitics must be [+animate] even in the third person, for otherwise no MLC violation occurs in these contexts. In general, that is the case in Romance; the only exceptions I am aware of are cases of locative dative clitics of the type exemplified in (i).

(i) Le puse la pata a la mesa.
CL.3DAT put the leg to the table
'I put the leg on the table.'

However, given their properties, these constructions cannot be the most appropriate ones to test whether combinations of two [+animate] objects are indeed impossible, because the type of direct object allowed in these constructions is restricted enough not to make an animate object pragmatically very felicitous. See Romero 1997 for an analysis of these constructions.

13. Some speakers have trouble getting the relevant direct object reading in (22a,b) when the animacy marker *a* is present; they tend to interpret it as a prepositional element heading an indirect object. This would in fact be rightly predicted by the account in this chapter if the animacy feature to be checked extends to all contexts where animacy is syntactically present. Unfortunately, I believe that these speakers' difficulty is due to purely pragmatic facts. In fact, given the right context, many speakers find (i) perfectly grammatical both with and without the *a* marker.

(i) Te dejamos (*a*) la niña en casa a las tres en punto.
CL.2DAT leave.1PL.NOM (to) the girl at home at three o'clock
'We will leave you the girl at home at three o'clock.'

14. Luján and Larson (1988) argue that Romance strong pronouns have the properties of English stressed pronouns and, unlike clitics, are focalized; if focus and syntactic scope are directly connected, these factors might also be playing a role in the contrasts discussed in the text. I leave this possibility open for further research.

15. I am indebted to Juan Uriagereka for this observation.

References

Aissen, Judith L. 1987. *Tzotzil clause structure*. Dordrecht: Kluwer.

Albizu, Pablo. 1997. Generalized Person-Case Constraint: A case for a syntax-driven inflectional morphology. In Amaya Mendikoetxea and Myriam Uribe-Etxebarria, eds., *Theoretical issues on the morphology-syntax interface*, 1 34. ASJU, San Sebastian.

Albizu, Pablo, and Luis Eguren. 1995. A Distributed Morphology based approach to ergative displacement in Basque. Ms., University of Southern California and Universidad Autónoma de Madrid.

Aoun, Joseph. 1982. The formal nature of anaphoric relations. Doctoral dissertation, MIT, Cambridge, Mass.

Bobaljik, Jonathan. 1993. Ergativity and ergative unergatives. In Colin Phillips, ed., *Papers on Case and agreement II*, 45 88. (MIT Working Papers in Linguistics 19.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Bonet, Eulàlia. 1991. Morphology after syntax: Pronominal clitics in Romance. Doctoral dissertation, MIT, Cambridge, Mass.

Bonet, Eulàlia. 1994. The Person-Case Constraint: A morphological approach. In Heidi Harley and Colin Phillips, eds., *The morphology-syntax connection*, 33 52. (MIT Working Papers in Linguistics 22.) MITWPL. Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Boskovic, Zeljko. * 1995. Principles of economy in nonfinite complementation. Doctoral dissertation, University of Connecticut, Storrs.

Campos, Héctor. 1994. Accusative and dative Case-checking in Spanish. Ms., Georgetown University, Washington, D.C.

Cheng, Lisa L.-S., and Hamida Demirdache. 1993. External arguments in Basque. In Hualde and Ortiz de Urbina 1993, 71 87.

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Collins, Chris, and Höskuldur Thráinsson. 1993. Object shift in double object constructions and the theory of Case. In Colin Phillips. ed., *Papers on Case and agreement II*, 131 174. (MIT Working Papers in Linguistics 19.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Comrie, Bernard. 1981. *Language universals and linguistic typology*. Oxford: Blackwell.

Demonte, Violeta. 1995. Dative alternation in Spanish. *Probus* 7, 5 30.

Elordui, Agurtzane. 1995. Hizkuntza heriotza eta aldaketa morfologikoa (Language death and morphological change). Doctoral dissertation, University of the Basque Country, Vitoria-Gasteiz.

Fernández, Beatriz. 1997. Kasu teoria euskaraz (Case theory in Basque). Doctoral dissertation, University of the Basque Country, Vitoria-Gasteiz.

Franco, Jon. 1993. On object agreement in Spanish. Doctoral dissertation, University of Southern California, Los Angeles.

Gómez, Ricardo and Koldo Sáinz. 1995. On the origin of the finite forms of the Basque verb. In Jose I. Hualde, Joseba A. Lakarra, and R.L. Trask, eds., *Towards a history of the Basque language*. 235 274. Amsterdam: John Benjamins.

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10, 279 326.

Hualde, Jose I., and Jon Ortiz de Urbina, eds. 1993. *Generative studies in Basque linguistics*. Amsterdam: John Benjamins.

Jaeggli, Osvaldo. 1973. *Issues in Romance linguistics*. Dordrecht: Foris.

Koizumi, Masatoshi. 1993. Object agreement phrases and the split VP hypothesis. In Jona than D. Bobaljik and Colin Phillips, eds., *Papers on Case and agreement I*, 99 148. (MIT Working Papers in Linguistics 18.) MITWPL, Department of Linguistics, MIT, Cambridge, Mass.

Laka, Itziar. 1993. The structure of inflection: A case study in X-zero syntax. In Hualde and Ortiz de Urbina 1993, 21 70.

Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19, 335 391.

Lasnik, Howard. 1995. Case and expletives revisited: on Greed and other human failings. *Linguistic Inquiry* 26, 615 633.

Luján, Marta, and Richard Larson. 1988. Paper presented at Richard Larson's seminar, MIT, Cambridge, Mass.

Martin, Roger. 1996. A minimalist theory of PRO and control. Doctoral dissertation, University of Connecticut, Storrs.

Murasugi, Kumiko. 1992. Crossing and nested paths: NP movement in accusative and ergative languages. Docotral dissertation, MIT, Cambridge, Mass.

Ormazabal. Javier and Juan Romero. 1998. Attract-F: A case against Case. Paper presented at the 21st GLOW Colloquium, Tilburg.

Ortiz de Urbina. Jon. 1989. *Parametters in the syntax of Basque*. Dordrecht: Foris.

Oyharçabal, Beñat. 1993. Verb agreement with non-arguments: On allocutive agreement. In Hualde and Ortiz de Urbina 1993, 89 114.

Perlmutter, David. 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.

Perlmutter, David and Janez Oresnik. * 1973. Language-particular rules and explanation in syntax. In Stephen R. Anderson and Paul Kiparsky, eds., *A festschrift for Morris Halle*. 419 459. New York: Holt, Rinehart and Winston.

Pesetsky, David. 1981. Paths and categories, Doctoral dissertation, MIT, Cambridge, Mass.

Romero, Juan. 1997. Construcciones de doble objeto i gramática universal (Double object constructions and Universal Grammar). Doctoral dissertation, Universidad Autónoma de Madrid.

Rosen, Carol. 1990. Rethinking Southern Tiwa: The geometry of a triple-agreement language. *Language* 66, 669 713.

Rouveret, Alain, and Jean-Roger Vergnaud. 1980. Specifying reference to the subject. *Linguistic Inquiry* 11, 97 202.

Uriagereka, Juan. 1988. On government. Doctoral dissertation, University of Connecticut, Storrs.

Uriagereka, Juan. 1994. Aspects of the syntax of clitic placement in Werstern Romance. *Linguistic Inquiry* 25, 79 123.

Uriagereka, Juan. 1997. Formal and substantive elegance in the Minimalist Program. In Chris Wilder, Hans-Martin Gärtner, and Manfred Bierwisch, eds., *The role of economy principles in linguistic theory*, 170 204. Berlin: Akademie-Verlag.

Zwart, C. Jan-Wouter. 1997. *Morphosyntax of verb movement: A minimalist approach to the syntax of Dutch*. Dordrecht: Kluwer.

Chapter 8
The Japanese Light Verb Construction and the Minimalist Program

Mamoru Saito and Hiroto Hoshi

The Japanese light verb construction has been discussed extensively in the literature. 1 The Japanese verb *su* functions exactly like the English main verb *do* in examples like (1).

(1) Mary-ga [NP suugaku-no) syukudai]-o sita. (= *su* + *ta* (past))
Mary-NOM (math-GEN homework-ACC did
'Mary did the (math) homework.'

But it can also be "void of meaning," as illustrated in (2) (3).

(2) a. Mary-ga John-to (kyonen) kekkon-sita.
Mary-NOM John-with last year marriage-did (married)
'Mary married John (last year).'
b. Mary-ga John-to (kyonen) [NP kekkon]-o sita.
Mary-NOM John-with last year marriage-ACC did
(same)

(3) a. Mary-ga John-to (kinoo) kaiwa-sita.
Mary-NOM John-with yesterday conversation-did (talked)
'Mary talked with John (yesterday).'
b. Mary-ga John-to (kinoo) [NP kaiwa]-o sita.
Mary-NOM John-with yesterday conversation-ACC did
(same)

In these examples both *Mary-ga* and *John-to* appear as clausal arguments: if they were within a projection of a noun (e.g., *kekkon* 'marriage' or *kaiwa* 'conversation'), they would have to be accompanied by the genitive Case marker *no*.2 In (2a) the incorporated noun *kekkon* assigns q-roles to these clausal arguments, and *su* functions merely as a "category-changing affix." In (2b) the q-role-assigning noun appears as the object of *su*, and *su* itself seems to be nothing more than an "expletive verb." The

construction exemplified by (2b) and (3b) is the so-called *Japanese light verb construction*.

In their classical analysis of this construction, Grimshaw and Mester (1988) propose the operation of *argument transfer*. According to this analysis, the noun *kekkon* 'marriage' in (2b) transfers its q-roles to the verb *su*, and the verb discharges them to the clausal arguments. Grimshaw and Mester discuss some peculiar properties of this construction and attribute them to the nature of the argument transfer operation. Further supporting arguments for this analysis are found in Miyagawa 1989 and Tsujimura 1990a.

In this chapter we argue instead for an LF incorporation analysis of the light verb construction. More specifically, we present evidence that *kekkon* 'marriage' in (2b) raises to the position of *su* in LF and discharges its q-roles at that level. 3 This analysis provides direct support for Chomsky's (1993) Minimalist Program, which proposes to eliminate the Projection Principle together with the levels of D-Structure and S-Structure. We argue that it also provides strong support for Chomsky's (1986b, 1993) Last Resort principle on movement.

In section 8.1 we briefly review the analysis proposed by Grimshaw and Mester (1988). In section 8.2 we present a piece of suggestive evidence for the LF incorporation analysis. In section 8.3 we return to the peculiar properties of the light verb construction discussed by Grimshaw and Mester and show that they follow directly from our analysis. In section 8.4 we consider the important observation made by Miyagawa (1989) and Tsujimura (1990a) that the q-role-assigning noun in the light verb construction cannot be ergative. There, we show that the LF incorporation analysis suggests plausible directions toward explaining this fact. In section 8.5 we summarize our discussion.

8.1
Grimshaw and Mester's (1988) Argument Transfer Analysis

One interesting property of the light verb construction is that some of the arguments of the q-role-assigning noun can remain within the projection of the noun, whereas others appear as clausal arguments. The examples in (4) illustrate this point.

(4) a. Honda-ga Amerika-de [NP Akoodo-no seisan]-o site-iru.
Honda-NOM U.S.-in Accord-GEN production-ACC doing-is
'Honda is producing Accords in the U.S.'

b. Mary-ga John-ni/-e [NP toti-no zyooto]-o sita.
Mary-NOM John-to/-to land-GEN giving-ACC did
'Mary gave a piece of land to John.'

In (4b), for example, the theme argument *toti* 'land' is within the NP headed by the q-role-assigning noun *zyooto* 'giving', as indicated by the presence of the genitive Case marker. The agent and goal arguments, on the other hand, appear as clausal arguments, outside the projection of *zyooto*.

Given this fact, Grimshaw and Mester (1988) propose that the q-role-assigning noun can transfer some (and possibly all) of its q-roles to the light verb *su*. Roughly put, when *zyooto* 'giving' transfers its agent and goal q-roles to *su*, we obtain the lexical entry in (6) from those in (5).

(5) a. *zyooto* (agent (goal (theme)))
b. *su* ( ) <ACC> 4

(6) *zyooto* (theme) + *su* (agent (goal)) <ACC>

This accounts for (4b), where *Mary-ga* and *John-ni/-e* appear in positions normally occupied by the arguments of the verb.

The theoretically important discussion in Grimshaw and Mester 1988 starts, however, with the authors' discovery that this operation of argument transfer needs to be constrained in a systematic way. If argument transfer is totally unconstrained, then the arguments of the q-role-assigning noun should be able to be realized either within the NP headed by the noun or outside it at the clausal level, quite freely. This is so because any q-role of the noun can optionally be transferred to *su*. But Grimshaw and Mester note the following rather surprising restrictions on the light verb construction:

(7) a. At least one internal q-role of the noun must be assigned to an argument outside the NP.
b. If a q-role T is assigned outside the NP, then all q-roles that are higher than T in the thematic hierarchy must also be assigned outside the NP.

In Grimshaw and Mester's terms, (7a) means that at least one internal q-role (in the sense of Williams 1981) must be transferred to the light verb *su*. The examples in (8) confirm this generalization.

(8) a. ?Mary-ga [NP John-to-no kekkon]-o sita. (cf. (2b))
Mary-NOM John-with-GEN marriage-ACC did
'Mary married John.'

b. ?Mary-ga [NP John-to-no kaiwa]-o sita. (cf. (3b))
Mary-NOM John-with-GEN conversation-ACC did
'Mary talked with John.'

As Grimshaw and Mester point out, examples of this kind are marginally allowed with the interpretation whereby *su* is a q-role-assigning verb, corresponding to the English main verb *do* (hence, ? instead of *). Under this interpretation, (8b), for example, roughly means that there was a specific act of engaging in conversation with John, and Mary did it. But (8a,b) are unacceptable with the "neutral interpretation" whereby *su* is taken as a light verb. This is what (7a) predicts, since in these examples only the external argument of the q-role-assigning noun appears at the clausal level.

If theme is the lowest q-role in the thematic hierarchy, (7b) implies that when the theme q-role is realized at the clausal level, all other q-roles must also be assigned at this level. Let us use the examples in (4) to see how this prediction is borne out. It should be noted first that when the theme q-role is realized outside the NP at the clausal level, the sentences in (4) become marginal, as shown in (9).

(9) a. ??Honda-ga Amerika-de Akoodo-o [NP seisan]-o
Honda-NOM U.S.-in Accord-ACC production-ACC
site-iru.
doing-is
'Honda is producing Accords in the U.S.'
b. ??Mary-ga John-ni/-e toti-o [NP zyooto]-o sita.
Mary-NOM John-to/-to land-ACC giving-ACC did
'Mary gave a piece of land to John.'

This is because of the presence of two accusative NPs within the sentence. The relevant "constraint" can be stated as in (10). 5

(10) *The Double*-o *Constraint*
A simple sentence cannot contain more than one *o*-marked phrase.

But when the theme argument appears outside the NP and some other argument occurs inside, the result is completely unacceptable, as shown in (11) (12).

(11) a. *Honda-ga Akoodo-o [NP Amerika-de-no seisan]-o
Honda-NOM Accord-ACC U.S.-in-GEN production-ACC
site-iru.
doing-is
'Honda is producing Accords in the U.S.'

b. *Amerika-de Akoodo-o [NP Honda-no seisan]-o
U.S.-in Accord-ACC Honda-GEN production-ACC
site-iru.
doing-is
(same)
c. *Akoodo-o [NP Honda-no Amerika-de-no seisan]-o
Accord-ACC Honda-GEN U.S.-in-GEN production-ACC
site-iru.
doing-is
(same)

(12) a. *Mary-ga toti-o [NP John-e-no zyooto]-o sita. 6
Mary-NOM land-ACC John-to-GEN giving-ACC did
'Mary gave a piece of land to John.'
b. *John-ni/-e toti-o [NP Mary-no zyooto]-o sita.
John-to/-to land-ACC Mary-GEN giving-ACC did
(same)
c. *Toti-o [NP Mary-no John-e-no zyooto]-o sita.
land-ACC Mary-GEN John-to-GEN giving-ACC did
(same)

(11a c) are variants of (9a), but in these examples at least one nontheme argument appears NP-internally: the locative *Amerika-de* 'in the U.S.' in (11a,c) and the agent *Honda* in (11b,c). The sharp contrast between (9a) and (11a c) clearly shows the validity of the generalization in (7b). The contrast between (9b) and (12a c) leads to the same conclusion.

(7a) and (7b) together imply that the subject q-role of the q-role-assigning noun must always be realized at the clausal level. This is so because at least one internal q-role must be assigned at the clausal level, and the subject (external) q-role is higher in the thematic hierarchy than any internal q-role. We have already provided ungrammatical examples that instantiate this generalization, namely, (11b,c) and (12b,c). More examples are shown in (13b,c) and (14b,c).

(13) a. Mary-ga John-to [NP kekkon]-o sita. (= (2b))
Mary-NOM John-with marriage-ACC did
'Mary married John.'
b. *[NP Mary-no John-to-no kekkon]-o sita.
Mary-GEN John-with-GEN marriage-ACC did
c. *John-to [NP Mary-no kekkon]-o sita.
John-with Mary-GEN marriage-ACC did

(14) a. Mary-ga John-to [NP kaiwa]-o sita. (= (3b))
Mary-NOM John-with conversation-ACC did
'Mary talked with John.'
b. *[NP Mary-no John-to-no kaiwa]-o sita.
Mary-GEN John-with-GEN conversation-ACC did
c. *John-to [NP Mary-no kaiwa]-o sita.
John-with Mary-GEN conversation-ACC did

Grimshaw and Mester (1988) argue that the restrictions in (7) should be attributed to the nature of argument transfer. They first note that in the light verb construction the object NP headed by the q-role-assigning noun is not an argument and hence must be licensed in some other way. They then propose that it is licensed by virtue of argument transfer itself. More precisely, they propose that the licensing obtains when the head noun transfers an obligatory argument, or an unsuppressed q-role, to *su*. This proposal straightforwardly accounts for (13b) and (14b), where no q-role is transferred to *su*. It also accounts for the generalization in (7a) if the subject q-role of a noun is suppressed as proposed by Zubizarreta (1985) and Grimshaw (1990), among others. Let us take the following examples for illustration:

(15) a. Mary-ga John-ni/-e [NP toti-no zyooto]-o sita. (= (4b))
Mary-NOM John-to/-to land-GEN giving-ACC did
'Mary gave a piece of land to John.'
b. ?Mary-ga [NP John-e-no toti-no zyooto]-o sita.
Mary-NOM John-to-GEN land-GEN giving-ACC did
(same)

Like (8a,b), (15b) is unacceptable with the light verb interpretation of *su*. The contrast in (15) is expected since *zyooto* 'giving' transfers an internal (unsuppressed) q-role to *su* in (15a), but not in (15b). This is shown in (16b,c), where unsuppressed q-roles are in full capitals.

(16) a. *zyooto* (agent (GOAL (THEME))), *su* ( ) <ACC> . . . input
b. *zyooto* (THEME) + *su* (agent (GOAL)) <ACC> . . . (15a)
c. *zyooto* (GOAL (THEME)) + *su* (agent) <ACC> . . . (15b)

For the generalization in (7b), Grimshaw and Mester propose that argument transfer applies in an outside-in fashion, transferring the outermost argument first and then affecting the inner ones in accordance with the hierarchy in the argument structure. If the external/internal distinction is represented structurally in the lexical specification of the predicate-

argument structure, as argued in Hale 1983, then the ungrammaticality of (17), in contrast with (15a), immediately follows.

(17) *John-ni/-e [NP Mary-no toti-no zyooto]-o sita.
John-to/-to Mary-GEN land-GEN giving-ACC did
'Mary gave a piece of land to John.'

In this example the goal q-role of *zyooto* 'giving' is transferred to *su*, but the outer agent q-role is not, as shown in (18).

(18) *zyooto* (agent (THEME)) + *su* (GOAL) <ACC>

Further, if the internal arguments themselves are hierarchically organized in the way proposed in Grimshaw 1990, this account extends to all instances of the generalization in (7b). Let us consider again the ungrammatical (12a), repeated in (19).

(19) *Mary-ga toti-o [NP John-e-no zyooto]-o sita.
Mary-NOM land-ACC John-to-GEN giving-ACC did
'Mary gave a piece of land to John.'

Here, the theme q-role, but not the goal q-role, is transferred to *su*, as shown in (20).

(20) *zyooto* (GOAL) + *su* (agent (THEME)) <ACC>

This kind of argument transfer is barred, as goal is structurally higher than theme in the representation of argument structure.

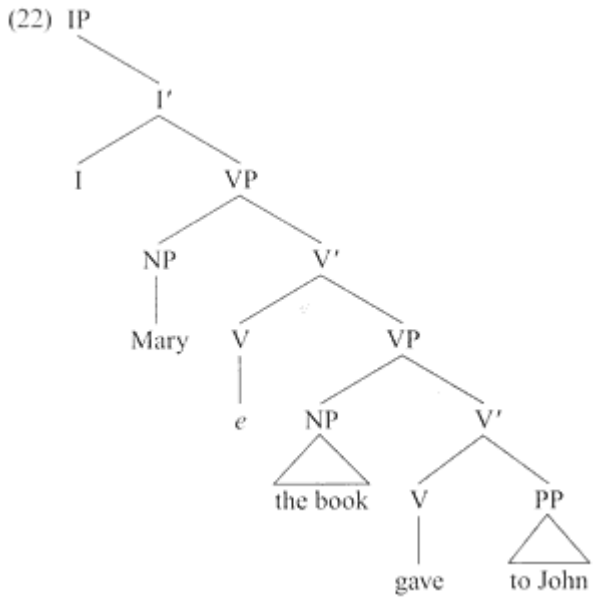
As Grimshaw and Mester note, their account for the generalizations in (7) relies crucially on some hypotheses regarding the lexical representation of argument structure. First, the external argument must be suppressed for nouns, as opposed to verbs. Second, not only the external/internal distinction but also the relative prominence relation among internal arguments must be represented structurally. Thus, the analysis illustrated above, if correct, provides support for these hypotheses, which are explored in detail in Grimshaw 1990.

8.2
LF Incorporation and the Double-o Constraint

Grimshaw and Mester's (1988) argument transfer analysis is based on the fundamental assumption that all q-roles must be assigned at D-Structure, owing to the Projection Principle. 7 But as noted in Chomsky 1993, works such as Larson 1988 cast doubt on this assumption. Larson proposes that the D-Structure representation of (21), for example, is roughly as in (22).

(21) Mary gave the book to John.

(22)



The most relevant aspect of this analysis is that the verb *gave* raises to the higher V position at S-Structure and assigns the agent q-role to the NP *Mary* at this level. This clearly implies that all q-roles need not be discharged at D-Structure. Chomsky (1993) takes a further step and proposes the Minimalist Program, which dispenses with the Projection Principle and eliminates the levels of D-Structure and S-Structure altogether. The only requirement on q-roles, then, is that they be properly assigned at LF.

Once we adopt the Minimalist Program, a completely new analysis of the light verb construction immediately suggests itself. That is, the q-role-assigning noun, heading the direct object NP, raises to the position of *su* in LF and discharges (some of) its q-roles at this level. According to this analysis, *zyooto* 'giving' in (23) assigns its theme q-role within the direct object NP and then raises to the position of *su* to discharge its goal and agent q-roles.

(23) Mary-ga John-ni/-e [NP toti-no zyooto]-o sita. (= (15a))
Mary-NOM John-to/-to land-GEN giving-ACC did
'Mary gave a piece of land to John.'

In this section we present a piece of suggestive evidence for this LF

incorporation analysis. Then, in section 8.3, we show that Grimshaw and Mester's generalizations in (7) follow straightforwardly from the analysis.

As noted above, the light verb construction is marginal when the theme argument of the q-role-assigning noun appears at the clausal level. The relevant examples in (9) are repeated in (24).

(24) a. ??Honda-ga Amerika-de Akoodo-o [NP seisan]-o
Honda-NOM U.S.-in Accord-ACC production-ACC
site-iru.
doing-is
'Honda is making Accords in the U.S.'
b. ??Mary-ga John-ni/-e toti-o [NP zyooto]-o sita.
Mary-NOM John-to/-to land-ACC giving-ACC did
'Mary gave a piece of land to John.'

It was also noted above that examples of this kind are ruled out by the Double-*o* Constraint, stated informally in (10). But as Sells (1988) points out, the examples in (24) are only mildly deviant, and they are much better than the "standard cases" of Double-*o* Constraint violations. Discussing the nature of Double-*o* Constraint violations in some detail, we suggest below that the better-than-expected status of (24a,b) is due to the LF incorporation of the q-role-assigning noun.

As shown by Harada (1973) and Kuroda (1978), among others, there are two distinct kinds of Double-*o* Constraint violations. The "standard" kind, called the *abstract* Double-*o* Constraint violation, obtains when two NPs require abstract objective Case, and it results in strong deviance. The most typical case of this type of violation is observed with the causative construction. As shown in (25b), the causee in this construction can be marked by the objective (accusative) Case marker *o* or by the postposition *ni*.

(25) a. John-ga aruku.
John-NOM walk
'John walks.'
b. Mary-ga [John-ni/-o aruk]-aseta.
Mary-NOM John-to/-ACC walk-made
'Mary made John walk.'

However, when the "embedded verb" is transitive and has its own object NP, the causee cannot appear in objective Case: (26b) with *John-o* is totally ungrammatical.

(26) a. John-ga hon-o yomu.
John-NOM book-ACC read
'John reads a book.'
b. Mary-ga [John-ni/*-o hon-o yom]-aseta.
Mary-NOM John-to/-ACC book-ACC read-made
'Mary made John read a book.'

The other kind of violation is called the *surface* Double-*o* Constraint violation. Roughly put, the marker *o* in Japanese is ambiguous between the objective Case marker and an adverbial locative postposition. The latter type of *o* appears in examples like (27).

(27) John-ga hamabe-o aruku.
John-NOM beach-on walk
'John walks on the beach.'

When this type of *o* coocuurs with the objective Case marker *o*, the sentence instantiates the surface Double-*o* Constraint violation, as in (28).

(28) Mary-ga [John-ni/??-o hamabe-o aruk]-aseta.
Mary-NOM John-to/-ACC beach-on walk-made
'Mary made John walk on the beach.'

This type of violation results only in marginality.

As discussed in detail in the references cited above, there is an important respect in which the abstract and surface Double-*o* Constraint violations differ, aside from their degrees of deviance. The abstract case obtains even when one of the "*o*-marked NPs" is an empty category, whereas the surface case can be circumvented by dislocating one of the NPs. The following examples of the cleft construction illustrate this point: 8

(29) a. *[CP Opi [IP Mary-ga John-o *t*i yom-aseta] no]-wa
Mary-NOM John-ACC read-made COMP-TOP
hon-oi da.
book-ACC is
'It is a book that Mary made John read.'
b. *[CP Opi [IP Mary-ga *t*i hon-o yom-aseta] no]-wa
Mary-NOM book-ACC read-made COMP-TOP
John-oi da.
John-ACC is
'It is John that Mary made read a book.'

(30) a. [CP Opi [IP Mary-ga John-o *t*i aruk-aseta] no]-wa
Mary-NOM John-ACC walk-made COMP-TOP
hamabe-oi da.
beach-on is
'It is the beach where Mary made John walk.'
b. [CP Opi [IP Mary-ga *t*i hamebe-o aruk-aseta] no]-wa
Mary-NOM beach-on walk-made COMP-TOP
John-oi da.
John-ACC is
'It is John who Mary made walk on the beach.'

The examples in (29) have two NPs that require abstract objective Case, and they are completely unacceptable even when one of them is a trace produced by empty operator movement. On the other hand, the examples in (30) show that an accusative NP and a locative NP with *o* are compatible as long as one of them is dislocated.

Given this background, let us consider the status of the light verb examples in (24) in more detail. These examples are only marginal and clearly are not as bad as (26b) with the *o*-marked causee. From this, Sells (1988) concludes that they are merely instances of the weak surface Double-*o* Constraint violation. It is predicted, then, that when the theme NP in the cleft construction is dislocated, the Double-*o* Constraint effect should disappear. The prediction is indeed borne out, as shown in (31).

(31) a. [CP Opi [IP Honda-ga Amerika-de *t*i seisan-o site-iru]
Honda-NOM U.S.-in production-ACC doing-is
no]-wa Akoodo-oi da. (cf. (24a))
COMP-TOP Accord-ACC is
'It is Accords that Honda is producing in the U.S.'
b. [CP Opi [IP Mary-ga John-ni/-e *t*i zyooto-o sita] no]-wa
Mary-NOM John-to/-to giving-ACC did COMP-TOP
toti-oi da. (cf. (24b))
land-ACC is
'It is a piece of land that Mary gave to John.'

The grammaticality of these examples provides strong confirmation for Sells's conclusion.

However, as Sells notes, this conclusion raises an interesting question for the analysis of the light verb construction. That is, why is it that

(24a,b) are not abstract Double-*o* Constraint violations? 9 Clearly, neither of the two *o*-marked NPs is a locative adverbial. Yet it seems that only one of them needs to be licensed by abstract objective Case assignment (or checking); otherwise, we should expect an abstract violation. Here, we suggest that LF incorporation provides an answer. More specifically, we suggest that the accusative Case on the q-role-assigning noun is licensed by its incorporation to the verb, basically along the lines proposed by Baker (1988) for other cases of noun incorporation. Then, only the accusative Case on the theme NP needs to be licensed by abstract Case assignment (or checking). Thus, we correctly predict that (24a,b) are only surface Double-*o* Constraint violations.10

Examples of cleft sentences such as those in (32) may seem problematic, but on the contrary, they constitute further evidence for the LF incorporation analysis.

(32) a. *[CP Opi [IP Honda-ga Amerika-de akoodo-o *t*i site-iru]
Honda-NOM U.S.-in Accord-ACC doing-is
no]-wa seisan-oi da. (cf. (31a))
COMP-TOP production-ACC is
b. *[CP Opi [IP Mary-ga John-ni/-e toti-o *t*i sita] no]-wa
Mary-NOM John-to/-to land-ACC did COMP-TOP
zyooto-oi da. (cf. (31b))
giving-ACC is

In (31) the theme NP is clefted, and the Double-*o* Constraint violation is circumvented, as expected. The examples in (32), on the other hand, show that when the accusative NP headed by the q-role-assigning noun is clefted, the result is completely unacceptable. This seems surprising, given that the weak Double-*o* Constraint effect disappears when one of the *o*-marked NPs is dislocated. But this is exactly as expected under the LF incorporation analysis. The noun *zyooto* 'giving' in (32b), for example, must incorporate into the verb *su* in LF. Otherwise, the subject NP *Mary-ga* and the goal PP *John-ni/-e* fail to receive their q-roles. But the required LF incorporation would clearly involve lowering, as the light verb *su* does not c-command the q-role-assigning noun *zyooto*. Consequently, this incorporation is excluded by the Proper Binding Condition or a general ban on lowering. The ungrammaticality of (32a,b) and the grammaticality of (31a,b) thus provide supporting evidence for the LF incorporation analysis.11

8.3
Some Consequences of the LF Incorporation Analysis

Although Grimshaw and Mester's (1988) argument transfer analysis has a number of attractive features, it also has some unsolved problems. One has to do with the status of the argument transfer operation itself. It is possible that an operation of this kind is allowed by Universal Grammar. But to the extent that it is employed only to account for the light verb construction, it is a description of a fact rather than an explanation. Thus, it seems desirable to eliminate it if possible. The LF incorporation analysis, as an alternative, clearly achieves this goal. In this section we show that this analysis also solves other potential problems that arise with the argument transfer analysis.

Recall first Grimshaw and Mester's important generalizations in (7), repeated in (33).

(33) a. At least one internal q-role of the noun must be assigned to an argument outside the NP.
b. If a q-role T is assigned outside the NP, then all q-roles that are higher than T in the thematic hierarchy must also be assigned outside the NP.

As briefly discussed above, their account for (33b) is that argument transfer applies in an outside-in fashion. The examples in (34) are ruled out because an inner argument of the q-role-assigning noun is transferred to *su*, leaving behind an outer argument, as illustrated in (35).

(34) a. *John-ni/-e [NP Mary-no toti-no zyooto]-o sita. (= (17))
John-to/-to Mary-GEN land-GEN giving-ACC did
'Mary gave a piece of land to John.'
b. *Mary-ga toti-o [NP John-e-no zyooto]-o sita. (= (12a))
Mary-NOM land-ACC John-to-GEN giving-ACC did.
(same)

(35) a. *zyooto* (agent (goal (theme))), *su* ( ) <ACC> . . . input
b. *zyooto* (agent (theme)) + *su* (goal) <ACC> . . . (34a)
c. *zyooto* (goal) + *su* (agent (theme)) <ACC> . . . (34b)

In (35b) the goal q-role is transferred to *su* "across" the agent q-role, and in (35c) the theme q-role is transferred, leaving behind the outer goal q-role.

Even if one accepts the hypothesis that argument transfer applies in a particular fashion (e.g., outside-in), one might wonder why this should be the case. As Grimshaw and Mester note, the effect of outside-in applica-

tion is that the basic hierarchy among the arguments is preserved even after argument transfer applies. But argument transfer makes the q-role-assigning noun and the light verb *su* independent q-role assigners. Thus, in (35b) *zyooto* 'giving' has two q-roles (agent and theme), whereas *su* has one (goal). Then, it is not clear why the thematic hierarchy must be observed, after argument transfer takes place, among the arguments of two independent q-role assigners. As the main point of the thematic hierarchy is to specify the order in which q-roles are discharged, it is natural to assume that it only affects the structural relation among the arguments of a single q-role assigner.

This question does not arise with the LF incorporation analysis, which in fact provides a much more straightforward explanation for the generalization in (33b). According to this analysis, there is only one q-role assigner, namely, the q-role-assigning noun. In the grammatical (36) the q-roles are discharged in the way specified by the thematic hierarchy.

(36) Mary-ga John-ni/-e [NP toti-no zyooto]-o sita. (= (15a))
Mary-NOM John-to/-to land-GEN giving-ACC did
'Mary gave a piece of land to John.'

The noun *zyooto* 'giving' discharges the theme q-role to *toti* 'land' within the NP. Then, in LF, it raises to the position of *su* and assigns the goal q-role to *John-ni/-e*. Finally, it discharges the agent q-role to *Mary-ga*, the argument in the highest structural position. On the other hand, in (34a,b) q-role assignment cannot take place consistently with the thematic hierarchy. In (34a) *zyooto* must assign the agent q-role within the NP before it raises in LF and discharges the goal q-role. Similarly, in (34b) it assigns the goal q-role within the NP before it discharges the theme q-role. Thus, the LF incorporation analysis enables us to explain the generalization in (33b) without any stipulation. The only assumption needed is that q-roles of a single predicate are discharged in a certain order, in accordance with the thematic hierarchy.

Another remaining problem for Grimshaw and Mester's (1988) analysis has to do with their account for (33a). As noted above, they hypothesize that the NP headed by the q-role-assigning noun is licensed by virtue of the transfer of an unsuppressed q-role. In the grammatical (36) the goal q-role is transferred to *su*, and thus the NP headed by *zyooto* 'giving' is licensed. But in (37) only the external agent q-role is transferred to *su*.

(37) ?Mary-ga [NP John-e-no toti-no zyooto]-o sita. (= (15b))
Mary-NOM John-to-GEN land-GEN giving-ACC did
'Mary gave a piece of land to John.'

Since the external q-role of a noun is suppressed, the NP headed by *zyooto* fails to be licensed. Thus, the example is unacceptable as an instance of the light verb construction.

The account illustrated above is straightforward if nouns do not have external q-roles at all. Then, no q-role is transferred in examples like (37). But as the external agent q-role is realized at the clausal level in the grammatical (36), it is clear that *zyooto* 'giving' has an external q-role and that the q-role can be transferred to *su*. Thus, as Grimshaw and Mester note, the external q-role of a noun must be visible to argument transfer even if it is suppressed. Then why doesn't the transfer of a suppressed q-role count for the licensing of the NP headed by the q-role-assigning noun? It seems difficult to come up with an answer as long as we assume the argument transfer analysis.

The LF incorporation analysis seems to provide a principled solution to this problem as well. Here, we would like to argue that the generalization in (33a) in fact can be derived as a consequence of the analysis if Chomsky's (1986b, 1993) principle Last Resort is assumed. Roughly put, Last Resort states that movement applies only when it is necessary to satisfy the morphological properties of the moved item. Thus, the movement in (38a) is possible, but that in (38b) is not.

(38) a. Maryi seems [*t*i to be intelligent].
b. *Johni seems to *t*i [that Mary is intelligent].

In (38a) *Mary* needs to move to [Spec, IP] to have its Case checked. In (38b), by contrast, *John* is already in a Case position before movement and there is no need for this NP to move to [Spec, IP]. Hence, Last Resort allows only the movement in (38a). The following contrast is explained in a similar way:

(39) a. There seems to be a man in the corner.
b. *There seems to a man [that Mary is intelligent].

The expletive *there* cannot be present at LF and hence must be replaced, as a result of Full Interpretation. In (39a) the "associate NP" *a man* needs to have its Case checked and hence must move to [Spec, IP] in LF. Thus, *there* is successfully replaced. In (39b), however, a *man* is already in a Case position, and there is no reason for this NP to move to [Spec, IP]. In this case Last Resort blocks the movement. As a result, the expletive remains at LF in violation of Full Interpretation. This example clearly shows that movement cannot take place to save the overall structure, but only to fulfill a need of the moved item.

Given Last Resort, we may ask how the LF incorporation in the light verb construction is motivated. We hypothesized above that in examples like (24b), repeated in (40), the accusative Case on *zyooto* 'giving' is licensed by virtue of the LF incorporation.

(40) ??Mary-ga John-ni/-e toti-o [NP zyooto]-o sita.
Mary-NOM John-to/-to land-ACC giving-ACC did
'Mary gave a piece of land to John.'

Thus, it seems possible that the incorporation takes place so that the accusative Case on the q-role-assigning noun can be checked. However, this hypothesis immediately faces a problem. If LF incorporation can take place for Case reasons, then we lose the account for the abstract Double-*o* Constraint violations. Let us again consider (26b), repeated in (41).

(41) *Mary-ga John-o hon-o yom-aseta.
Mary-NOM John-ACC book-ACC read-made
'Mary made John read a book.'

This example is totally impossible since there are two NPs that must be licensed by abstract objective Case assignment (or checking). We suggested above that (40) does not show the same violation because one of the accusative Cases (i.e., the one on *zyooto* 'giving') is licensed by incorporation. But if Case licensing can motivate LF incorporation, there seems to be no way to prevent the accusative Case on *hon* 'book' in (41) from being licensed in the same way. That is, the noun *hon* should be able to incorporate into the verb in LF and have its Case licensed exactly like *zyooto* in (40). We then incorrectly predict that (41) has the same grammatical status as (40). Hence, Case licensing cannot be a "proper motivation" for LF incorporation (as opposed to NP-movement to [Spec, IP]).

When we compare *zyooto* 'giving' in (40) and *hon* 'book' in (41), an obvious alternative emerges. The former noun, as opposed to the latter, is a q-role assigner and hence must discharge its q-roles. Then, it is quite plausible that q-role assignment motivates the LF incorporation. The noun *zyooto* raises to discharge its theme, goal, and agent q-roles in (40), and to assign its goal and agent q-roles in the grammatical (36), repeated in (42).

(42) Mary-ga John-ni/-e [NP toti-no zyooto]-o sita.
Mary-NOM John-to/-to land-GEN giving-ACC did
'Mary gave a piece of land to John.'

In (41), on the other hand, there is no reason for the noun *hon* to raise to the position of the verb, since the noun is not a q-role assigner. Consequently, Last Resort prohibits the incorporation in this case, and as a result the accusative Case on the noun cannot be licensed by incorporation. 12

Let us now return to the problematic (37), repeated in (43), which is unacceptable with the light verb interpretation of *su*.

(43) ?Mary-ga [NP John-e-no toti-no zyooto]-o sita.
Mary-NOM John-to-GEN land-GEN giving-ACC did
'Mary gave a piece of land to John.'

Here, as in Grimshaw and Mester's analysis, we can take advantage of the fact that the external q-role of a noun is suppressed, but in a more direct way. In fact, all we need to assume is that nouns need not discharge their external q-roles. Given this uncontroversial assumption, there is no reason for *zyooto* 'giving' in (43) to incorporate, since it discharges all of its internal q-roles within its own projection. Thus, given Last Resort, the incorporation cannot take place, and the subject NP *Mary-ga* fails to receive a q-role. The only possible way for *Mary-ga* to receive a q-role is from the verb *su*. Then, *su* cannot be a light verb, but must be a q-role assigner exactly like the English main verb *do*, a correct result.

As we attributed the ungrammaticality of (43) (with the light verb interpretation) to the failure of *Mary-ga* to receive a q-role, a question can be raised with respect to examples like (44), where no argument (overt NP or nonexpletive *pro*) appears in the clausal subject position.

(44) *[NP John-e-no toti-no zyooto]-o sita.
John-to-GEN land-GEN giving-ACC did
'(Someone) gave a piece of land to John.'

As there is no argument NP in the subject position, there should be no problem with q-marking even if *zyooto* 'giving' does not raise to the position of *su*. However, the example will be ruled out in exactly the same way as the English (39b). Since by assumption there is no argument NP in [Spec, IP], this position must be occupied by an expletive *pro*. And this expletive remains at LF in violation of Full Interpretation, as no NP can replace it. In particular, Last Resort prevents the NP headed by *zyooto* from moving to [Spec, IP], since its Case is licensed by *su*.

The analysis outlined above for (43) accounts for all cases of Grimshaw and Mester's generalization that at least one internal argument of the q-role-assigning noun must appear at the clausal level. If there is no internal

argument at the clausal level, Last Resort prevents the q-role-assigning noun from incorporating into *su*, and hence, the subject NP fails to receive a q-role. If the subject position is occupied by an expletive, which does not need a q-role, then the expletive ends up violating Full Interpretation at LF, again because of Last Resort. Thus, this analysis, if correct, provides strong support for the general idea that movement applies only as a last resort.

We have shown in this section that the LF incorporation account enables us to improve on Grimshaw and Mester's analysis. This constitutes further evidence for the LF incorporation analysis and consequently for the minimalist hypothesis that the q-Criterion applies only at LF. In a sense, the analysis proposed here provides a clearer case than Larson's (1988) analysis of the English dative/double object constructions, since it implies that q-roles need not be discharged even at S-Structure, as long as they are assigned at LF.

In addition, our analysis provides support for the major part of Grimshaw and Mester's hypothesis regarding argument structure. Exactly as in their analysis, we assumed crucially that nouns need not assign their external q-roles. More importantly, we relied on the specific version of the thematic hierarchy assumed by Grimshaw and Mester (e.g., agent > goal > theme). The latter assumption, interestingly, contradicts Larson's analysis, which implies the hierarchy agent > theme > goal. (See (22) above.) Thus, although our analysis directly supports Larson's general idea on the way in which q-roles are discharged, it indirectly supports alternative analyses of the dative/double object constructions (e.g., those argued for in Pesetsky 1995, Kitagawa 1994, and Takano 1996).

Our account for generalization (33a) in terms of Last Resort implies that q-roles constitute part of the "morphological properties" of a q-role-assigning head. That is, in the terminology of Chomsky 1995, they are "formal features" on the q-role-assigning heads. This is so because according to our analysis, q-role-assigning nouns are allowed to move to discharge (or check) their q-roles. Chomsky (1994, 1995) presents evidence that the failure of q-role assignment results in a nonconvergent derivation, not just semantic deviance. Lasnik (1995) suggests the stronger hypothesis that q-roles are formal features of the assigner. Boskovic * and Takahashi (1998) argue for a yet stronger conclusion that q-roles are formal features on both the assigner and the assignee. (See also Boskovic 1994.) Our analysis is consistent with these proposals and more specifically provides direct support for Lasnik's suggestion.13

8.4
Miyagawa's (1989) and Tsujimura's (1990a) Ergativity Constraint

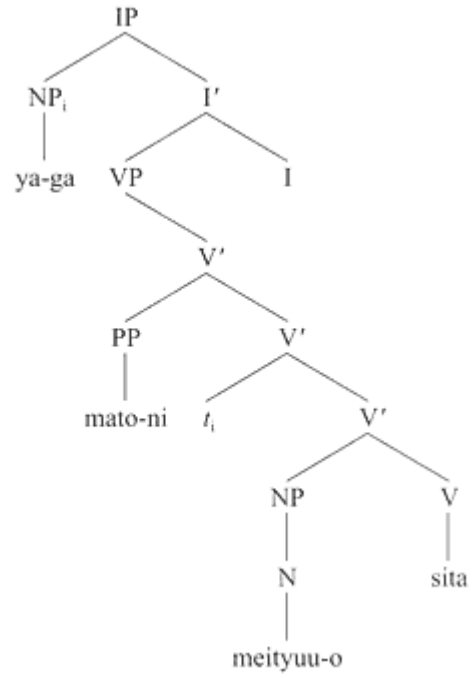
In this section we discuss an important fact noted by Miyagawa (1989) and Tsujimura (1990a), namely, that the q-role-assigning noun in the light verb construction cannot be ergative. Our purpose here is to show that the LF incorporation analysis suggests some promising directions toward explaining this mysterious fact.

A relevant contrast is shown in (45). 14

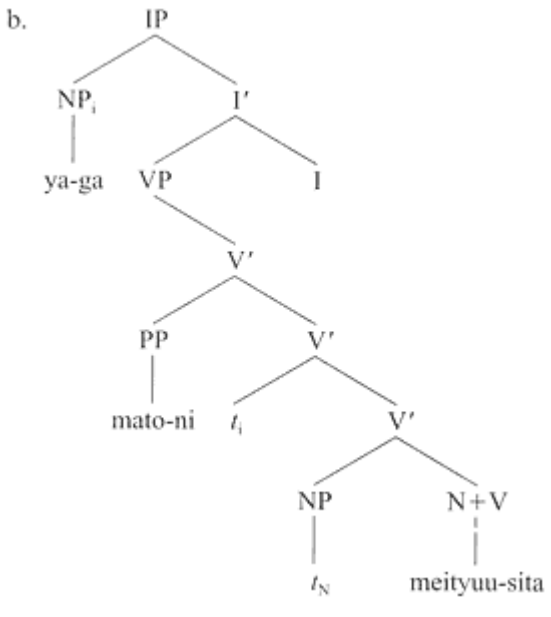
(45) a. ?*Ya-ga mato-ni [NP meityuu]-o sita.
arrow-NOM target-to strike-ACC did
'The arrow struck the target.'
b. Ya-ga mato-ni [V meityuu-sita].
arrow-NOM target-to strike-did
'The arrow struck the target.'

(45a) illustrates Miyagawa's and Tsujimura's generalization: *meityuu* 'strike' is ergative and the example is degraded. (45b), on the other hand, shows that an ergative noun is allowed as long as it is combined with the light verb in overt syntax. If we assume, following Kageyama (1982), Terada (1990), Tsujimura (1990b), and others, that this case involves syntactic incorporation, the structures of (45a,b) at S-Structure (i.e., at the point of Spell-Out) can be represented as in (46a,b), respectively.

(46) a.



b.



Miyagawa (1989) and Tsujimura (1990a) argue that this contrast is expected given Burzio's (1986) Generalization, which is stated in (47).

(47) A verb assigns an external q-role iff it can assign Case.

Assuming Grimshaw and Mester's (1988) analysis for (45a), Miyagawa and Tsujimura first point out that argument transfer creates the lexical entry in (48b) from those in (48a). 15

(48) a. *meityuu* (Æ (goal (theme))), *su* ( ) <ACC> . . . input
b. *meityuu* ( ) + *su* (Æ (goal (theme))) <ACC> . . . (45a)

But (48b) is inconsistent with Burzio's Generalization since *su* is an accusative Case assigner and yet has no external q-role. The N V complex, *meityuu-su* 'strike', in (45b) does not have this problem: it assigns neither an accusative Case nor an external q-role.

As far as we can see, this account for (45a) can be maintained with the LF incorporation analysis, provided that Burzio's Generalization holds at this level. Let us again consider the example in (40), repeated in (49).

(49) ??Mary-ga John-ni/-e toti-o [NP zyooto]-o sita.
Mary-NOM John-to/-to land-ACC giving-ACC did
'Mary gave a piece of land to John.'

This example indicates that after the LF incorporation of *zyooto* 'giving' to *su*, the N V complex can still assign accusative Case to (or check the accusative Case of) the theme NP *toti* 'land'. That is, the Case feature of *su* is retained by the N V complex. But the N V complex formed in the LF representation of (45a) does not have an external q-role, since *meityuu* 'strike' is an ergative noun. Hence, its properties are inconsistent with Burzio's Generalization.

On the basis of the contrast in (45), Miyagawa and Tsujimura argue convincingly for their conclusion that Japanese has ergative nouns. But the task of providing a precise account for examples like (45a) still remains. Those examples are particularly interesting because Chomsky's (1986b) explanation for the standard cases of Burzio's Generalization does not seem to cover them. The relevant part of the generalization is schematized in (50).

(50) *[$_{IP}$ NP [$_{VP}$ V NP]] (order irrelevant)

no θ-role Case

Chomsky (1986b) proposes to explain (50), to the extent that it is correct, by Last Resort. Since the object is in a Case position, this principle prevents it from moving to the subject position in overt syntax. (See (38b) above.) This implies that the subject position must be occupied by an expletive. But Last Resort prevents the movement of the object NP not only in overt syntax but also in LF. Hence, the expletive fails to be replaced and remains in LF in violation of Full Interpretation. (See (39b) above.)

This explanation for (50) does not seem to extend to examples like (45a). If nominative Case in Japanese must be licensed within the I projection at S-Structure (or by the point of Spell-Out), as argued by Saito (1982) and Takezawa (1987), and assumed here, then the NP *ya-ga* 'arrow-NOM' should be able to move from a VP-internal position to [Spec, IP] without violating Last Resort. The movement is illustrated in (46a). Here, there is no need for the accusative NP to move to [Spec, IP] either overtly or in LF. Hence, Chomsky's account for (50) in terms of Last Resort does not cover examples of this type. 16

Thus, Miyagawa's and Tsujimura's generalization regarding the light verb construction still constitutes a problem to be explained. More importantly, if Chomsky's explanation for (50) is correct, (45a) should not be grouped together with the standard cases of Burzio's Generalization. We must therefore look for an independent way to explain cases like (45a).

Although we do not have a precise proposal to offer at this point, we would like to show in the remainder of this section that the LF incorporation analysis allows us to approach this problem from a new perspective.

One difference clearly stands out between (45a) and (45b). The NP *ya-ga* 'arrow-NOM' moves overtly to [Spec, IP] in both (45a,b), as shown in (46a,b). But this NP-movement differs in the two examples in one important respect. Since the q-role-assigning noun adjoins to *su* overtly in (46b), *ya-ga* is already assigned a q-role when it moves. On the other hand, in (46a) the q-role-assigning noun incorporates only in LF. Thus, when *ya-ga* moves to [Spec, IP], it is not yet assigned a q-role. The contrast between (45a) and (45b) is then captured if the following generalization holds: 17

(51) An argument NP cannot move from a $\bar{\theta}$ (non-q) position to a Case position.

If we attribute the ungrammaticality of (45a) to (51), we are led to a specific analysis for an external argument subject. Let us again consider example (42), repeated in (52).

(52) Mary-ga John-ni/-e [NP toti-no zyooto]-o sita.
Mary-NOM John-to/-to land-GEN giving-ACC did
'Mary gave a piece of land to John.'

According to the standard VP-internal subject hypothesis, the subject *Mary-ga* in (52) is generated in [Spec, VP] and is raised overtly to [Spec, IP]. But since the NP receives a q-role only after the q-role-assigning noun *zyooto* 'giving' raises to the position of *su* in LF, the movement of the nominative NP originates in a $\bar{\theta}$ position exactly as in the case of (45a). Thus, (51) entails that the subject in (52), as opposed to that in (45a), can be generated directly in [Spec, IP].

It seems that this analysis of (52) in fact should be allowed if our proposals on the light verb construction are correct. That is, even if *Mary-ga* in (52) is generated directly in [Spec, IP], it should be able to successfully receive its q-role. The q-role-assigning noun *zyooto* 'giving' can first raise to the position of *su* and discharge its goal q-role to *John-ni/-e*, then raise to I and assign its agent q-role to *Mary-ga*. Hence, nothing seems to prevent the direct insertion of the subject NP in [Spec, IP].18

What we must ensure is that the nominative NP in (46a) cannot be generated directly in [Spec, IP] and receive a q-role in the same way. Otherwise, *ya-ga* 'arrow-NOM' in (46a) need not move from a $\bar{\theta}$ position to

a Case position, and the generalization in (51) becomes irrelevant for this case also. Here, according to Miyagawa and Tsujimura, the crucial difference between (46a) and (52) is that in the former the q-role-assigning noun is ergative. That means that only in (46a) is the nominative NP assigned an internal q-role. Thus, the following reasonable assumption will have the desired effect:

(53) An internal q-role is assigned within a projection of a lexical category.

This forces the nominative NP to originate internal to VP only when the q-role-assigning noun is ergative.

The approach to Miyagawa's and Tsujimura's generalization outlined above is promising if (51) has a plausible interpretation in the overall syntactic theory. We would like to suggest two possibile interpretations of this generalization without commiting ourselves to either at this point.

First, it seems possible to group (51) with the general ban on improper movement from an $\bar{A}$ position to a Case position. There are examples that clearly show that movement of an adjunct to a Case position is prohibited. Thus, (54) contrasts sharply with (55).

(54) Yesterdayi seemed [IP *t*i to be the best day to hold the meeting].

(55) *Yesterdayi seemed *t*i [CP that Mary was the leading candidate].
(cf. It seemed yesterday that Mary was the leading candidate.)

(54) shows that *yesterday* can move to [Spec, IP] if it originates in an A position. (55), on the other hand, shows that it cannot move from an adjunct position to [Spec, IP]. 19

The standard cases of improper movement also fall under this generalization. Examples such as the following have been discussed by Takahashi (1992) and Sakai (1994), among others:

(56) *[IP Johni seems [CP *t*i' [IP it is likely [IP *t*i to win the race]]]].

It is not clear what condition on representation this example violates, especially if the intermediate trace can delete in LF. (See also Fukui 1993 for relevant discussion.) What appears to be wrong in (56) is the second step of the movement: it takes place from an $\bar{A}$ position ([Spec, CP]) to a Case position (matrix [Spec, IP]). Thus, examples like (56) also suggest that movement from an $\bar{A}$ position to a Case position cannot take place.

Let us now return to the generalization in (51) and example (45a), which is repeated in (57) with a more precise structure.

(57) ?*[IP Ya-gai [VP mato-ni *t*i [NP meityuu]-o sita]].
arrow-NOM target-to strike-ACC did
'The arrow struck the target.'

Suppose, as seems reasonable, that a VP-internal position counts as an Ā position in the relevant sense, unless it is a q position. Then, when *ya-ga* 'arrow-NOM' moves overtly from the VP-internal position to [Spec, IP], it originates in an Ā position. Hence, the movement takes place from an Ā position to a Case position, exactly as in (55) and (56). Viewed this way, (51) will be a subcase of a more general, well-established generalization.

The second, totally independent way to interpret (51) is in terms of q-theory. Recall first that in (46b), the representation of the grammatical (45b), *ya-ga* 'arrow-NOM' moves to [Spec, IP] after it is assigned a q-role. But in (46a), the structure of (45a), the same NP is not yet assigned a q-role when it moves to [Spec, IP]. Suppose then that a q-role can be assigned only to the head of a chain. In (46b) *ya-ga* heads its singleton chain when it receives a q-role, and it carries this q-role along to [Spec, IP]. On the other hand, in (46a), by the time the q-role-assigning noun *meityuu* 'strike' incorporates into the light verb *su* in LF, *ya-ga* is already in [Spec, IP], heading a two-member chain. If only the head of a chain can receive a q-role, *meityuu* is unable to assign a q-role to *ya-ga*.

The success of this second approach, in turn, relies on whether there is a principled explanation for the condition that only the head of a chain can receive a q-role. We believe that there is. Note first that it is possible to generalize this condition to (58).

(58) An operation on a chain must take place at its head position.

This would make sense if we adopt Chomsky's (1993) copy theory of movement in a specific way. Chomsky proposes that movement consists of copying and deletion. Thus, movement leaves a copy of the moved item behind. Let us assume that a chain consists literally of the head and its copies. It is then possible that an operation on the head also automatically affects the other members of the chain, which are by definition copies of the head. The chain remains uniform, consisting of the head and its copies. On the other hand, an operation on a copy is illegitimate, since it does not affect the "original." The head of the chain is not a copy of the element to which the operation applies. Hence, the operation has no effect on the head, and as a result, it creates a nonuniform chain where the copy is no longer identical to the head.

There is also another explanation for (58), which seems to us more plausible. That is, (58) follows if A-movement does not leave a trace (or a copy), and hence, there are no A-chains. Then, *ya-ga* 'arrow-NOM' in (46a) fails to receive a q-role at LF through its trace simply because the trace does not exist. Although the elimination of A-traces and A-chains appears to be a radical move, Lasnik and Saito (1992), for example, come very close to proposing just that.

Lasnik and Saito first propose a revision of Chomsky's (1986b) Uniformity Condition as in (59).

(59) a. a assigns inherent Case to b only if a q-marks b.
b. Suppose that b bears a q-role assigned by a. Then, if g is a barrier for a, g dominates b.

(59a) is identical to the first part of Chomsky's condition. (59b), on the other hand, is a generalized version of the second part. Unlike Chomsky's original formulation, it extends to examples such as (60), attributed to Mark Baker, and (61).

(60) *Johni seems [CP that [IP it is told *t*i [CP that [IP Mary is intelligent]]]].

(61) *Johni seems [CP[IP[NP the belief [IP *t*i to be intelligent]] is crazy]].

(59b) amounts to saying that A-movement cannot take place across a barrier. Hence, it straightforwardly rules out (60) and (61), where *John* is moved across a CP/IP pair. For A-movement, it has the same empirical coverage as Chomsky's (1986a) proposal that all traces must be antecedent-governed. One thing that it is designed to explain is Aoun's (1982) generalization that "S' (CP) breaks A-chains."

Having motivated (59b), Lasnik and Saito (1992) point out that there are considerable overlaps among the effects of (59b), Condition A of the binding theory, the Empty Category Principle (ECP) (in their formulation), and the Locality Condition on Chains. Then, they show that (59b) in fact suffices to rule out most, if not all, of the relevant ungrammatical examples. Let us consider (62) (64) to illustrate this point.

(62) a. *Johni seems [CP[IP Mary visited *t*i]].
b. *Johni seems [CP[IP *t*i is intelligent]].

(63) *Johni seems [CP that [IP it is believed [IP *t*i to be intelligent]]].

(64) *Johni seems [CP[IP[NP hisi belief [IP *t*i to be intelligent]] is crazy]].

(62a) and (62b) instantiate the Specified Subject Condition case and the Nominative Island Condition case of Condition A, respectively. (63) is an example of superraising, which Lasnik and Saito (1984) propose to rule out by the ECP. And Lasnik (1985) appeals to the Locality Condition on Chains to account for examples such as (64). Examples (62) (64) are all excluded by (59b): in all cases the movement crosses a CP/IP pair. On the basis of examples like these, Lasnik and Saito (1992) suggest that A-traces are subject neither to Condition A nor to the ECP, and further that the Locality Condition on Chains can be dispensed with. If this is tenable, then A-traces and A-chains may in fact be redundant. Unlike Condition A and the ECP, (59) does not require the presence of A-traces. And unlike the Locality Condition on Chains, it does not refer to A-chains. 20

The hypothesis that there are no A-traces goes well with our conception of q-role assignment. If q-roles are significant only in semantic interpretation and if thematic relations are simply read off the LF representations, A-traces and A-chains are clearly needed. However, as noted above, our analysis of the light verb construction implies that q-roles are formal features of q-role-assigning heads. Then, it is plausible that they can be assigned (or discharged) in the course of the derivation, and the configuration for q-role assignment need not be preserved at LF. This in turn implies that A-traces are not necessary to represent thematic relations at this level.

We have shown so far that there are two plausible ways to approach Miyagawa's (1989) and Tsujimura's (1990a) ergativity constraint on the light verb construction. Both are based on the assumption that in the ungrammatical (46a) the nominative NP raises to [Spec, IP] before it is assigned a q-role. This assumption does not hold under Grimshaw and Mester's (1988) argument transfer analysis. In this analysis argument transfer takes place at D-Structure; hence, the nominative NP in (46a) is assigned a q-role before it moves. On the other hand, the assumption is an automatic consequence of the LF incorporation analysis. Thus, to the extent that they are tenable, the approaches to the ergativity constraint suggested in this section provide further support for the LF incorporation analysis.

8.5 Conclusion

In this chapter we argued for an LF incorporation analysis of the light verb construction in Japanese. In section 8.2 we motivated the analysis on

the basis of the pattern shown by double-*o* examples. In section 8.3 we argued that the analysis enables us to capture Grimshaw and Mester's (1988) insights in a principled way. Finally, in section 8.4 we showed that it suggests promising approaches to Miyagawa's (1989) and Tsujimura's (1990a) ergativity constraint on the construction.

The LF incorporation analysis implies that q-roles need not be assigned at D-Structure or S-Structure, but can be discharged at LF. Thus, it provides strong support for Chomsky's (1993) Minimalist Program, which proposes to eliminate D-Structure as a pure representation of GF-q relations, together with the Projection Principle. We also argued that LF incorporation is subject to Chomsky's (1986b, 1993) Last Resort, which was initially motivated for movement of maximal projections. Our analysis thus broadens the empirical scope of this principle. The status of this principle has been debated extensively in the recent literature. Boskovic * (1997), for example, presents arguments for it, whereas Chomsky (1995) proposes to eliminate it, deriving its effects from the theory of formal features and the definition of the operation Attract. (See note 20 above.) Our analysis, as it stands, supports the former view, although it may be possible to reformulate it and make it consistent with Chomsky's (1995) proposals.

Our analysis implies also that q-roles are part of the "morphological properties" of their assigners: it is crucial in our analysis that a q-role-assigning noun can move to discharge its q-roles without violating Last Resort. As noted above, this conclusion accords with Lasnik's (1995) suggestion that q-roles are formal features of their assigners.

Our discussion of Miyagawa's and Tsujimura's generalization in section 8.4 was more speculative. We suggested two possibilities for explaining it. One was to reduce it to the general ban on improper movement from an $\bar{A}$ position to a Case position. The other was based on the elimination of A-traces and A-chains. Further research is needed before we can choose between these two possibilities.

Notes

This is a revised version of a paper originally written in 1994. A still earlier version, Hoshi and Saito 1993, was written as a chapter of the research report of the 1991 93 project on Japanese grammar at the University of Connecticut, submitted to BBN Systems and Technologies. We thank BBN for its financial support, and the other participants in the project, Jun Abe and Yoichi Miyamoto, for their suggestions and comments. We also benefited from discussions with many other people, including Zeljko Boskovic*, Howard Lasnik, Hideki Maki, Keiko

Murasugi, Myung-Kwan Park, and Daiko Takahashi. The early version was presented in the spring 1993 syntax seminar at the University of Connecticut, and in colloquia at Cornell University (1994) and Universität zu Köln (1995). This revised version, in its initial form, was given at the Special GLOW Meeting in Hyderabad (January 1998), and also in the spring 1998 syntax seminar at Nanzan University. We thank the audiences at these places for comments, and also Chris Collins, Joseph Emonds, Jay Jayaseelan, Roger Martin, and an anonymous reviewer for suggestions leading up to the final revision. The preparation of this revised version was supported in part by the Nanzan University Pache Research Grant (Saito) and by grants from the British Academy, the Japan Foundation Endowment Committee, and SOAS at the University of London (Hoshi).

We are delighted that this final version is included in a volume to honor Howard Lasnik, a great teacher and good friend to whom both of us owe so much.

1. See, for example, Grimshaw and Mester 1988, Sells 1988, Miyagawa 1989, Terada 1990, Tsujimura 1990a,b, Aoyagi 1991, Hasegawa 1991, Kajihara 1992, Matsumoto 1992, 1996, Kageyama 1993, Uchida and Nakayama 1993, Sato 1993, Nakajima 1993, and Dubinsky 1994.

2. As shown in (i), not only NPs but also PPs are followed by the genitive Case marker within Japanese NPs.

(i) a. [NP Mary-no John-to-no kekkon]
Mary-GEN John-with-GEN marriage
'Mary's marriage with John'
b. [NP Mary-no John-to-no kaiwa]
Mary-GEN John-with-GEN conversation
'Mary's conversation with John'

These examples are ungrammatical without the genitive markers.

3. See, for example, Kageyama 1982, 1993, Larson 1988, Terada 1990, Tsujimura 1990b, Koizumi 1995, and especially Hale and Keyser 1993 for relevant discussion on incorporation and q-role assignment. An LF incorporation analysis, somewhat similar to ours, is proposed independently in Dubinsky 1994. Since Dubinsky's account for the relevant data differs from ours in many respects, he draws different theoretical conclusions. But he suggests, as we do, that the q-role-assigning noun discharges its q-roles after its incorporation at LF.

4. The light verb *su* does not have any q-role of its own, but is a Case assigner: it assigns accusative Case to the NP headed by the q-role-assigning noun.

5. See, for example, Shibatani 1973, Harada 1973, Kuroda 1978, and Poser 1981 for detailed discussion on this constraint. We will consider the exact nature of the marginality of the examples in (9) in section 8.2.

6. A clausal goal argument can be realized either with the postposition *ni* or with the postposition *e*. However, only the latter is available NP-internally, since *ni* is for some reason incompatible with the genitive Case marker, as shown in (i).

(i) a. [IP Mary-ga Nara-ni/Nara-e syuppatu-sita].
Mary-NOM Nara-to/Nara-to departed
'Mary departed for Nara.'

b. [NP Mary-no Nara-e-no /*Nara-ni-no syuppatu]
Mary-GEN Nara-to-GEN/*Nara-to-GEN departure
'Mary's departure for Nara'

This is why we use only *e* in (12a) and (12c): those examples with *ni* are ruled out for an independent reason.

7. See Chomsky 1981 for detailed discussion of this assumption.

8. See Kuroda 1978, Hoji 1990, and Murasugi 1991 for detailed discussions on the cleft construction in Japanese. We follow the latter two works and assume that the examples in (29) (30) involve movement of an empty operator to [Spec, CP].

9. It should be noted here that Sells (1988) suggests an answer quite different from ours. See Uchida and Nakayama 1993 and Dubinsky 1994 for discussion of Sells's analysis.

10. The exact mechanism of "Case licensing by incorporation" still needs to be worked out. The cases of noun incorporation discussed in Baker 1988 involve movement of the N head of an argument NP at S-Structure. Since our case has to do with incorporation of a q-role-assigning noun in LF, it is not clear that Baker's mechanism can be applied directly.

One possibility is that in the configuration in (i) the lower segment V2 checks the Case of the incorporated N and the two-segment category V1 V2 checks the objective Case of the theme NP.

(i) V_1
N V_2

This seems consistent with Chomsky's (1994) proposal that the lower segment and the two-segment category are independent terms in an adjunction structure. If this is the correct way to analyze "Case licensing by incorporation," then the two accusative NPs in (24a,b) are after all both licensed by abstract Case checking. It is just that incorporation creates a configuration in which a single verb can check two objective Cases.

11. Grimshaw and Mester (1988) note that the NP headed by the q-role-assigning noun resists relativization and topicalization. As shown in (i), it cannot be passivized either. (See Sells 1988 for relevant discussion.)

(i) *Kekkon-ga Mary-ni yotte John-to sareta. (= *su* + *rare* (passive) + *ta*)
marriage-NOM Mary-by John-with done
'Mary married with John.'

As far as we can see, our account of (32) generalizes to all of these cases.

12. This implies that the Case licensing of *zyooto* 'giving' in (40) is "accidental." The noun incorporates into the light verb in LF only to discharge its q-roles, but as a result of this incorporation, the accusative Case on the noun happens to be licensed.

13. As an anonymous reviewer points out, if q-role assignment involves feature checking, interesting questions arise with respect to its locality. It is widely assumed

that q-role assignment is possible in the head-complement configuration. On the other hand, the complement position is excluded from the checking domain of the head (e.g., in Chomsky 1993). (See also Chomsky 1995 for much relevant discussion.) We may take this as evidence that the definition of checking domain should be revised. Alternatively, it may be possible to exclude the head-complement structure from the configurations of q-role assignment. Unfortunately, the examination of these possibilities is beyond the scope of this chapter.

14. Miyagawa and Tsujimura show that other ergative nouns, such as *zyoohatu* 'evaporation' and *tootyaku* 'arrival', exhibit the same pattern.

15. The absence of an external q-role is indicated by f.

16. Even if nominative is licensed VP-internally, as argued in Fukui 1986, the problem is not resolved. This hypothesis implies that the Extended Projection Principle (EPP) does not hold and there are no expletives in Japanese. This is so because if *ya-ga* 'arrow-NOM' in (45b) can have its Case licensed within VP and [Spec, IP] is occupied by an expletive *pro*, the example is incorrectly ruled out by Last Resort/Full Interpretation. In fact, given the VP-internal subject hypothesis, a simple sentence like (i) would lead to the same conclusion.

(i) John-ga hon-o yonda.
John-NOM book-ACC read
'John read a book.'

But if Japanese is not subject to the EPP and does not have an expletive, there is no expletive *pro* in the subject position of (45a). Consequently, Chomsky's Last Resort account for (50) does not extend to this example.

17. (51) also subsumes the nonpassivizability of the NP headed by the q-role-assigning noun in the light verb construction. See note 11.

18. The specifics of the head movement in (52) depend on how V and I are merged in Japanese. If V raises to I covertly, as proposed in Park 1991 and Saito 1992, then the q-role-assigning noun incorporates to V and the N V complex raises to I in LF. On the other hand, if V-to-I raising is overt in Japanese, as argued by Tada (1990), Hoshi (1994), Koizumi (1995), and Takahashi (1996), among others, then the q-role-assigning noun moves to the position of V and then incorporates into the V-I complex in LF.

Note that the account suggested here does not necessarily deny the VP-internal subject hypothesis for all cases. First, it seems needed for cases like (ia), where the main verb *read* does not raise to the position of I because of the presence of the auxiliary verb *has*.

(i) a. John has read the book.
b. John read the book.

If the main verb in a simple sentence like (ib) raises to I in LF, as argued by Chomsky (1991, 1993) and more recently by LaPorte-Grimes (1996), the subject in this sentence can be generated directly in [Spec, IP]. But even in this case it is possible that the subject can be generated in either [Spec, IP] or [Spec, VP], unless the latter possibility is excluded by an economy principle.

19. See Saito and Murasugi 1990 for facts similar to (55) in Japanese.

We use the terms *argument* and *A position* in a rather strict sense. Tada (1990), for example, discusses Japanese scrambling and concludes that the VP-adjoined position is an A position in a broader sense. Similarly, Murasugi (1991, 1992) argues on the basis of the distribution of *pro* in Japanese that locative and temporal phrases are arguments, again, in a broader sense. (See also Miyamoto 1994 for relevant discussion.)

20. Two remarks are in order here.

First, the conclusion reached here does not necessarily hinge on the validity of the condition in (59b). Note first that (62a,b) violate Last Resort, and (60), (62a), (63), and (64) are ruled out by Rizzi's (1990) Relativized Minimality. Chomsky (1995) proposes to derive the effects of these two principles by refining the theory of formal features and reinterpreting movement as "attraction." His Attract F is defined as in (i).

(i) K *attracts* F if F is the closest feature that can enter into a checking relation with a sublabel of K.

As far as we can tell, this can accommodate (61) as well.

If we accept this proposal, (59b) itself should be dispensed with in favor of (i). But our conclusion is unaffected since (i), like (59b), does not require A-traces or A-chains. We appeal to (59b), not (i), in the discussion in the text, since it is not clear to us at this point how the analysis proposed in section 8.3 in terms of Last Resort should be restated with (i).

Second, Lasnik and Saito (1992) do not go on to propose the elimination of A-traces because of examples like (i).

(i) *[How likely *t*i to be taken of John]j is advantagei *t*j?

Following a suggestion by Anthony Kroch (personal communication), they argue that examples of this kind are ruled out by the Proper Binding Condition as applied to A-traces. (See also Saito 1989.) If A-movement does not produce a trace, this case requires an independent explanation.

Another relevant case, which was brought to our attention by Chris Collins, is Huang's (1993) analysis of the contrast between (iia) and (iib).

(ii) a. [Which picture of himselfi/j]k does Johni think Billj likes *t*k?
b. [How *t*j proud of himself*i/j]k does Johni think Billj will be *t*k?

The reflexive in (iib) cannot take the matrix subject *John* as its antecedent. Huang proposes that the trace *t*j of *Bill* in the AP-internal subject position blocks this binding relation as a "specified subject." This analysis is incompatible with the Proper Binding Condition analysis of (i) and hence cannot be taken as *additional* evidence for A-traces. Yet, if A-traces are dispensed with, an independent account is needed for this case also.

References

Aoun, Joseph. 1982. The formal nature of anaphoric relations. Doctoral dissertation, MIT, Cambridge, Mass.

Aoyagi, Hiroshi. 1991. Light verb constructions and restructuring at LF. MS., University of Southern California, Los Angeles, and Nanzan University.

Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Boskovic, Zeljko. * 1994. D-Structure, Theta-Criterion, and movement into theta-positions. *Linguistic Analysis* 24, 247 286.

Boskovic, Zeljko. 1997. *The syntax of nonfinite complementation*. Cambridge, Mass.: MIT Press.

Boskovic, Zeljko, and Daiko Takahashi. 1998. Scrambling and Last Resort. *Linguistic Inquiry* 29, 347 366.

Burzio, Luigi. 1986. *Italian syntax: A government-binding approach*. Dordrecht: Kluwer.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1986a. *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1986b. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freidin, ed., *Principles and parameters in comparative grammar*. Cambridge, Mass.: MIT Press. [Also in Noam Chomsky, *The Minimalist Program*. Cambridge, Mass.: MIT Press (1995).]

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale and Samuel Jay Keyser, eds., *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press. [Also in Noam Chomsky, *The Minimalist Program*. Cambridge, Mass.: MIT Press (1995).]

Chomsky, Noam. 1994. Bare phrase structure. (MIT Occasional Papers in Linguistics 5.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass. [Also in Gert Webelhuth, ed., *Government and Binding Theory and the Minimalist Program*. Oxford: Blackwell (1995).]

Chomsky, Noam. 1995. Categories and transformations. In *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Dubinsky, Stanley. 1994. Syntactic underspecification: A minimalist approach to light verbs. In Hiroyuki Ura and Masatoshi Koizumi, eds., *Formal approaches to Japanese Linguistics 1*. (MIT Working Papers in Linguistics 24.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Fukui, Naoki. 1986. A theory of category projection and its applications. Doctoral dissertation, MIT, Cambridge, Mass.

Fukui, Naoki. 1993. A note on improper movement. *The Linguistic Review* 10, 111 126.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, Mass.: MIT Press.

Grimshaw, Jane, and Armin Mester. 1988. Light verbs and q-marking. *Linguistic Inquiry* 19, 205 232.

Hale, Kenneth. 1983. Warlpiri and the grammar of nonconfigurational languages. *Natural Language & Linguistic Theory* 1, 5 47.

Hale, Kenneth, and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Kenneth Hale and Samuel Jay Keyser, eds., *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press.

Harada, Shin-Ichi. 1973. Counter equi NP deletion. *Annual Bulletin* 7, 113 147. The Research Institute of Logopedics and Phoniatrics, University of Tokyo.

Hasegawa, Nobuko. 1991. On head movement in Japanese: The case of verbal nouns. In Sadayoshi Ogawa and Masahiro Akiyama, eds., *Proceedings of the Sophia Linguistic Society 6*. Department of Linguistics, Sophia University.

Hoji, Hajime. 1990. Theories of anaphora and aspects of Japanese syntax. Ms., University of Southern California, Los Angeles.

Hoshi, Hiroto. 1994. Passive, causative, and light verbs: A study on theta role assignment. Doctoral dissertation, University of Connecticut, Storrs.

Hoshi, Hiroto, and Mamoru Saito. 1993. Japanese light verb construction: A case of LF q-role assignment. In Mamoru Saito, ed., *A formal grammar of Japanese II*. Department of Linguistics, University of Connecticut, Storrs.

Huang, C.-T. James. 1993. Reconstruction and the structure of VP: Some theoretical consequences. *Linguistic Inquiry* 24, 103 138.

Kageyama, Taro. 1982. Word formation in Japanese. *Lingua* 57, 215 258.

Kageyama, Taro. 1993. *Bunpoo to gokeisei* (Grammar and word formation). Kasukabe, Japan: Hituzi Syobo.

Kajihara, Satomi. 1992. The syntactic nature of argument transfer with *suru*. Master's thesis, University of Washington, Seattle.

Kitagawa, Yoshihisa. 1994. Shells, yolks, and scrambled e.g.s. In Mercè Gonzalez, ed., *NELS 24*, vol. 1. GLSA, University of Massachusetts, Amherst.

Koizumi, Masatoshi. 1995. Phrase structure in minimalist syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Kuroda, S.-Y. 1978. Case-marking, canonical sentence patterns, and counter equi in Japanese. In John Hinds and Irwin Howard, eds., *Problems in Japanese syntax and semantics*. Tokyo: Kaitakusha.

LaPorte-Grimes, Laurel. 1996. What "what about answers" say about verbal morphology. In Asako Uchibori and Kazuko Yatsushiro, eds., *Papers in honor of Mamoru Saito*. (University of Connecticut Working Papers in Linguistics 7.) Department of Linguistics, University of Connecticut, Storrs. [Distributed by MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.]

Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19, 335 391.

Lasnik, Howard. 1985. Illicit NP movement: Locality conditions on chains? *Linguistic Inquiry* 16, 481 490.

Lasnik, Howard. 1995. A note on pseudogapping. In Rob Pensalfini and Hiroyuki Ura, eds., *Papers in minimalist syntax*. (MIT Working Papers in Linguistics 27.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15, 235 289.

Lasnik, Howard, and Mamoru Saito. 1992. *Move* a*: Conditions on its application and output*. Cambridge, Mass.: MIT Press.

Matsumoto, Yo. 1992. On the wordhood of complex predicates in Japanese. Doctoral dissertation, Stanford University, Stanford, Calif.

Matsumoto, Yo. 1996. A syntactic account of light verb phenomena in Japanese. *Journal of East Asian Linguistics* 5, 107 149.

Miyagawa, Shigeru. 1989. Light verbs and the ergative hypothesis. *Linguistic Inquiry* 20, 659 668.

Miyamoto, Yoichi. 1994. Secondary predicates and tense. Doctoral dissertation, University of Connecticut, Storrs.

Murasugi, Keiko. 1991. Noun phrases in Japanese and English: A study in syntax, learnability and acquisition. Doctoral dissertation, University of Connecticut, Storrs.

Murasugi, Keiko. 1992. Locative/temporal vs. manner/reason phrases. *Kinjo Gakuin Daigaku Ronshu* 33, 153 170.

Nakajima, Takashi. 1993. Syntactic analysis of the light verb construction. In Edward J. Rubin and Michael Bernstein, eds., *Cornell working papers in linguistics 11*. Department of Linguistics, Cornell University, Ithaca, N.Y.

Park, Myung-Kwan. 1991. Experiencer movement and anaphor binding in causative psych-verb constructions in Korean. Ms., University of Connecticut, Storrs.

Pesetsky, David. 1995. *Zero syntax*. Cambridge, Mass.: MIT Press.

Poser, William. 1981. The "Double-*o* Constraint": Evidence for a direct object relation in Japanese. Ms., MIT, Cambridge, Mass.

Rizzi, Luigi. 1990. *Relativized Minimality*, Cambridge, Mass.: MIT Press.

Saito, Mamoru. 1982. Case marking in Japanese: A preliminary study. Ms., MIT, Cambridge, Mass.

Saito, Mamoru. 1989. Scrambling as semantically vacuous A'-movement. In Mark Baltin and Anthony Kroch, eds., *Alternative conceptions of phrase structure*. Chicago: University of Chicago Press.

Saito, Mamoru. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1, 69 118.

Saito, Mamoru, and Keiko Murasugi. 1990. N'-deletion in Japanese. *Japanese/Korean Linguistics* 1, 285 301.

Sakai, Hiromu. 1994. Uniformity in syntactic derivation. Ms., University of California, Irvine.

Sato, Yutaka. 1993. Argument transfer: All or nothing. *Gengo Kenkyuu* 103, 92 127.

Sells, Peter. 1988. More on light verbs and q-marking. Ms., Standford University, Stanford, Calif.

Shibatani, Masayoshi. 1973. Semantics of Japanese causativization. *Foundations of Language* 9, 327 373.

Tada, Hiroaki. 1990. Scrambling(s). Ms., MIT, Cambridge, Mass.

Takahashi, Daiko. 1992. Improper movements and chain formation. Ms., University of Connecticut, Storrs.

Takahashi, Daiko. 1996. Antecedent-contained deletion in Japanese. In Asako Uchibori and Kazuko Yatsushiro, eds., *Papers in honor of Mamoru Saito*. (University of Connecticut Working Papers in Linguistics 7.) Department of Linguistics, University of Connecticut, Storrs. [Distributed by MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.]

Takano, Yuji. 1996. Movement and parametric variation in syntax. Doctoral dissertation, University of California, Irvine.

Takezawa, Koichi. 1987. A configurational approach to Case-marking in Japanese. Doctoral dissertation, University of Washington, Seattle.

Terada, Michiko. 1990. Incorporation and argument structure in Japanese. Doctoral dissertation, University of Massachusetts, Amherst.

Tsujimura, Natsuko. 1990a. Ergativity of nouns and Case assignment. *Linguistic Inquiry* 21, 277 287.

Tsujimura, Natsuko. 1990b. Unaccusative hypothesis and noun classification. *Linguistics* 28, 925 957.

Uchida, Yoshiko, and Mineharu Nakayama. 1993. Japanese verbal noun constructions. *Linguistics* 31, 623 666.

Williams, Edwin. 1981. Argument structure and morphology. *The Linguistic Review* 2, 84 114.

Zubizarreta, Maria Luisa. 1985. The relation between morphophonology and morphosyntax: The case of Romance causatives. *Linguistic Inquiry* 16, 247 289.

Chapter 9
Move F and Raising of Lexical and Empty DPs

Daiko Takahashi

In the principles-and-parameters framework it has become a standard assumption that a (or the) transformational operation Move is triggered by the necessity to check features (see Chomsky 1993, 1995). In Chomsky's (1995) view, for instance, a category a moves into the checking domain of a head H (say, the specifier position of H) so that some feature F of a may check the corresponding feature F' of H. A question that should arise naturally in this conception of movement is why it is not just F, instead of a, that moves. Although language offers abundant cases of category movement such as DP-raising, *wh*-movement, and head movement, it is quite difficult to find cases of pure feature movement (namely, cases where only features move out of DPs, *wh*-phrases, or heads), which nonetheless should be expected under Chomsky's (1995) theory of movement.

The answer Chomsky (1995) offers is that although only F should move in the ideal situation, category movement is allowed if forced by some extraneous factor. Specifically, Chomsky (pp. 262 263) phrases this idea in terms of the following conditions ((2) is slightly modified without affecting the content for the purpose of this discussion):

(1) F carries along just enough material for convergence.

(2) Words whose features are isolated or scattered may not be subject to PF rules, making the derivation crash (that is, not converge) at PF.

According to the condition in (2), hereafter called the *PF Integrity Condition*, words whose features are extracted cannot be subject to rules of the PF component and hence cannot receive PF interpretation, ending up as illegitimate objects at PF. This defines the extraneous factor that forces less economical category movement: ideally, only F should move out of a, but this would make a an illegitimate PF object, causing the derivation to crash. As an alternative, therefore, a is permitted to move by the economy condition in (1).

The feature movement (or Move F) hypothesis implemented this way makes a number of predictions. One of them, noted by Chomsky (1995), is that since the PF Integrity Condition is simply inapplicable in LF, LF movement should be restricted to feature movement. On the basis of facts regarding expletive constructions, Lasnik (1995) argues that this prediction is indeed borne out (for detailed discussion, see Lasnik's paper). A second prediction is that if a, a category containing a feature F that is about to enter into a checking relation, is phonologically null, that is, an empty category, then movement of F alone out of a should be possible (or, in fact, obligatory given (1)) even in overt syntax. 1 This is because a, an empty category, is by nature exempt from PF interpretation, so that it should not cause any trouble even if its features are isolated or scattered. This leads to the expectation that other things being equal, empty categories should behave differently from their lexical counterparts when they contain features that need to undergo checking.

In this chapter I confirm the second prediction on the basis of a raising construction in Japanese. As I discuss in detail, a certain kind of empty subject, unlike its lexical counterparts, appears to be unable to undergo raising. I show that this asymmetry directly follows from, and hence confirms, the Move F hypothesis.

The remainder of this chapter is organized as follows: In section 9.1 I introduce an adverbial temporal clause in Japanese that plays key roles in later sections. In section 9.2, by using the adjunct clause as a probe, I demonstrate the existence of quasi-argumental empty pronoun (*pro*) subjects as well as argumental *pro* subjects in Japanese. In section 9.3 I examine the behavior of various types of DPs in a raising construction and observe that whereas lexical DPs, argumental or quasi-argumental, are uniformly able to raise, empty subjects exhibit heterogeneous behavior; to be specific, I observe that whereas argumental *pro* subjects can raise like lexical subjects, quasi-argumental *pro* subjects cannot. In section 9.4 I provide an analysis for the asymmetry between lexical subjects and quasi-argumental *pro* subjects in terms of the Move F hypothesis, pointing out some of its consequences. In section 9.5 I consider argumental *pro* subjects, trying to accommodate their behavior under the analysis given in section 9.4. In section 9.6 I summarize the chapter.

9.1 Nagara Clauses

The conjunction *nagara* 'while' heads synchronous adverbial clauses in Japanese. The following is an example:

(3) Taroo1-ga [PRO1 biiru-o nomi nagara] benkyoosita.
Taroo-NOM beer-ACC drinking while studied
'Taroo studied while drinking beer.'

This sentence roughly means that Taroo's studying was synchronous with his drinking of beer. I assume that the subject position in the *nagara* clause is occupied by PRO controlled by the matrix subject, as depicted in (3). This is supported by the following two facts. First, the subject of the adverbial clause must be taken to be identical to the matrix subject. Thus, (3) cannot be understood to mean that Taroo's studying was synchronous with someone else's drinking of beer. Second, no lexical DP can appear in the subject position. Compare the following example with (3):

(4) *Taroo1-ga [Hanako-ga/kare1-ga/zibun1-ga biiru-o
Taroo-NOM Hanako-NOM/he-NOM/himself-NOM beer-ACC
nomi nagara] benkyoosita.
drinking while studied
'*Lit*. Taroo studied while Hanako/he/himself drinking beer.'

Here the appearance of the lexical DPs in the embedded subject position results in ungrammaticality. 2

As is often the case with this sort of control construction, the controller of the PRO subject of a *nagara* clause is restricted to a c-commanding subject. Consider the following examples:

(5) Taroo1-ga Hanako2-o [PRO1/*2 biiru-o nomi nagara] sikatta.
Taroo-NOM Hanako-ACC beer-ACC drinking while scolded
'Taroo1 scolded Hanako2 while PRO1/*2 drinking beer.'

(6) [Taroo1-no tomodati]2-ga [PRO*1/2 biiru-o nomi
Taroo-GEN friend-NOM beer-ACC drinking
nagara] benkyoosita.
while studied
'Taroo1's friend2 studied while PRO*1/2 drinking beer.'

(7) Taroo1-ga [PRO1/*2 biiru-o nomi nagara]
Taroo-NOM beer-ACC drinking while
[Hanako2-ga benkyoosita to] itta.
Hanako-NOM studied that said
'Taroo1 said [that Hanako2 studied] [while PRO1/*2 drinking beer].'

Example (5) indicates that the object cannot qualify as the controller of PRO. In (6) the genitive phrase in the matrix subject does not c-command PRO and is unable to be its antecedent. In (7) the *nagara* clause is intended

to modify the matrix predicate headed by *itta* 'said', so that the lower subject *Hanako* fails to c-command and hence control PRO. With these facts in hand, I next discuss two types of *pro* subjects.

9.2 Two Types of Pro Subjects

It has been recognized since Kuroda 1965 that empty pronominal subjects exist in Japanese. Consider the following example:

(8) *pro* Hanako-o sikatta.
Hanako-ACC scolded
'I/We/You/He/She/They scolded Hanako.'

Here the subject is not overtly expressed but is understood to refer to someone established elsewhere in the context.

There are at least three reasons for postulating a null subject in (8). First, the q-Criterion requires something to be present in the subject position to receive the external q-role from the verb in (8).

Second, the existence of *pro* helps explain why (8) conforms to the following condition and hence is grammatical: 3

(9) Every tensed sentence must have at least one nominative phrase in Japanese. (See Shibatani 1977.)

This condition is assumed to be in effect in Japanese syntax to account for the Case pattern in so-called stative predicate constructions (see Kuno 1973). To illustrate this pattern, I begin with the following example:

(10) Taroo-ga eigo-o hanasita.
Taroo-NOM English-ACC spoke
'Taroo spoke English.'

This is a simple transitive sentence and exhibits the canonical Case pattern for transitive sentences in Japanese, where the subject is marked nominative and the object accusative. To the stem of the verb, the potential morpheme *e* 'can' may be attached, making the resulting complex predicate stative.

(11) Taroo-ga eigo-o hanas-e-ta.
Taroo-NOM English-ACC speak-can-PAST
'Taroo could speak English.'

Although (11) preserves the canonical Case array for transitive sentences, the attachment of the potential morpheme allows additional Case pat-

terns for the subject and object. First, the accusative Case for the object may be converted into nominative, as shown in (12).

(12) Taroo-ga eigo-ga hanas-e-ta.
Taroo-NOM English-NOM speak-can-PAST
'Taroo could speak English.'

Further, the nominative Case for the subject in (12) may be changed to dative, giving rise to (13).

(13) Taroo-ni eigo-ga hanas-e-ta.
Taroo-DAT English-NOM speak-can-PAST
'Taroo could speak English.'

However, the subject cannot be dative if the object remains accusative.

(14) *Taroo-ni eigo-o hanas-e-ta.
Taroo-DAT English-ACC speak-can-PAST

The impossibility of (14) is readily accounted for by the condition in (9), since the sentence does not contain a nominative phrase. In contrast, all the grammatical examples in (10), (11), (12), and (13) contain at least one nominative phrase, satisfying (9).

With (9) in mind, let us consider (8). In clear contrast with (14) the sentence is perfectly grammatical, although it appears to lack a nominative phrase. By assuming that the postulated null subject bears nominative Case, we can straightforwardly account for its grammaticality.

Third, the well-formedness of the following example reinforces the validity of positing *pro* in (8):

(15) *pro*1 [PRO1 biiru-o nomi nagara] Hanako-o sikatta.
beer-ACC drinking while Hanako-ACC scolded
'I/We/You/He/She/They scolded Hanako while drinking beer.'

In section 9.1 I observed that the *nagara* clause contains a PRO subject that needs to be controlled by the matrix subject. Postulating the *pro* matrix subject directly provides the requisite controller for PRO in (15).

Let us now turn to null subjects in sentences such as these: 4

(16) a. *pro* sigureta.
rained
'It rained.'
b. *pro* fubuita.
snowed
'It snowed.'

(17) *pro* FBI-niyotte [Taroo-ga yuuzai dearu to] syoomeisareta.
FBI-by Taroo-NOM guilty is that was proved
'It was proved by the FBI that Taroo was guilty.'

It is impossible or at least difficult to give a q-theoretic reason for postulating *pro* in these cases, since it is unclear whether the predicates assign any q-role to the subjects in (16) and it is reasonable that the matrix subject position is dethematized owing to passivization in (17).

Nonetheless, it is necessary to posit *pro* subjects in cases like (16) and (17) for two reasons. First, recall condition (9), according to which some nominative phrase must be present in (16) and in the matrix clause in (17). As noted by Kuroda (1978), the *pro* subjects in (16) and (17) should count as the nominative phrases that make the sentences well formed with respect to (9).

The second reason comes from examining cases where the expressions in (16) and (17) are embedded in adverbial clauses headed by *nagara*. First consider the following example:

(18) *pro* [PRO sigure nagara] fubuita.
raining while snowed
'It snowed while raining.'

Suppose that the weather sentences in (16) were literally subjectless, hence that the *nagara* clause in (18) were also subjectless: in other words, that the PRO subject did not exist. If this were correct, it would be difficult to explain why the following example is ungrammatical although (18) is grammatical:

(19) *Taroo-ga [PRO sigure nagara] hasitta.
Taroo-NOM raining while ran
'*Taroo ran while raining.'

If there were no PRO subject in the *nagara* clause, we could not take recourse to any explanation in terms of PRO to rule (19) out. If, on the other hand, PRO is present as indicated above, then we can readily exclude the sentence by saying that the matrix subject *Taroo* does not qualify as the proper controller of PRO, a claim supported by the following ungrammatical example:

(20) *Taroo-ga sigureta.
Taroo-NOM rained
'*Taroo rained.'

Returning to (18), we expect it to be grammatical since the matrix *pro* subject serves as the proper controller of PRO in the adjunct clause.

The same argument can be constructed for (17). Consider the following examples:

(21) *pro* [PRO FBI-niyotte [Taroo-ga yuuzai dearu to]
FBI-by Taroo-NOM guilty is that
syoomeisare nagara] CIA-niyotte [Hanako-ga muzai
being proved while CIA-by Hanako-NOM innocent
dearu to] syoomeisareta.
is that was proved
'It was proved by the CIA that Hanako was innocent while being proved by the FBI that Taroo was guilty.'

(22) *Zisin-ga [PRO FBI-niyotte [Taroo-ga yuuzai
earthquake-NOM FBI-by Taroo-NOM guilty
dearu to] syoomeisare nagara] okita.
is that being proved while occurred
'*An earthquake occurred while being proved by the FBI that Taroo was guilty.'

(23) *Zisin-ga FBI-niyotte [Taroo-ga yuuzai dearu
earthquake-NOM FBI-by Taroo-NOM guilty is
to] syoomeisareta.
that was proved
'*An earthquake was proved by the FBI that Taroo was guilty.'

Examples (21) and (22) show that the expression in (17), with the verb inflected properly, can be embedded as the complement of *nagara* only if the matrix clause is a CP-extraposition construction of the same kind. This fact can be analyzed as follows: just as (17) has *pro* as the matrix subject, so the *nagara* clauses in (21) and (22) have PRO subjects. In (21) the PRO subject is properly controlled by the matrix *pro* subject, which, given the above reasoning, must be present too. In (22), by contrast, the PRO subject lacks the proper controller; note that the matrix subject in (22) cannot serve as the controller, as is evident from the ungrammaticality of (23). This way of analyzing (21) and (22) would not be available if the *pro* subject in (17) and hence the PRO subjects in (21) and (22) were literally absent. Therefore, to the extent that postulating these null subjects helps us explain the fact in question, we have an argument for doing so.

To sum up, I have observed that two types of empty pronominal subjects exist in Japanese: one type functions as the recipient of a q-role and is exemplified in (8); the other seems to occur independently of q-roles and is found in weather sentences like (16) and in CP-extraposition constructions like (17). I call the first type of null subject *argumental pro*, and the second type *quasi-argumental pro*, following the terminology in Chomsky 1981, 1986. 5

9.3
Raising in Japanese

In this section I consider the behavior of lexical argumental and quasi-argumental DPs and their null counterparts in a raising construction, examining whether they can undergo raising. The raising verb I use here is *hazime* 'begin', which, in contrast to the typical control verb *wasure* 'forget', is insensitive to type of subject, a hallmark of raising predicates (see Koizumi 1995 and the references therein for detailed discussion of raising and control predicates in Japanese). Compare (24) with (25).

(24) a. Hanako-ga tegami-o kaki hazimeta.
Hanako-NOM letter-ACC writing began
'Hanako began writing a letter.'
b. Komakai tyuui-ga nittyuu kankei-ni
close attention-NOM Japan-China relationship-to
haraware hazimeta.
being paid began
'Close attention began being paid to the Japan-China relationship.'

(25) a. Hanako-ga tegami-o kaki wasureta.
Hanako-NOM letter-ACC writing forgot
'Hanako forgot to write a letter.'
b.*Komakai tyuui-ga nittyuu kankei-ni
close attention-NOM Japan-China relationship-to
haraware wasureta.
being paid forgot
'*Close attention forgot to be paid to the Japan-China relationship.'

Whereas *hazime* 'begin' allows both the argumental DP *Hanako* and the quasi-argumental idiom chunk *komakai tyuui* 'close attention' to appear

as the subjects in (24), *wasure* 'forget' permits only the former as its subject in (25).

It is a little hasty, however, to conclude that raising really takes place in (24), since the word order does not tell us much about whether the subjects are actually moved. For example, (24a) can be analyzed either as in (26a) or as in (26b).

(26) a. [IP Hanako1-ga [IP *t*1 tegami-o kaki] hazimeta]
b. [IP D [IP Hanako1-ga tegami-o kaki] hazimeta]

Whereas the subject raises in (26a), it stays in situ in (26b). Both of these representations are consistent with the actual word order of the sentence. In order to show that the subject can indeed raise into the matrix clause in (24a), an adverbial clause headed by *nagara* may be added to the sentence. Consider the following example:

(27) Hanako1-ga [PRO1 biiru-o nomi nagara] tegami-o
Hanako-NOM beer-ACC drinking while letter-ACC
kaki hazimeta.
writing began
'Hanako began writing a letter while drinking beer.'

This sentence is ambiguous with respect to which predicate the *nagara* clause modifies. When (27) has the interpretation that at some point during Hanako's drinking of beer, Hanako's writing of a letter began, the adjunct clause modifies the matrix predicate headed by *hazimeta* 'began' and thus is located in the matrix clause. On the other hand, when (27) means that Hanako began the action of drinking beer and writing a letter simultaneously, the adjunct clause modifies the lower predicate headed by *kaki* 'writing' and hence belongs to the embedded clause. Of importance to the present discussion is the availability of the first reading. This interpretation stems from the *nagara* clause's being located in the matrix clause. In order for the PRO subject in the adjunct clause to be controlled, the antecedent *Hanako* must be in the matrix clause as well. This is schematized in (28).

(28) [IP Hanako1-ga [PRO1 biiru-o nomi nagara][IP *t*1 tegami-o kaki]
hazimeta]

This consideration ensures that the subject can indeed raise in (24a).

The same sort of ambiguity is obtained if a *nagara* clause is added to (24b).

(29) Komakai tyuui1-ga [PRO1 nitibei kankei-ni
close attention-NOM Japan-USA relationship-to
haraware nagara] nittyuu kankei-ni(-mo)
being paid while Japan-China relationship-to(-also)
haraware hazimeta.
being paid began
'Close attention began being paid (also) to the Japan-China
relationship while being paid to the Japan-USA relationship.'

Just like (27), this sentence is ambiguous with respect to the location of the adjunct clause. (29) means either that at some point during the careful treatment of the Japan-USA relationship, close attention began being paid to the Japan-China relationship, or that close attention began being paid to the Japan-USA relationship and to the Japan-China relationship simultaneously. Of relevance here is the possibility of the first reading, which indicates that the idiom chunk subject *komakai tyuui-ga* 'close attention-NOM' can raise into the matrix clause so that it can control (c-command) PRO in the adjunct clause, as depicted in (30).

(30) [IP komakai tyuui1-ga [PRO1 nitibei kankei-ni haraware nagara]
[IP *t*1 nittyuu kankei-ni(-mo) haraware] hazimeta]

From the discussion so far, it is clear that lexical DPs, argumental or quasi-argumental, can undergo raising. 6

Let us now turn to the empty subjects discussed in section 9.2 and consider whether they can raise. The following are relevant examples:

(31) *pro* Hanako-o sikari hazimeta.
Hanako-ACC scolding began
'I/We/You/He/She/They began scolding Hanako.'

(32) *pro* fubuki hazimeta.
snowing began
'It began snowing.'

(33) *pro* FBI-niyotte [Taroo-ga yuuzai dearu to]
FBI-by Taroo-NOM guilty is that
syoomeisare hazimeta.
being proved began
'It began being proved by the FBI that Taroo was guilty.'

Whereas the subject in (31) is an argumental *pro*, those in (32) and (33) are quasi-argumental *pros*. It is not obvious from these examples where

the null subjects are located. In order to resolve the indeterminacy, let us add *nagara* clauses to the sentences.

The following example is derived by attaching a *nagara* clause to (31):

(34) *pro*1 [PRO1 biiru-o nomi nagara] Hanako-o
beer-ACC drinking while Hanako-ACC
sikari hazimeta.
scolding began
'I/We/You/He/She/They began scolding Hanako while drinking beer.'

This sentence is ambiguous just like (27) and (29), and especially can be understood to mean that at some point during drinking of beer, I/we/you/he/she/they began scolding Hanako. The availability of this reading ensures that the argumental *pro* subject can raise so that it controls PRO in the adjunct clause modifying the matrix predicate, as schematized in (35).

(35) [IP *pro*1 biiru-o nomi nagara][IP *t*1 Hanako-o sikari]
hazimeta]

Let us next consider the following examples, which are obtained by adding *nagara* clauses to (32) and (33):

(36) *pro*1 [PRO1 sigure nagara] fubuki hazimeta.
raining while snowing began
'It began snowing while raining.'

(37) *pro*1 [PRO1 CIA-niyotte [Hanako-ga muzai dearu to]
CIA-by Hanako-NOM innocent is that
syoomeisare nagara] FBI-niyotte [Taroo-ga yuuzai
being proved while FBI-by Taroo-NOM guilty
dearu to] syoomeisare hazimeta.
is that being proved began
'It began being proved by the FBI that Taroo was guilty while being proved by the CIA that Hanako was innocent.'

Significantly, these sentences are unambiguous, in clear contrast with (27), (29), and (34). Specifically, (36) does *not* mean that at some point during the rainfall, snow began to fall, but only that it began raining and snowing simultaneously; likewise, (37) does *not* mean that at some point during the CIA's proof of Hanako's innocence, the FBI began proving Taroo's guilt, but only that synchronous proofs of Hanako's innocence and Taroo's

guilt began. The interpretations that should arise from the adjunct clauses' modifying the matrix predicates are absent in these cases.

This fact indicates that (36) and (37) cannot be represented as in (38) and (39), respectively.

(38) *[IP *pro*1 [PRO1 sigure nagara] [IP *t*1 fubuki] hazimeta]

(39) *[IP *pro*1 [PRO1 CIA-niyotte [Hanako-ga . . . to] syoomeisare nagara] [IP *t*1 FBI-niyotte [Taroo-ga . . . to] syoomeisare] hazimeta]

The impossibility of these representations follows if the quasi-argumental *pro*s are unable to raise and hence cannot control PRO in the adjunct clauses.

Note incidentally that the readings available for (36) and (37) stem from the adjunct clauses' modifying the lower predicates and hence from their being located in the embedded clauses. Thus, they are allowed even if the *pro* subjects stay in situ, as shown in (40) and (41).

(40) [IP D [IP *pro*1 [PRO1 sigure nagara] fubuki] hazimeta]

(41) [IP D [IP *pro*1 [PRO1 CIA-niyotte [Hanako-ga . . . to] syoomeisare nagara] FBI-niyotte [Taroo-ga . . . to] syoomeisare] hazimeta]

Here, there is nothing wrong with the control relation between the *pro* subjects in situ and the PRO subjects, which is also clear from (18) and (21).

To summarize, I have observed in this section that whereas lexical DPs, argumental or quasi-argumental, can raise uniformly, empty subjects exhibit heterogeneous behavior: argumental *pro*s are able to raise, but quasi-argumental *pro*s are not.

9.4
A Move F Analysis

Let us now attend to the asymmetry between lexical DPs and quasi-argumental *pro*s, postponing the analysis of the behavior of argumental *pro*s until the next section. Let us start by considering the case where lexical DPs appear in the raising construction and are about to raise, as schematically represented in (42), where *RV* stands for the raising verb and the order is irrelevant.

(42) [I' I(. . . F' . . .) [VP RV [IP . . . DP(. . . F . . .) . . .]]] (X(. . . F . . .) indicates that X consists of F and other features.)

By assumption, DP contains a feature F that needs to check the corresponding feature F' of the matrix I. 7 Ideally, just F should raise to I to check F', but this would make DP uninterpretable at PF and hence cause the derivation to crash in PF, owing to the PF Integrity Condition. As the second option, therefore, the economy condition in (1) permits the entire DP to raise to the specifier position of the matrix I. These two options are illustrated in (43).

(43) a. *[I' F-I(. . . F' . . .) [VP RV [IP . . . DP(. . . *t*F . . .) . . .]]] (* is due to (2))
b. [IP DP(. . . F . . .) [I' I(. . . F' . . .) [VP RV [IP . . . *t*DP . . .]]]]

This way, the possibility of raising of lexical DPs is accommodated under the Move F hypothesis.

Turning to quasi-argumental *pro*s, suppose that they occur in the raising construction and contain a feature F that must check F' of I, as shown in (44).

(44) [I' I(. . .F' . . .) [VP RV [IP . . . *pro*(. . . F . . .) . . .]]]

Again, the minimal operation should raise just F to the matrix I unless such raising yields an illegitimate PF object. Unlike in (42), the interfering PF factor does not play a role in this case, since *pro*, by nature, is exempt from PF interpretation. Therefore, only F raises to I to check F', blocking the alternative derivation in which *pro* raises in its entirety to the specifier position of IP. This is schematically indicated in (45).

(45) a. [I' F-I(. . .F' . . .) [VP RV [IP . . . *pro*(. . .*t*F . . .) . . .]]]
b. *[IP *pro*(. . . F . . .) [I' I(. . . F' . . .) [VP RV [IP . . . *tpro* . . .]]]] (* is due to (1))

Thus, the Move F hypothesis predicts that there is an asymmetry between lexical DPs and quasi-argumental *pro*s: lexical DPs raise to the specifier position of IP as feature movement out of them is not permitted, whereas quasi-argumental *pro*s do not raise; only F moves out of the latter.

Let us now recall (28), (30), (38), and (39), all of which are repeated here.

(28) [IP Hanako1-ga [PRO1 biiru-o nomi nagara][IP *t*1 tegami-o kaki] hazimeta]
'Hanako began writing a letter while drinking beer.'

(30) [IP komakai tyuui1-ga [PRO1 nitibei kankei-ni haraware nagara] [IP *t*1 nittyuu kankei-ni(-mo) haraware] hazimeta]
'Close attention began being paid to the Japan-China relationship while being paid to the Japan-USA relationship.'

(38) *[IP *pro*1 [PRO1 sigure nagara][IP *t*1 fubuki] hazimeta]
'It began snowing while raining.'

(39) *[IP *pro*1 [PRO1 CIA-niyotte [Hanako-ga . . . to] syoomeisare nagara][IP *t*1 FBI-niyotte [Taroo-ga . . . to] syoomeisare] hazimeta]
'It began being proved by the FBI that Taroo was guilty while being proved by the CIA that Hanako was innocent.'

I noted in section 9.3 that whereas the lexical DPs can raise and control PRO in the adjuncts in the matrix IP in (28) and (30), the quasi-argumental *pro*s cannot raise and hence fail to serve as the controller of the PRO subjects of the *nagara* clauses in (38) and (39). Although it is straightforward that (28) and (30) are grammatical, the impossibility of (38) and (39) needs some qualification. According to the Move F analysis, (38) and (39) should be schematically represented as follows (in English words and word order):

(46) *[I' F-I(. . .F' . . .) [VP began [IP *pro*1(. . .*t*F . . .) . . .][while PRO1 . . .]]]

In (46) the *pro* subject itself remains in the embedded IP, only F being moved to the matrix I. Notice that F is in a position high enough to c-command PRO. Since this representation is somehow ungrammatical, either of the following two assumptions must hold:

(47) a. Features cannot control.
b. The raised F in (46) lacks some qualification necessary to control.

Assumption (47a) amounts to saying that control is a relation between categories (or XPs). This conforms with Lasnik's (1995) conclusion that features cannot be binders, but directly conflicts with Chomsky's (1995) suggestion that features can control, which is based on the alleged grammaticality of (48).

(48) There arrived three men1 without PRO1 identifying themselves.

Assuming that only the Case and f-features of the associate *three men* raise to I in LF, Chomsky suggests that those features can function as the controller of PRO. It seems, however, that more data need to be examined to reach the conclusion that features can control, since (49) points to the opposite direction and thus supports (47a).

(49) *There seemed to be someone in the bathroom without PRO seeming to be in the bedroom.

(50) Someone seemed to be in the bathroom without PRO seeming to be in the bedroom.

If the features of *someone*, raised covertly, could control PRO in (49), it would be difficult to distinguish (49) from (50).

Leaving the issue unsettled, let us suppose that (48) represents the core case, and hence that features can control. Then, the challenge is to exclude (46) while allowing (48). In order to achieve this, we must assume (47b). The question to be asked is what the qualification is that F lacks in (46) but the covertly raised features of the associate possess in (48).

The answer should lie in the difference between Japanese and English in terms of the features checked under raising. It is clear that in English, f-features are included in the set of those features. It has occasionally been questioned whether f-feature checking or agreement plays any role in Japanese syntax (see Fukui 1986, Kuroda 1988, Fukui and Saito 1992). If indeed f-features are not included in the set of the features checked under raising in Japanese, we can note a contrast between (46) and (48): the raised F in (46) consists of the Case and Extended Projection Principle (EPP) features but lacks the f-features, 8 whereas the Case and f-features of the associate raise in (48). This contrast follows if only f-features or DPs containing f-features qualify as controllers (see Cardinaletti 1997, Ura 1996).

As it is beyond the scope of this chapter to settle the issue of whether features serve as controllers, I simply assume that either (47a) is correct or (47b) holds along the lines discussed above. It should be clear, however, that whichever turns out be the case, the Move F hypothesis enables us to account for the asymmetry between lexical subjects and empty quasi-argumental subjects.

9.5 Why Can Argumental *Pros* Raise?

In section 9.3 I observed that argumental *pros* behave like lexical DPs. This observation is based on the fact that in (34) the adjunct clause can modify the matrix predicate headed by *hazimeta*, which entails that the sentence can be represented as in (35). (34) and (35) are repeated here.

(34) *pro*1 [PRO1 biiru-o nomi nagara] Hanako-o sikari
beer-ACC drinking while Hanako-ACC scolding
hazimeta.
began

'I/We/You/He/She/They began scolding Hanako while drinking beer.'

(35) [IP *pro*1 [PRO1 biiru-o nomi nagara][IP *t*1 Hanako-o sikari] hazimeta]

This causes something of a problem for the Move F analysis, according to which argumental *pros* should behave like quasi-argumental *pros*, rather than like lexical subjects.

I would like to point out, however, that this behavior of argumental *pros* does not undermine the Move F analysis. It has been argued that what have traditionally been called raising predicates are in fact ambiguous between raising and control predicates (see Lasnik and Saito 1992, Martin 1992). As noted by Lasnik (1992a), Martin (1992), and myself (Takahashi 1994), VP-ellipsis serves as a way to distinguish between control and raising/exceptional-Case-marking (ECM) infinitives: VP can be elided if it is a complement of a control I, but not if it is a complement of a raising/ECM I. Compare the examples in (51) with those in (52).

(51) a. John tried to be smart, and Mary tried to [VP *e*] also.
b. I persuaded John to be smart, and I persuaded Mary to [VP *e*] also.

(52) a. *John said there was likely/seemed to be a riot, and in fact there is likely/seems to [VP *e*].
b. *John said advantage was likely/seemed to have been taken of Mary, and in fact advantage is likely/seems to [VP *e*].
c. *I believe John to be smart, and I believe Mary to [VP *e*] also.

The contrast here shows the validity of the generalization.

Bearing this in mind, consider the following examples:

(53) a. John is likely to win the race, and Mary is likely to [VP *e*] also.
b. Although John seems not to be smart, Mary seems to [VP *e*].

(53a,b) sound much better than (52a,b), though all of them involve the standard raising predicates *be likely* and *seem*. The acceptability of (53) indicates that those predicates sometimes behave as control predicates. The comparison between (52a,b) and (53) suggests that they can function as control predicates only if their subjects are full arguments. In (52a,b) they must be raising predicates since their subjects are not full arguments (the pure expletive *there* and the quasi-argumental idiom chunk *advan-*

tage); consequently, VP-ellipsis is not allowed. In (53), on the other hand, the full argument *Mary* appears as the subject of both *be likely* and *seem*, so that they can be control predicates, permitting VP-ellipsis.

If this sort of ambiguity holds for raising predicates in Japanese as well, it is expected that the adjunct clause can be located in the matrix clause in (34). 9 This is not because the argumental *pro* subject can raise as in (35), but because it can be base-generated in the matrix clause as in (54).

(54) [IP *pro*1 [PRO1 biiru-o nomi nagara][IP PRO1 Hanako-o sikari]
hazimeta]

Here the "raising" predicate *hazimeta* 'began' is in fact a control predicate, and thus the *pro* subject is directly base-generated in the matrix clause, controlling PRO in the embedded IP. Importantly, the *pro* subject can control PRO in the adjunct in the matrix clause, yielding the interpretation in question.

Notice that I am not excluding the raising option for argumental *pro* subjects. When the verb *hazimeta* 'began' in (34) happens to choose to be a raising predicate, the sentence should have the following representation according to the Move F analysis:

(55) *[I' [VP [PRO1 biiru-o nomi nagara][IP *pro*1 (. . . *t*F . . .) Hanako-o sikari]
hazimeta] F-I(. . . F' . . .)]

Although raising of F per se should be allowed, PRO in the adjunct clause fails to be properly controlled in (55). Therefore, the raising option happens not to surface in the presence of the grammatical control option. This means that without the adjunct clause, the sentence should be structurally ambiguous as shown in (56).

(56) a. [I' [VP [IP *pro* (. . . *t*F . . .) Hanako-o sikari] hazimeta] F-I(. . . F' . . .)]
b. [IP *pro*(. . . F . . .) [I' [VP [IP PRO Hanako-o sikari] hazimeta]
I(. . . F' . . .)]]

Unfortunately, I have not yet found any way to detect this structural difference.

Recall that the cases in which quasi arguments (and expletives) appear as the subjects of "raising" predicates necessarily involve raising (see (52a,b)). Thus, when the "raising" verb *hazime* 'begin' appears with quasi-argumental *pro* subjects, the option of being a control predicate is not possible. Consequently, this offers the ideal situation to observe the effect of pure feature movement (recall the discussion in section 9.4).

9.6
Summary

In this chapter I have provided an argument for the Move F hypothesis based on the behavior of various types of subjects in a raising construction in Japanese. I have observed that whereas lexical DPs can undergo raising, quasi-argumental *pros* apparently cannot. I have shown that this asymmetry follows from the Move F hypothesis. Although argumental *pros* appear not to conform to the prediction of my analysis, I have shown that they do not undermine it, since the seemingly unexpected cases may not involve raising at all. It has become an important research topic in the minimalist framework to investigate the roles that features play in syntax, and I believe that the material in this chapter contributes to a further understanding of their workings.

Notes

It is quite a privilege for me to have this chapter included in a volume dedicated to Howard Lasnik, to whom I am indebted for the many ways in which he has taught me linguistics.

The material in this chapter was presented at the COE Syntax Workshop held at Kanda University of International Studies in March 1997 and in a seminar at Tohoku University. An earlier version was read at the second Formal Approaches to Japanese Linguistics conference and appears as Takahashi 1996. I would like to thank the participants at those meetings, especially Jun Abe, Susan Fischer, Nobuko Hasegawa, Kazuko Inoue, Yasuo Ishii, Shigeyuki Kuroda, Roger Martin, Masatake Muraki, Yuji Takano, and Hiroyuki Ura, for helpful comments and questions. I am also grateful to Zeljko Boskovic, * and Howard Lasnik for their comments and suggestions on an earlier version. Needless to say, remaining inadequacies are mine.

1. In Takahashi 1997 I provide several arguments that this prediction too is borne out, on the basis of null operator constructions.

2. In addition to synchronous clauses, the conjunction *nagara* may introduce concessive/contrastive adverbial clauses. The following is an example:

(i) [Taroo-ga ittoosyoo-o totte i nagara] kare-no tomodati-wa
Taroo-NOM the first prize-ACC taken having while he-GEN friends-TOP
daremo syoo-o toranakatta.
anyone prize-ACC did not take
'While Taroo had taken the first prize, none of his friends took any prize.'

As is clear from (i), in this usage the subject of the *nagara* clause can be overtly realized and be independent of the matrix subject, showing that control is not involved. In this chapter I ignore *nagara* clauses with concessive/contrastive interpretation.

3. The condition in (9) can be recast in the current framework as the requirement that the Case feature of Tense must be checked at least once in Japanese. Note incidentally that Japanese allows more than one nominative phrase in a clause (the so-called multiple nominative construction). See Koizumi 1995, Kuno 1973, Takezawa 1987, and the references therein for detailed discussion of the construction.

4. The verbs in (16a,b), *sigure* 'rain intermittently' and *fubuk* 'snow and blow hard', presumably derive from the nouns *sigure* 'intermittent rainfall (particularly in late autumn)' and *fubuki* 'blizzard', respectively. Only for ease of exposition, I gloss and translate them here simply as 'rain' and 'snow', respectively.

5. Chomsky (1981, 1986) devises the term *quasi arguments* for the pleonastic subjects in weather sentences and CP-extraposition constructions, partly because it is necessary to distinguish them from pure expletives. Although these two types of "dummy" subject appear to be alike in that they occur in (seemingly) $\bar{\theta}$ (non-q) positions, they do behave differently in one respect: quasi arguments can be realized as PRO whereas pure expletives cannot. Compare (i) with (ii).

(i) a. It snowed without PRO raining. (Chomsky 1981)
b. It is true that John is sick without PRO being obvious that he is. (Chomsky 1986)

(ii) a. *There was a man in the bathroom without PRO being a woman in the bedroom.
b. *It is true that John is sick without PRO seeming that he is. (Chomsky 1986)

One way to approach the difference is to assume that PRO requires a q-role (see Safir 1985, Lasnik 1992b) and that quasi arguments, but not pure expletives, bear some q-role (that must be distinguished from standard q-roles). Note that the contrast between (ib) and (iib) shows that the pleonastic pronoun *it* associated with CPs has (at least) two distinct usages, one as a quasi argument and the other as a pure expletive. The *pro* subject in (17) corresponds to the former. Also, suggesting that idiom chunks like *advantage* in *take advantage of* bear some sort of q-role, Chomsky (1981) puts them in the group of quasi arguments.

6. Note that this merely indicates that raising is possible in Japanese, which suffices for the purpose of this chapter. Whether raising is obligatory in the language is not so obvious.

7. I am not concerned for the moment with the question of whether F is a single feature or a collection of features, though it becomes important later in this section.

8. Although I assume here that the feature responsible for the EPP effect is present in Japanese as well, I do so only for the sake of discussion. Even if it turns out to be absent in Japanese, its absence does not affect my analysis.

9. See Koizumi 1995 and the references therein for arguments that some raising verbs in Japanese including *hazime* 'begin' are in fact ambiguous.

References

Cardinaletti, Anna. 1997. Agreement and control in expletive constructions. *Linguistic Inquiry* 28, 521 533.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale and Samuel Jay Keyser, eds., *The view from Building 20*, 1 52. Cambridge, Mass.: MIT Press. [Also in Noam Chomsky, *The Minimalist Program*, 167 217. Cambridge, Mass.: MIT Press (1995).]

Chomsky, Noam. 1995. Categories and transformations. In *The Minimalist Program*, 219 394. Cambridge, Mass.: MIT Press.

Fukui, Naoki, 1986. A theory of category projection and its applications. Doctoral dissertation, MIT, Cambridge, Mass.

Fukui, Naoki, and Mamoru Saito. 1992. Spec-head agreement, X'-compatibility, and optional movement. Ms., University of California, Irvine, and University of Connecticut, Storrs.

Koizumi, Masatoshi. 1995. Phrase structure in minimalist syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, Mass.: MIT Press.

Kuroda, S.-Y. 1965. Generative grammatical studies in the Japanese language. Doctoral dissertation, MIT, Cambridge, Mass.

Kuroda, S.-Y. 1978. Case-marking, canonical sentence patterns, and counter equi in Japanese. In John Hinds and Irwin Howard, eds., *Problems in Japanese syntax and semantics*, 30 51. Tokyo: Kaitakusha.

Kuroda, S.-Y. 1988. Whether we agree or not: A comparative syntax of English and Japanese. *Lingvistic & Investigationes* 12, 1 47.

Lasnik, Howard. 1992a. Government, Case, and the distribution of PRO. Ms., University of Connecticut, Storrs.

Lasnik, Howard. 1992b. Two notes on control and binding. In Richard K. Larson, Sabine Iatridou, Utpal Lahiri, and James Higginbotham, eds., *Control and grammar*, 235 252. Dordrecht: Kluwer.

Lasnik, Howard. 1995. Last Resort and Attract F. In Leslie Gabriele, Debra Hardison, and Robert Westmoreland, eds., *Proceedings of the Sixth Annual Meeting of the Formal Linguistics Society of Mid-America*, 62 81. Indiana University Linguistics Club, Bloomington.

Lasnik, Howard, and Mamoru Saito. 1992. *Move* a. Cambridge, Mass.: MIT Press.

Martin, Roger. 1992. On the distribution and Case features of PRO. Ms., University of Connecticut, Storrs.

Safir, Kenneth. 1985. *Syntactic chains*. Cambridge: Cambridge University Press.

Shibatani, Masayoshi. 1977. Grammatical relations and surface cases. *Language* 53, 789 809.

Takahashi, Daiko. 1994. Sluicing in Japanese. *Journal of East Asian Linguistics* 3, 265 300.

Takahashi, Daiko. 1996. Move-F and *pro*. In Masatoshi Koizumi, Masayuki Oishi, and Uli Sauerl and, eds., *Formal approaches to Japanese linguistics* 2, 255 265. (MIT Working Papers in Linguistics 29.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Takahashi, Daiko. 1997. Move-F and null operator movement. *The Linguistic Review* 14, 181 196.

Takezawa, Koichi. 1987. A configurational approach to Case-marking in Japanese. Doctoral dissertation, University of Washington, Seattle.

Ura, Hiroyuki. 1996. Multiple feature-checking. Doctoral dissertation, MIT, Cambridge, Mass.

Chapter 10
On the Grammar of Polish Nominals

Ewa Willim

In research following the introduction of the DP analysis of the structure of nominals and the split-I hypothesis of clausal structure, distinct pieces of grammatical information associated with lexical categories have been analyzed as distinct functional features heading their own syntactic projections. A host of functional heads have been proposed in the literature, raising the question whether there is a universal set of functional features associated with lexical categories, which furthermore play the same role in the overt and covert syntax across languages, that is, are universally heads projecting into the syntax. In particular, the question is whether it is reasonable to assume a particular functional head if a language does not have suitable functional morphology and there is no evidence for a syntactically active head position or for a syntactically active specifier position that is expected to be provided by such a head.

In this chapter I will investigate some implications of the DP analysis of nominal structure for the grammar of nominals in Polish, an articleless language. In section 10.1 I will discuss the functional categories associated with the noun, with special reference to Polish. 1 Drawing on data from a number of unrelated languages, I will take the stand that there is no universal set of functional features associated with the noun and further, that functional heads need to be assumed in a language only if they are manifested.

In section 10.2 I will show that Polish lacks any lexical candidates to fill the D position besides lexical articles and that there is no evidence for syntactic N-raising to D, observed overtly in some DP languages. I will also argue that extraction out of nominals does not target [Spec, DP] in Polish. Furthermore, I will discuss the unavailability of two genitive arguments supported by the noun in Polish and argue that this fact is predicted if, compared with nominals in DP languages, Polish nominals

have one less domain in which specifier-head (Spec-head) agreement could be identified via genitive Case. As the DP projection is not manifested in Polish,. I will suggest that the noun lacks it in the overt syntax.

In section 10.3 I will investigate some implications of the DP analysis for the interpretation of nominal expressions. Drawing on the observation that the DP projection is neither necessary nor sufficient for the (in)definite reference of nominals to be established in (some) DP languages, I will suggest that Polish nominals lack the DP projection also at LF and that rather than having the D head perform an LF operation on the NP, its predicate, Polish resorts to some other means to interpret the reference of its nominals.

10.1
The Extended Syntactic Projection of NP

10.1.1
The Functional Categories of the Noun

In Chomsky 1995 the formal features (FF) complex of a lexical noun includes the head-related categorial feature and the f-features person, number, and gender. The features number and gender are associated with functional morphology in a number of unrelated languages. By hypothesis, each of the two features could be taken to project categorially into the syntax. For example, Picallo (1991) analyzes Catalan nominals as including the functional head Gen(der), which selects NP. Gen projects into GenP, which is selected by the functional head Num(ber), the locus of number specification of NP. Adapted to the framework in Chomsky 1995, this analysis entails that the noun checks its gender and number features, which it picks up as it is pulled from the lexicon, by raising successive cyclically into the respective heads Gen and Num. The heads are subsequently eliminated from the derivation, satisfying the principle of Full Interpretation.

However, Ritter (1992, 1993) argues that gender is not a functional head in the nominal complex in Modern Hebrew. Rather, in this language gender is an inherent property of the lexical stem and it is checked already in the lexicon. Consequently, gender is specified on the head of NP, not on the head of an appropriate functional projection. By contrast, number is specified in Modern Hebrew on the head of an appropriate functional projection, NumP. In Spanish, however, the single head Num specifies both the number of NP and its grammatical gender. For Haitian, a language that does not manifest gender distinctions, Ritter (1992) suggests that the noun lacks the feature [+/ Feminine] in its complex of FFs

altogether. Thus, gender seems to be a purely semantic feature of the lexical stem, closely associated with the animacy/inanimacy and natural gender distinctions; as such, it is assessed only at LF in Haitian. However, it is absent both morphologically and syntactically in Haitian. By contrast, both semantic gender and grammatical gender are manifested in Spanish and Modern Hebrew.

Ritter's (1992, 1993) analyses suggest that a particular feature need not be found in the FF complex of the noun in a given language and that a particular functional feature need not be associated with a separate functional head. If correct, these analyses imply that crosslinguistically, the extended syntactic projection of NP may contain different functional categories.

That different NPs may contain different functional heads within the same language can be observed with respect to the inflectional category of possession. In Hungarian (see Ritter 1992) and Turkish (see Haegeman 1994), the noun manifests agreement with the genitive possessor. If this agreement is mediated through an agreement head intermediate between N and D, only nominals with a possessor can be assumed to contain this agreement head, as in the absence of the possessor subject, the agreement head could never check its feature(s), causing the derivation to crash. Arguably, a phonologically unrealized possessor cannot license a phonologically unrealized agreement head. Then, although possession may be an inflectional category of the noun in a language, it need not always project categorially into the syntax. An important implication of this analysis is that functional heads are assumed in the derivation only if they are manifested.

Languages also differ in how they define the content of D(eterminer), another functional head associated with NP. If D is the locus of features pertaining to the referentiality/deixis of nominal expressions, the noun must contain a relevant, D-related feature (e.g., [+/ R(eferential)]; see Longobardi 1994), triggering N-to-D movement in the overt or covert syntax. As the substantive definite article is a semantic determiner, D is also endowed with the feature [+/ Definite]. A lexical D is inherently specified for definiteness. However, as Ritter (1992) and Siloni (1991) point out, a phonologically unrealized D is not inherently specified as (in)definite in Modern Hebrew, and a nominal without a lexical article receives a definiteness specification from the genitive subject phrase in Hebrew construct states. 2

In fact, a degree of parametric variation in the system of functional categories associated with the noun may be expected under the assump-

tion that functional morphemes are added to the lexical noun stem in the lexicon and that what is checked in the course of the derivation are the functional features associated with such morphemes (also see Speas 1991). On this view, there is no motivation for assuming that functional features correspond uniquely to separate functional heads, that there is a universally fixed set of such heads, and that they are inherently ordered with respect to one another. Rather, if a language does not have strong functional morphology associated with a given lexical category, and there is no evidence for a syntactically active functional head or heads associated with it or a syntactically active Spec associated with the functional head, there is no reason to expect that the category should have the particular head or heads in its extended syntactic projection (also see Uriagereka 1994). In fact, on the assumption that the only functional features the category has in a language are interpretable, as interpretable features must survive until LF and need not enter checking relations overtly or covertly (see Chomsky 1995, 285), the extended syntactic projection of this category need not contain any functional heads whatsoever. Thus, rather than being assumed, functional heads must be motivated in a language.

Apart from the categorial feature and the f-features, the noun is endowed with the head-related feature of Case. If this feature is uninterpretable, it must enter a checking relation. For example, the feature [+Nominative] can be checked as a free rider on agreement in finite clauses and [+Objective] can be checked under Spec-head agreement in a functional projection of a verb assigning structural Case. If Case is a weak feature, as in English, unless checked as a free rider in the overt syntax, it is checked at LF, when the noun raises into D (see Chomsky 1995, 279). Consequently, in English the noun need not have any functional categories other than D (and perhaps Num) associated with it. On the other hand, if Case is strong, feature checking will take place in a functional projection.

Furthermore, the noun is assumed to assign inherent (genitive) Case, which is Spec-related (see Chomsky 1995). If inherent Case is tied to q-marking, then genitive Case assigned by the noun is interpretable and need not enter a checking relation. Even if it didfor example, as a free riderthe feature would not delete when checked, unlike structural Case, which is uninterpretable and deletes for convergence when checked. As inherent Case may be checked overtly in a language, I will assume here following Lasnik (1995) that inherent Case, like structural Case, is

licensed in a Spec-head relation of a functional head associated with the relevant lexical category. That is, genitive Case is a Spec-related feature checked (and/or identified) under abstract Spec-head agreement.

10.1.2 The Syntactic Structure of Polish Nominals

Nominal inflection is rich in Polish, where the noun inflects for gender, number, and m(orphological)-case. The features of gender, number, and m-case participate in morphological agreement between the noun and any adjective-like modifier it occurs with, including demonstratives, 1st and 2nd person possessive pronouns, most numerals, and both attributive and predicative adjectives. Number, gender, and m-case are marked on the noun in the form of a portmanteau affix.

The Polish noun has seven m-cases (N(ominative), G(enitive), D(ative), A(ccusative), I(nstrumental), Ab(lative), and V(ocative)), three genders (M(asculine), F(eminine), and N(euter)), and two numbers (Sg and Pl). 3 Almost every combination of these features may be expressed through more than one suffix. For example, the features M, Sg, N are expressed by the suffixes *-a*, *-Æ*, *-o*, and the features F, Sg, N are expressed by the suffixes *-a*, *-i*, *-Æ* (see Klemensiewicz 1986, 70 71). As the same suffixes can realize different sets of features, there is a high degree of syncretism in the language.

The paradigmatic set consisting of the seven suffixes that each nominal stem takes in the singular and the seven suffixes it takes in the plural forms its declension, representing a particular declensional type. All nouns that decline according to the same declensional type are grouped in a declensional class, subclass, and paradigm (see Grzegorczykowa, Laskowski, and Wróbel 1984). The rules assigning a noun to a declensional class are morphological rules accessing information about the gender of the noun, the phonological shape of the stem, the meaning of the stem, and the morphological structure of the noun (see Grzegorczykowa, Laskowski, and Wróbel 1984, 263). For example, I Sg may be expressed by the suffix *-Æ* or the suffix *-em*. The suffix *-Æ* attaches to F nouns whereas the suffix *-em* attaches to all other nouns. In N Sg the suffix *-Æ* is found in M nouns except for diminutive proper names and collective nouns in *-o*, and in F nouns with a soft stem.

Unlike gender, number, and m-case, the declensional class feature is invisible to the overt syntax in that it does not participate in agreement. I propose to capture this distinction by assuming that declensional class is a feature of the lexical stem and that it is checked in the lexicon. In other

words, the declensional class feature is a purely morphological feature. On the other hand, the features gender, number, and m-case are visible both in the morphological component and in the syntax in Polish.

That gender is a feature of the inflectional suffix rather than the lexical stem is suggested by the fact that an inflectional suffix can change the gender of the noun. The suffix *-a*, typical of feminine nouns in Polish, has a derivational function in *markiz/markiza* 'marquis/marchioness', *chrzestny/chrzestna* 'godfather/godmother', and so on. Notice, furthermore, that an animate noun assigned to a given gender category may have referents of either sex: for example, *osoba* 'person' and *postac* * 'character' are morphologically feminine but may refer to persons of both sexes. Moreover, an animate noun, typically a diminutive or an augmentative, assigned to a given gender category may have a referent whose sex need not correspond to its grammatical gender. For example, *kociak* 'young, attractive female' is grammatically masculine, *babsko* 'virago' is neuter but has a female referent, and *chlopisko* 'hunk of a man' is neuter but has a male referent. I thus submit that gender is both semantic/notional and grammatical in Polish. Semantic/notional gender is an inherent feature of the lexical stem in Polish and is interpretable. Grammatical gender is an uninterpretable feature. In fact, it is simply the kind of agreement the noun triggers, and it is a feature of the inflectional suffix.

On the assumption that number is an interpretable feature of the noun (see Chomsky 1995), even if the feature [+/ P1] were checked in the overt syntax, it would not be deleted. [+/ P1] participates in agreement, but arguably, [+/ P1] is not an interpretable feature of the adjective. On the assumption that an interpretable feature of a lexical head cannot check an uninterpretable feature of a functional head, as there would be feature mismatch, it is necessary to distinguish the semantic from the grammatical feature of number. Then, just like gender, number may be an interpretable feature of the stem, correlated with the grammatical feature of number, itself an uninterpretable feature of the suffix, in effect, a cover term for a specific kind of inflectional morphology. In this way, nouns that are morphologically singular but notionally plural (e.g., *profesorostwo Kowalscy* 'Professor and Mrs. Kowalski') can be accounted for straight-forwardly. In the case of a stem/suffix feature mismatch, either the features of the stem or the features of the suffix determine agreement in a language. Unlike gender, which is intrinsic, number is an optional feature of the lexical stem.

Although all the agreement features of the noun are realized by means of a single portmanteau suffix on all its agreeing modifiers, gender and number are realized by means of distinct suffixes in nominative agreement in finite clauses. Booij and Rubach (1987, 36 37) analyze the inflectional verb forms in (1) as composed of a separate gender marker, *-e* for M and *-a* for F, and the number and person markers, *-m* for 1Sg and *-ś* for 2Sg. Consequently, the order of morphemes in (1) is stem + thematic vowel + tense + gender + person and number.

(1) a. rob + i + ł + e + m
'I-M was doing'
b. rob + i + ł + a + m
'I-F was doing'
c. rob + i + ł + e + ś
'you-M were doing'
d. rob + i + ł + a + ś
'you-F were doing'

The facts in (1) suggest that gender and number are distinct inflectional features in Polish, but they do not necessarily suggest that each of these features is associated with a separate functional head. Nevertheless, given that the noun tends to precede its arguments and that it precedes classificatory adjective modifiers, I suggest that there are two functional heads in the extended maximal projection of the noun in Polish and that the noun raises twice in the overt syntax in Polish. 4 Assuming the NP-internal subject hypothesis and assuming that adjective modifiers are base-generated to the left of NP, the fact that the noun surfaces before its arguments (see (26), (28) (29)) and after (some) adjective modifiers (see (3) (4), (9) (10)) constitutes evidence that it undergoes (partial) N-raising.5 Following Ritter (1992), I take the grammatical gender and number of NP to be specified on the functional head Num. Since to be used in syntactic structures, the noun must be inflected for m-case in Polish, I suggest that Num is in turn selected by a functional category encoding m-case, which I will call *K(ase)*. I take K to be the highest position of the morphological checking domain of the noun in the syntax in Polish. Notice that m-case closes the noun off derivationally and inflectionally and completes the noun phrase syntactically in Polish, very much as the determiner closes off a syntactically well formed noun phrase in English.

Following Cinque (1995), I assume that adjectives are Specs in the functional projection(s) of the noun; that is, they are Specs of Num and/

or K in Polish. As adjective modifiers can enter the nominal complex in a stack, by analogy with Holmberg's (1993) analysis of (in)definite adjectives in the nominal complex in Scandinavian, I propose the following procedure for feature checking involved in nominal agreement in Polish: The noun raises successively through the functional heads to check the functional features gender, number, and m-case. The features are checked in the overt syntax and are eliminated from the functional heads, but are not deleted from the lexical stem, as the inflectional features of the lexical noun are visible to the morphological component. Except for classificatory adjectives, agreeing modifiers check their N-related features under Spec-head agreement with the noun in K. If there is a stack of modifiers, the noun-K complex continues moving through the successive empty K positions, as shown in (2), forming a chain. All links of the chain share the same features. The relevant partial derivation of a nominal with two adjective modifiers is shown in (2). 6

(2) [KP AP [K'[K e] [KP AP [K'[K e] [KP[K N] . . . [NP tN]]]]]]

Following Bosque and Picallo (1996), I assume here that the relative ordering of thematic, attributive, and classificatory adjective modifiers reflects the structure of semantic inclusion: the adjective denoting a subclass of entities with a certain property is closer to the noun than the adjective denoting a class of which the former is a proper subset. In other words, the syntactic ordering of adjective modifiers reflects the semantic hierarchy of specification. I assume that this is at the heart of the relative ordering of demonstratives, possessive adjectives, and adjective modifiers appearing in this order in Polish nominals. Hence, possessive and numeral modifiers may appear in either order in Polish, with the resulting difference in meaning, as shown in (3) and (4).

(3) moje trzy książki
my three books
'my three books'

(4) trzy moje książki
three my books
'three of my books'

Notice that if feature checking proceeds as in (2), and the noun is preceded rather than followed by its agreeing modifiers except for classificatory adjectives, the checking operation must be partly covert. To account for the postnominal position of classificatory adjectives, I suggest that they

enter into Spec-head agreement with the noun as it passes through a NumP shell on its way from Num to K. In other words, the Spec-head agreement involved in the relation between classificatory adjective and noun takes place in the overt syntax in Polish. As attributive adjectives precede the noun in an uninterrupted sequence in Polish, unlike in Italian (see Cinque 1995), they check their features against the noun after Spell-Out. If this is correct, Polish offers evidence that the mere presence of inflectional morphology in a language is not coextensive with strength and the necessity to check features overtly.

10.2
The D Paradigm

10.2.1
Evidence for D Position

Apart from lexical articles, evidence for the D position in the nominal complex derives from the occurrence of other morphemes in D. In English, D may be lexicalized with the clitic *'s*. In Dutch (see Corver 1990) and colloquial German (Josef Bayer, personal communication), D may be filled with a possessive pronoun, as illustrated in (5) and (6), respectively. Furthermore, Abney (1987) suggests that personal pronouns are base-generated in D, as shown in (7).

(5) [DP [DP dit meisje] [D d'r] [NP fiets]]
this girl her bike
'this girls's bike'

(6) [DP [DP dem Professor] [D sein] [NP Buch]]
the Professor his book
'the professor's book'

(7) [DP[D we] [NP students of linguistics]]

Polish lacks lexical articles; it has no clitic attaching to a genitive subject phrase and no analogue of either (5) or (6). If personal pronouns had to be analyzed as base-generated in D, only personal pronouns would offer morphological evidence for the D position in Polish nominals. However, there seems to be some parametric variation as regards the categorial status of personal pronouns. In Hebrew, 3rd person pronouns spell out the functional head Num, as evidenced by the fact that demonstrative pronouns are aggregates of the definite article (in D) and personal pronouns (in Num). Rouveret (1991) suggests in his account of (anti-)agreement in Welsh finite clauses that whereas lexical subjects are full DPs, pronominal subjects are

NumPs. Den Dikken (in press) extends this analysis to the treatment of possessive pronouns in English and Hungarian. If these analyses are correct, there is reason to suppose that personal pronouns do not lexicalize D. If personal pronouns are NumPs in Polish as well, their occurrence is not evidence for the existence of items of the category D in the lexicon in Polish (also see Corver 1990, Witkos * 1993).

In my opinion, the fact that an appositive can stand in relation to a proper name, as in (8), analogously to (7), casts additional doubt on the analysis of pronouns as elements in D.

(8) Paul Jones, a/the student of linguistics

Longobardi (1994) argues that English has no overt N-to-D raising and consequently, proper names never surface in D position in overt syntax in English. Then, *Paul Jones* is not in D in (8), and the appositive is not in relation to an element in D in (8) and, given the similarity of interpretation, possibly also not in (7). Then, it is not necessarily the case that in the Polish equivalent of (7), *my studenci jezykoznawstwa** the pronoun is in D.

In Polish (in)definiteness can be marked via a demonstrative or indefinite pronoun: for example, *ten* 'this', *ci* 'these', *jakis** 'a/some'. If demonstrative and indefinite pronouns are universally analyzed as determiners heading the phrasal projection from D to DP, structures with such elements manifest the need for the DP layer in nominals in Polish. However, it is debatable whether demonstratives are inserted in D or in [Spec, DP]. Chomsky (1995) analyzes determiners as elements of D, but Giusti (1994) argues that the definite article and the demonstrative in fact do not compete for the same position, that is, for the D slot. Rather, demonstratives appear in a Spec position, either in [Spec, DP] or in the Spec of the lower functional head selected by D. On this analysis, the cooccurrence of a definite article and a demonstrative observed in a number of languages (e.g., Hungarian, Romanian, Old Icelandic, and Gothic) is unsurprising (see Giusti 1994).

That demonstratives and indefinite pronouns are not heads in Polish is suggested by similarities between their agreement morphology and the agreement morphology of adjective modifiers. If nominal agreement is uniformly seen as an instance of a Spec-head relation, then the agreement exhibited by demonstrative and indefinite pronouns, which is similar to the agreement manifested by adjective modifiers, as shown in (9) and (10), is unexpected if demonstrative and indefinite pronouns were to occupy

a higher head position (e.g., the D slot in Polish) rather than a Spec position.

(9) a. ta nowa ksiazki *
this-F.SG.N new-F.SG.N book-F.SG.N
b. ten nowy dom
this-M.SG.N new-M.SG.N house-M.SG.N

(10) a. tej nowej ksiazki
this-F.SG.G new-F.SG.G book-F.SG.G
b. tego nowego domu
this-M.SG.G new-M.SG.G house-M.SG.G

Witkos* (1993) provides additional evidence for analyzing demonstratives as Specs rather than heads from elliptical constructions like (11), where the nouns following the interrogative and the demonstrative are phonologically empty.

(11) a. Ktora* wybierasz?
which-F.SG.A choose-2SG.PRES.IMPERF
'Which one are you choosing?'
b. Tę zieloną.
this-F.SG.A green-F.SG.A
'The green one.'

The rich inflectional morphology of noun modifiers in Spec arguably identifies the empty *pro* in the head position in Polish, which thus lacks lexical analogues of English *one/ones*.

If D were a functional category associated with the noun, then in the absence of functional morphology, it might be expected to be manifested otherwise, that is, by being syntactically active in Polish. Such evidence is found in various DP languages. For example, the noun moves into D overtly in the Semitic construct states (see Ritter 1992, Siloni 1991), in Romanian (see Giusti 1994), and in Italian nominals with a proper name (see Longobardi 1994). In all these languages, when the noun moves into D overtly, adjective modifiers surface following the noun. Polish lacks an analogue of the Semitic construct state. A lexical possessor is postnominal in Polish, but attributive adjectives precede the noun, as shown in (9) and (10), suggesting that if there is N-raising in Polish, it is not into D. Unlike in Italian, an attributive adjective modifying a personal name precedes the personal name, suggesting that there is no N-to-D raising in overt syntax in Polish.7 The relevant contrast is illustrated in (12) and (13).

(12) a. Cameresi vecchio
Cameresi old
'old Cameresi'
b. *vecchio Cameresi

(13) a. maly * Kowalski
young Kowalski
'the young/little Kowalski'
b. *Kowalski maly

In conclusion, there is neither functional morphology to support the functional head D in Polish, nor evidence for a phonologically unrealized but syntactically active D head in the language. In sections 10.2.2 and 10.2.3 I will look at whether there is evidence for a syntactically active [Spec, DP].

10.2.2
Evidence for [Spec, DP]: Extraction from NP

In languages with the D paradigm, D provides a position that an element can target, substituting for [Spec, DP]. The function of D is to provide a structural configuration in which features can be checked. Then, as suggested by Rizzi (1990), if an element raises out of NP, it must pass through [Spec, NP] to trigger abstract agreement with the head, which licenses the trace. A filled Spec blocks extraction out of NP. Since N is a genitive Case assigner, only a genitive NP can pass through [Spec, NP]. As genitive Case assigned by N is inherent, hence interpretable, it is unaffected by checking and can arguably be assessed again in [Spec, D], which licenses the trace of the element moved out of DP, as shown in (14).8

(14) Who did you see [DP *t* a [NP *(t)* picture of *t*]]?

Giorgi and Longobardi (1991, 231) and Rizzi (1990, 52) show that prepositional arguments and adjuncts cannot be extracted out of nominals in Italian whereas genitive arguments can, as expected in this scenario. Examples (15) and (16) illustrate the relevant contrast.

(15) l'uomo politico di cuii ho registrato [la telefonata *t*i]
the politician of/by whom (I) recorded the phone call

(16) *l'uomo politico a cuii ho registrato [la telefonata *t*i]
the politician to whom (I) recorded the phone call

Then, if a language allows *wh*-extraction of a PP argument or adjunct of the noun, the moved element cannot pass through [Spec, NP] (and

further up, [Spec, DP]). Polish allows such "extractions," as illustrated in (17) (19).

(17) Od kogoi czytasz [list *t*i]?
from who read-2SG.PRES letter
'*From whom are you reading a letter?'

(18) Do kogoi masz [zaufanie *t*i]?
to who have-2SG.PRES confidence
'*In whom do you have confidence?'

(19) Z kimi nagrywasz [rozmowy *t*i]?
with whom record-2SG.PRES conversations
'*With whom do you record conversations?'

If extraction is indeed involved in structures like (17) (19), their acceptability shows that it cannot target [Spec, NP] and/or a higher Spec position.

Extraction out of nominals is sensitive to the so-called Specificity Condition. Diesing (1992) argues that extraction is possible out of NP only if the NP has no presuppositional reading or does not require one. In contexts allowing extraction, the indefinite article allows extraction whereas the definite article (weakly) disallows it. Compare (14) with (20) in this respect (see Diesing 1992, 97).

(20) ?*Who did you see [DP *t* the [NP (*t*) picture of *t*]]?

In Polish, on the other hand, "extraction" is possible out of a non-presupposed NP containing the demonstrative *ten* 'this'. In such contexts the demonstrative is nondeictic and indefinite; therefore, it functions like a weak determiner introducing a nonpresuppositional NP. 9

(21) O czymi piszesz [te swoja ksiazke *t*i]?*
about what write-2SG.PRES.IMPERF this self book
'*About what are you writing this your book?/What are you writing this book of yours about?'

If the extended projection of NP were to include the DP layer in Polish, the demonstrative would arguably be generated in [Spec, DP]. The presence of the demonstrative in this Spec should make it impossible for the PP argument to move out of the nominal in (21) if extraction out of nominals were constrained by the requirement that the trace of the moved element be licensed through genitive Case under Spec-head agreement in Polish.10 That displacement of elements out of NP does not target the Spec position is arguably also at the heart of so-called left branch extrac-

tion phenomena observed in Polish. I therefore conclude that there is no direct evidence that D is a head providing a position targeted by a phrase undergoing overt movement in Polish.

10.2.3
Evidence for [Spec, DP]:
The Syntax of the Genitive in Polish

In Polish, where adnominal genitive is synthetic and there are no dummy prepositions to realize the genitive Case of the subject and object arguments, a nominal complex cannot contain two lexical genitive arguments. In complex event nominals the object argument is in the genitive and the subject argument is introduced by the preposition *przez*, the equivalent of the English *by*. In English (see Anderson 1983), German, 11 Catalan (see Picallo 1991), and Arabic (see Carstens 1993), two genitive arguments are possible. The relevant examples are provided in (22) (25). The Polish example (26) is given to illustrate the contrast.12

(22) John's reconstruction of an 18th-century French village

(23) Hannibals Zerstörung der Stadt
Hannibal's destruction of the city

(24) l'avaluació de la comissió dels resultats
the evaluation of the committee of the results
'the committee's evaluation of the results'

(25) taxriib-u al-ruumaan-i li-kartaaz-a*
destruction-N DEF-Romans-G of-Carthage-G
'the Romans' destruction of Carthage'

(26) a. odkrycie Ameryki przez Kolumba
discovery Ameryki-G by Columbus-A
'the discovery of America by Columbus'
b. *odkrycie Ameryki Kolumba
discovery America-G Columbus-G
'Columbus's discovery of America'
c. *odkrycie Kolumba Ameryki
d. *Kolumba odkrycie Ameryki
e. *Ameryki odkrycie Kolumba

The availability of nominals with two genitive arguments in some languages suggests that a lack of such nominals in a language cannot in general be due to a semantic constraintfor example, a reduction of the argument structure of the nominal and the necessity for the external q-role to be encoded in an adjunct introduced by a semantically contentful

preposition (*by* in English, *durch* in German, *przez* in Polish, etc.). Rather, complex event nominals and/or result nominals with two genitives are unavailable in Polish as a result of a morphological requirement that cannot be met within the nominal complex. Following the insights of Chomsky (1986) and Longobardi (1996), who suggest that the genitive Case feature assigned to a nominal phrase must be both licensed and identified within the syntactic projection of the noun, I suggest that the reason that only one argument of the noun may realize genitive Case in Polish is precisely that Polish nominals lack the DP layer and consequently, only one argument can enter the relation relevant for licensing and identification of genitive Case, in contrast to what we find in the DP languages illustrated in (22) (25).

For Longobardi (1996), Case may be identified either formally, through synthetic morphology or an adposition, or positionally, in the Spec of an appropriate functional category. The German example (23) is especially interesting in that synthetic morphology identifies two arguments, but distinguishes them in terms of position, contrasting minimally with the Polish example (26). On the plausible assumption that if a language has only one morphological mechanism for identifying the genitive Case of the noun (i.e., genitive Case can be identified only through synthetic morphology), then the genitive arguments must be distinguished in terms of occupying different Specs where they can be licensed, Polish could also identify two arguments through synthetic morphology if it could distinguish them positionally (i.e., if one were in the Spec of a functional category selecting NP in which genitive Case could be checked and the other could raise to [Spec, DP], again for reasons of feature checking/identification). If Polish lacks the DP layer, only one genitive argument can be licensed and identified. Examples like (26a) are grammatical, because once the head noun has raised into Num, establishing a checking relation, the object argument can move into [Spec, NumP], where its Case feature is licensed. The external argument checks Case within the prepositional phrase in (26a). By contrast, (26b) and (26c) are ill formed, because only one argument can check genitive Case in [Spec, NumP] whereas these structures involve two arguments in need of checking their genitive Case feature. If the object checks its Case feature, the subject is unable to check/identify its Case, feature, yielding (26b). If the subject checks its Case feature, the object is unable to check/identify genitive Case, yielding (26c). The relevant partial derivation of (26b) is shown in (27). 13

(27) *[KP[K odkryciei] [NumP Amerykij [Num *t*i] [NP Kolumba [N *t*i] *t*j]]]

In this scenario the only well-formed nominals with a genitive argument are those in which if the external argument is expressed through a genitive-marked phrase, the object argument is either sentential, prepositional, or marked for oblique m-case. Subject-Experiencer nominals typically involve a genitive-marked subject and an oblique object, as shown in (28).

(28) zafascynowanie Piotra grami komputerowymi
fascination Piotr-G computer games-I
'Piotr's fascination with computer games'

To account for the contrast between (26) and (28), I suggest that some case morphology is semantic in Polish. That is, the category K of the instrumental KP *grami komputerowymi* in (28) is headed by the appropriate Case feature, which not only checks the Case feature of the noun but also assigns a q-role to the NP it selects, very much the way the preposition *with* can be analyzed as checking the Case of and assigning a q-role to its complement in the PP *with computer games* in the English gloss to (28). Then, the oblique object in (28) does not check Case under Spec-head agreement in [Spec, NumP] and consequently, the genitive subject can check its Case in this configuration.

Notice that there is no ban on two synthetic genitive phrases appearing in a sequence in Polish nominals, although such nominals are often ambiguous. However, in structures with two genitives one is a semantically restricted Poss(essor).

(29) kolekcja znaczków Piotra
collection stamps-G Piotr-G
'Piotr's collection of stamps'

Given that Poss cannot bind an anaphor in Polish, a language with subject-oriented anaphors, I suggest that in structures with two synthetic genitives, Poss is an adjunct (also see Willim 1995). Clark (1978) suggests that Poss is a location for the object possessed and that with human possessors the location is animate. I propose that Poss encompasses both the prototypical possessor, as in *my son's friends*, and its metaphorical extension, the so-called genitive of origin denoting the Author, as in *my son's letter*. Poss is identified uniquely as this modifier of the head noun, which may be related to it across the copula in the ascriptive *X is Y's* construction, hence as the modifier of the object-denoting noun *kolekcja*

'collection' in (29), but not of a relational, abstract noun like *odkrycie* 'discovery' in (26). 14

I submit here that like the oblique Case in (28), genitive Case associated with Poss is semantic in Polish. By analogy with the proposal of Anderson (1983), who analyzes the Case of Poss to be assigned by the lexical, contentful *'s* rather than by the noun, I suggest that the interpretation of *Piotra* 'Piotr's' in (29) is tied to its inflectional morphology; that is, the suffix carries the feature [+G] associated with a functional head K, which assigns the Poss q-role to its complement NP. Being interpretable, this feature need not enter a checking relation, just as an interpretable feature of Case assigned by a semantically contentful preposition need not enter a checking relation. Nevertheless, Poss (i.e., *Piotra* 'Piotr's') in (29) identifies the Case of the abstract K of its inclusive genitive KP and not the Case of the noun *kolekcja* 'collection'. In this scenario the object argument checks its genitive Case in [Spec, NumP] after the noun raises into Num in (29). On the other hand, the Poss KP, possibly generated left-adjoined to NP, remains in its base-generated position throughout the syntactic derivation.

The analysis of the conditions on the occurrence of genitive Case-marked elements in the nominal complex in Polish reveals that only one genitive argument can be supported by the noun, in contrast to what is found in some DP languages.15 This further undermines the DP analysis of the structure of Polish nominals. For lack of functional morphology to manifest D and for lack of evidence that D is a position or that it provides a position in which features can be checked overtly, I conclude that the extended syntactic projection of NP does not include the DP layer in the overt syntax in Polish.16

10.3
The DP Hypothesis and (In)definiteness in Polish

10.3.1
Determiners, Reference, and Argumenthood

The DP hypothesis makes it possible to draw a strict parallel between the nominal and the verbal domains with respect to their interpretation. Both NP and VP may be seen as predicates whose extensions are restricted by their respective functional heads, D and T. In this scenario, just as the VP provides a class of events and the function of T is to locate a given event in time, so the NP provides a set of entities and the function of D is to determine which member of the set of entities is being referred to.17

In several languages, including English and Italian, a countable noun in the singular cannot be used in referring expressions unless it is combined with a determiner whose function is to determine the reference of the expression in which the noun appears in terms of the identity of the referent, as shown in (30) (Longobardi 1994, (6b)) for Italian and in its gloss for English.

(30) Ho incontrato *(un/il) grande amico di Maria ieri.
'I met *(a/the) great friend of Maria yesterday.'

Thus, some languages provide empirical evidence that determiners combine with nouns to produce referential expressions. On the basis of this evidence, Longobardi (1994) suggests that nominals lacking a determiner, as well as nominals in which the determiner is not a semantic restrictor but instead an expletive determiner, can only be interpreted predicatively, as in *She is a teacher*. To be an argument and hence a referring expression, a nominal must be predicated of a head selecting it. This requirement is independent of the lexical content of D. If this is correct, the DP hypothesis is a virtual necessity at LF in all languages, including all the languages that lack lexical determiners. The implication is that if D is not present in syntactic structures in Polish, it must be introduced at LF, the level at which the article is relevant and required for convergence. Then, morphological parameters would have to be augmented with lexical expansion parameters, as discussed by Uriagereka (1994).

Although lexical items are typically selected in the overt component, the operation Select may in principle insert lexical items at LF as well (see Chomsky 1995, 232). An empty article could be selected at LF in Polish, as it has no PF features; furthermore, it could be inserted at the root of the phrase marker on the assumption that nominal structures are not merged with the functional heads Agrs/T and V/n before Spell-Out, as has been suggested for Russian, a related language, by Hoffman (1998). Then, only sub-phrase markers would be assembled before Spell-Out and full merger might arguably be required only at LF for reasons of propositional semantics. Merger of an empty article at the root of the phrase marker would extend it, thus complying with the restrictions on the operation Select. Since Polish is a flexible word order language allowing scrambling as well as stylistic reorderings, this view is theoretically attractive. If correct, this hypothesis would imply that surface word order is largely determined by the conditions under which information is struc-

tured in the utterances of a language, a view discussed widely in the literature (see, e.g., Szwedek 1976, Siewierska 1993).

As I am only marginally concerned with word order considerations here, I will remain neutral with respect to the possibility that syntactic derivations may or may not involve full merger in the overt syntax in Polish. In the remainder of the chapter I will discuss some implications of the hypothesis that nominal expressions must be analyzed as DP structures at LF in Polish.

10.3.2
A Minimalist Approach to (In)definiteness of Reference

In Longobardi's (1994) approach, the element in D position is marked with the feature [+/ R] and the D slot may be filled with an article or a pronoun or may be left empty. Proper names and pronouns are object-referring and check the interpretable feature [+R]. In Italian, where this feature is strong, an object-referring noun checks it by overtly filling the empty D slot. That is, Italian has an empty as well as a lexical expletive determiner. In English [+R] is weak, banning overt N-to-D raising, and it, too, has an empty and a lexical expletive article.

Other nouns are denotative and check their [R] feature via one of a handful of strategies. If D is filled with a nonexpletive lexical article, the article has operator-like properties and the common noun provides a range for the operator. That is, the expression *the book* has the interpretation [D [N]] at LF: the *x* such that *x* belongs to the class of books. If D is filled with a lexical article that is expletive, as in the English *the beaver*, the noun checks its [R] interpretation and the structure is licensed with the generic interpretation. The function of the lexical expletive article is to block the default indefinite, so-called existential interpretation obtaining in lexically governed contexts, as in *I found [[e] fresh water/[e] fresh oranges]*. The empty expletive determiner precludes singular count nouns, as it precludes quantification over individuals. Rather, by default it selects the mass or plurality interpretation and can only be satisfied by a mass noun or a plural count noun.

In Italian, where the feature [+/ R] is strong, only proper nouns raise to D in the overt syntax. Common nouns do not raise in the overt syntax; rather, they check their [R] feature without overt movement. The empty D is "frozen" until LF and a nominal with an empty D is interpreted as existentially closed. This is why a nominal with a bare singular mass or a plural count noun can only be interpreted existentially in Italian. In

English, on the other hand, both [+R] and [R] are checked at LF. A common noun marked [R] can check its feature by raising into D, and the resulting structure has the generic interpretation, as in *Fresh water is often drinkable*. If a mass or plural noun does not raise at LF, [R] is checked if the empty D is interpreted as being in a CHAIN with the noun. In this scenario the noun in the foot of the CHAIN is interpreted, but the head of the CHAIN is not. The nominal is interpreted as existentially closed in this case, as in *I found fresh oranges*.

Longobardi's (1994) account of the interpretation of bare nouns crucially depends on the uninterpretability of the expletive article, which must be devoid of semantic content to be expletive. Expletives are invisible at LF. A true expletive has an associate, the noun if the expletive is the determiner. Chomsky (1995, 288) distinguishes two cases of expletives: pure expletives exemplified by *there* in English and impure expletives exemplified by *it* in English. The lexical entry for the expletive may in principle specify two forms: a lexical, overt expletive and a null, phonetically empty expletive. A pure expletive lacks any features apart from the strong categorial feature and attracts the associate, or rather its features, to move. Impure expletives have the relevant features, blocking associate raising. According to Chomsky (1995), a lexical expletive appears only when it is required for PF convergence, and in fact its content is added by phonological operations at PF. In this scenario, in the generically interpreted *the beaver* in (31) compared with the generically interpreted *beavers* in (32), that the definite article is lexical rather than empty if the noun is singular must be required for PF convergence, as the generic *the beaver* and *beavers* are the same objects at LF, with the noun's features filling the D slot in both cases.

(31) The beaver/*Beaver is a hard-working animal.

(32) Beavers/*The beavers are hard-working animals.

Then, it is unclear why the lexical expletive is not required for PF convergence, hence barred, with the generically interpreted plural count noun *beavers*.

In English *the beavers* cannot be interpreted generically, unlike in Italian, where a generic noun always requires a lexical expletive article. Conversely, the empty expletive into which a proper noun raises at LF can only be filled with a singular proper noun, allowing *Jones* and dis-allowing **Joneses*: hence the contrasts between *Jones* and **the Jones* and between **Joneses* and *the Joneses*. Furthermore, the empty expletive

strictly selects a mass interpretation for a singular common noun it combines with, as witnessed by the contrast between *I ate lion* and *I ate a lion* (see Longobardi 1994, 633). The lexical expletive, by contrast, is incompatible with (singular) mass nouns, as witnessed by the ungrammaticality of (33).

(33) *The fresh water is often drinkable.

The relevant contrasts are based on the concepts of number and mass interpretation, which are surely semantic concepts. Then, both the lexical and the null expletive articles seem to be inherently vested with semantic content, that is, with the features [+/ P1] and [+/ Count]. In conclusion, the empty article is not a pure expletive lacking any feature specification other than the categorial feature; rather, it is the impure expletive, which bans associate raising both in English and in Italian. This, together with the near complementary distribution of the lexical and the empty expletive articles, may help explain how an empty expletive article can be acquired by both English- and Italian-learning children.

In Polish not only proper names, pronouns, mass nouns, and plural count nouns are bare. Singular count nouns are bare as well, as shown in (34), the Polish analogue of (30).

(34) Spotkalem * przyjaciolke* Marii wczoraj.
met-lSG.PAST friend-A Maria-G yesterday
'I met a friend of Maria's/Maria's friend yesterday.'

If it were necessary to insert the empty article at LF in Polish in compliance with the restrictions on Select, endowed with operator force, the substantive article would restrict the reference of a nominal as specific definite or indefinite. In contrast to the empty expletive in English and Italian, the empty expletive would be a pure expletive in Polish, lacking any feature specification, as it could be satisfied both by singular countable and mass nouns as well as by plural nouns. Polish lacks a lexical analogue of the English expletive *there*, a lack that casts some doubt on the availability of a pure expletive article in the lexicon. Furthermore, if the pure expletive article must be erased at LF for convergence, it is hard to imagine how a child could ever learn that [KP/NP[*Mary*]] is an illegitimate object at LF but [DP[D *Mary*] [NP[N *t*]] is legitimate if *Mary* identifies the referent directly through its lexical content. Then, although the insertion of the empty substantive article might be necessary to determine the referential content of the common noun with specific reference, nominals

with proper nouns, pronouns, and possibly also common nouns with generic interpretation, which are names of kinds of things in Carlson 1977, would not require the empty article at LF; that is, they would be KPs in overt syntax and also at LF. Then, only referential nominals with a common noun would have to be represented as DP structures with an article at LF.

However, at least in two DP languages, the Bantu language Bemba and Basque, referential nominals reach LF as DP structures with a lexical article in D position, but nevertheless, a given DP is interpreted as definite or indefinite depending on its grammatical function (subject or object) and the type of predicate, as illustrated for Bemba in (35) (from Givón 1978, 301) and for Basque in (36) (from Laka 1993, 157).

(35) Umu-ana a-à-somene ici-tabo.
DET-child he-PAST-read DET-book
'The child read the/a book.'

(36) Emakume-a-k arto-a erein du.
woman-DET-ERG corn-DET planted has
'The woman has planted (the) corn.'

In other words, (in)definiteness of reference is not the function of the article in D in Bemba or Basque. On the basis of this fact, I want to suggest that it is not logically necessary that Polish nominals should have the DP layer at LF. 18 Rather, (in)definiteness of reference may be directly related to the properties of syntactic configurations in which nominals find themselves at LF in Polish, just as in Bemba and Basque.

Generally speaking, Polish encodes through word order the contrast between so-called predicational sentences, in which the identity of the referent of the theme of the utterance (i.e., its semantic subject) is pragmatically presupposed, and so-called presentational sentences, which assert the existence of the referent of a given expression in the universe of discourse. In predicational sentences the theme of the utterance is interpreted as definite or generic; and to be interpreted as indefinite, the theme may have to be introduced by an indefinite pronoun. Object shift is strongly correlated with definite interpretation. In so-called neutral word order (i.e., noncontrastive word order correlated with main sentence stress falling at the end of the sentence), a sentence-final bare nominal tends to be indefinite; and to be interpreted as definite, it has to be introduced by a demonstrative (see Szwedek 1976). Nevertheless, a definite description is possible sentence-finally as long as it introduces a new discourse referent.

In presentational sentences the theme is in sentence-final position and the indefinite interpretation is facilitated by subject inversion. Suppose that by default, kind-referring nominals are interpreted as indefinite at LF unless lexically marked as definite. 19 Then, to receive definite interpretation, they must be mapped onto Spec positions of appropriate functional heads correlated with definite interpretation (e.g., Ref(erential)P suggested by É. Kiss (1996) and possibly AgroP/nP). These Specs may be multiple or recursive to allow multiple elements with definite interpretation both sentence- and VP-initially. If this is correct, it is unnecessary to resort to D-insertion at LF in Polish to determine the reference of nominals.

10.4 Conclusion

In conclusion, there is no evidence to motivate the DP analysis of the structure of nominals in the overt syntax in Polish. The DP analysis of nominals at LF might have a simpler, more economical alternative in an approach deriving (in)definite interpretation from the properties of structural configurations encoding aspects of information structure in the utterances of the language. As a result of parametric variation in the inventory of the functional features of the noun, Polish nouns are under-specified for referentiality and (in)definiteness in the lexicon and project into KPs. Their (in)definiteness is not determined by the LF operation of the element in the D slot on the NP, its predicate; rather, it derives from the different relations that nominals hold at LF to certain functional heads correlated with the interpretation of referentiality and deixis. In other words, nominals become specified for (in)definiteness of reference as a result of overt and covert movement operations in Polish.

Notes

I had the good fortune to study syntax with Howard, if only for a short time. To me, his standard for scientific rigor will always be the standard to which all linguistic research should aspire. I thank him for sharing his insight into language and linguistic argumentation, for all the feedback and encouragement he gave me on my work, and for his sense of fair play.

This research was partly supported by a TEMPUS-PHARE individual mobility grant IMG-96-PL-2157 to the Grammatical Models Group at Tilburg University, The Netherlands. I wish to thank Elena Anagnostopoulou, Josef Bayer, Bob Borsley, Christine Erb, Henk van Riemsdijk, Craig Thiersch, Juan Uriagereka, and Jacek Witkos* for their help with language data, discussion, and helpful criticism. Usual disclaimers apply.

1. I will not be concerned here with quantifiers and numerals in Polish, a complex issue deserving a separate study. With respect to their morphological properties, they are adjective-like and manifest agreement with the head noun.

2. See, for example, Siloni 1991 for further discussion of the parametric variation in the featural composition of a phonologically unrealized D in several unrelated languages.

3. Linguists disagree on the number of genders in Polish. Klemensiewicz (1986) distinguishes three genders (M, F, and N); Szober (1963) talks of three genders in the singular (M, F, and N) and two genders in the plural (so-called M personal and non-M personal). Manczak * (1956) argues that Polish nouns should be analyzed into five gender classes (M personal, M animate, M inanimate, F, and N), and Corbett (1983) adds a sixth gender (M divirilized). The reason for these differences is that in the case form exhibiting most variation (i.e., the accusative), adjective and adjective-like modifiers take four distinct suffixes in the singular and two in the plural and that the agreements distinguished in the singular are not matched by the agreements distinguished in the plural. There are also differences of opinion concerning declinational classes in Polish.

4. Unlike attributive adjectives, classificatory adjectives add a semantic function to the noun (see Bosque and Picallo 1996). For example, *milk* in *milk chocolate* is a classificatory modifier. In Polish such adjectives follow the noun: for example, *czekolada mleczna* 'milk chocolate'.

5. This line of argumentation has been used in studies on Hebrew (see Ritter 1992, Siloni 1991), Catalan (see Picallo 1991), Spanish (see Bosque and Picallo 1996), and Italian (see Cinque 1995). In Polish, where the complement of a noun appears in the prepositionless, synthetic genitive, if the genitive Case feature must be formally licensed, the complement must raise to the Spec of a functional head. But because the noun surfaces before the complement, the noun must be raising twice in Polish.

6. Third person possessive pronouns are morphologically invariant. However, in order to maintain the integrity of the possessive pronoun paradigm, I will assume here that all possessive pronouns are possessive adjectives. Third person possessive adjectives are morphophonologically deficient, unlike 1st and 2nd person possessive adjectives.

7. The head N precedes ordinal adjectives in dynastic names and titles (e.g., *Mieszko Pierwszy* 'Mieszko the First') and classificatory adjectives. Crucially, however, if there is N-movement in such nominals at all, it is not into D, as the ordering facts are the same when the nominal contains an attributive adjective and/or a quantifier (e.g., *niedzwiedz* *polarny* 'polar bear', *kazdy mlody** *niedzwiedz polarny* 'each young polar bear').

8. If no material fills [Spec, NP] in (14) and (20), the Spec position need not be projected and movement proceeds directly through [Spec, DP] in both these cases. Then, however, it is [Spec, DP] rather than [Spec, NP] where genitive Case is licensed and identified.

9. Diesing (1992, 100) discusses an indefinite use of *this* in English, but provides no example of extraction out of NP with this determiner.

10. Extraction is also possible out of a nominal with a possessive pronoun, as shown in (i). Such movement is banned both in Giorgi and Longobardi 1991 and in Rizzi 1990.

(i) Jakiej książki$_i$ chciałabyś przeczytać [jego recenzję t_i]? what-G book-G want-2SG.COND (to) read his review-A
'*What booki would you like to read [his review of *t*i]?'

11. I owe (23) to Josef Bayer, who also informs me that a genitive argument cannot be extracted out of NP in German. This may tie in with the observation that prenominal genitives are restricted to proper and kinship nouns in German. Then, a *wh*-trace cannot identify the Spec-head agreement between D and the element in its Spec in German.

12. See also Rozwadowska 1995 for further examples and discussion.

> (22) is ambiguous between the event and the result reading in English, as English can pack the external argument into the '*s*-marked genitive in both types of nominals. In Catalan, on the other hand, (24) is a result nominal. Complex events use the analogue of the English *by* phrase to express the external argument in Catalan.

13. The unacceptability of (26d) and (26e) suggests that an argument cannot raise into [Spec, KP] to check genitive Case in Polish.

14. In Polish, like in Italian (see Giorgi and Longobardi 1991) but unlike in English (see Anderson 1983), Poss and Author may be linked to distinct positions and cooccur in the nominal complex. Thus, it seems that when Author is generated in the lexical projection of a noun as an aspect of the interpretation of the so-called R-argument of noneventive nouns, Poss can still be adjunct-adjoined to the nominal complex in Polish.

15. I do not claim that two genitive arguments are licensed in nominal structures in all DP languages. Rather, it seems that the availability of two nonprepositional genitives correlates with whether a language has a (morphological) means to distinguish between them.

16. Corver (1990) also suggests that Polish nominals lack the DP projection.

17. This is a simplification. Nominal expressions display an ambiguity between the referential and the nonreferential reading, both when introduced by the definite article and when introduced by the indefinite article, as discussed by Carlson (1977) and Givón (1978). According to Vergnaud and Zubizarreta (1992), only token-denoting (i.e., referential) nominals are denotative. In type-denoting nominals the determiner is expletive (i.e., has no denotative meaning).

18. To distinguish arguments from predicative nominals, it might be assumed by analogy with Chomsky's treatment of predicative adjectives (see Chomsky 1995, 283) that a predicative nominal is related to a predicative small clause with an open position. By contrast, an argument is a closed expression, even if it is not a DP projection at LF.

19. Object-referring nominals are interpreted as definite by virtue of their lexical content.

References

Abney, Steven. 1987. The English noun phrase in its sentential aspect. Doctoral dissertation, MIT, Cambridge, Mass.

Anderson, Mona. 1983. Prenominal genitive NPs. *The Linguistic Review* 3, 1 24.

Booij, Geert, and Jerzy Rubach. 1987. Postcyclic versus postlexical rules in lexical phonology. *Linguistic Inquiry* 18, 1 44.

Bosque, Ignacio, and Carme Picallo. 1996. Postnominal adjectives in Spanish DPs. *Journal of Linguistics* 32, 349 385.

Carlson, Gregory. 1977. A unified analysis of English bare plurals. *Linguistics and Philosophy* 1, 413 457.

Carstens, Vicki. 1993. Feature-types, DP-syntax and subject agreement. *GLOW Newsletter* 30, 20 21.

Chomsky, Noam. 1986. *Knowledge of language*. New York: Praeger.

Chomsky, Noam. 1995. Categories and transformations. In *The Minimalist Program*, 219 394. Cambridge, Mass.: MIT Press.

Cinque, Guglielmo. 1995. On the evidence for partial N-movement in the Romance DP. In Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi, and Raffaella Zanuttini, eds., *Paths towards Universal Grammar*, 85 110. Washington, D.C.: Georgetown University Press.

Clark, Eve V. 1978. Locationals: Existential, locative and possessive constructions. In Greenberg 1978, 84 126.

Corbett, Greville G. 1983. The number of genders in Polish. *Papers and Studies in Contrastive Linguistics* 16, 83 89.

Corver, Norbert. 1990. The syntax of left branch extractions. Doctoral dissertation, University of Tilburg, The Netherlands.

Diesing, Molly. 1992. *Indefinites*. Cambridge, Mass.: MIT Press.

Dikken, Marcel den. In press. (Anti-)agreement in Hungarian possessed nominal phrases. In István Kenesei, ed., *Crossing boundaries: Advances in Central and Eastern European languages*. Amsterdam: John Benjamins.

Giorgi, Alessandra, and Giuseppe Longobardi. 1991. *The syntax of noun phrases: Configuration, parameters and empty categories*. Cambridge: Cambridge University Press.

Giusti, Giuliana. 1994. Enclitic articles and double definiteness. *The Linguistic Review* 11, 241 255.

Givón, Talmy. 1978. Definiteness and referentiality. In Greenberg 1978, 292 330.

Greenberg, Joseph, ed. 1978. *Universals of human language*, vol. 4. Stanford, Calif.: Stanford University Press.

Grzegorczykowa, Renata, Roman Laskowski, and Henryk Wróbel, eds. 1984. *Gramatyka współczesnego jezyka * polskiego: Morfologia*. Warsaw: Pinstwowe* Wydawnictwo Naukowe.

Gussmann, Edmund, ed. 1995. *Licensing in syntax and morphology*. Lublin: Wydawnictwo Folium.

Haegeman, Liliane. 1994. *Introduction to Government and Binding Theory*. 2nd ed. Oxford: Blackwell.

Hoffman, Joel. 1998. Two types of paratactic word-order variation. Paper presented at the CLITE-1 Conference, Szeged, Hungary, April 1998.

Holmberg, A. 1993. On the structure of predicative NP. *Studia Linguistica* 47, 126 138.

É. Kiss, Katalin. 1996. Two subject positions in English. *The Linguistic Review* 13, 119 142.

Klemensiewicz, Zenon. 1986. *Podstawowe wiadomo ci z gramatyki jezyka* polskiego* (Essentials of Polish grammar). Panstwowe Wydawnictwo Naukowe.

Laka, Itziar. 1993. Unergatives that assign ergative. In Jonathan Bobaljik and Colin Phillips, eds., *Papers on Case and agreement I*, 150 172. (MIT Working Papers in Linguistics 18.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Lasnik, Howard. 1995. Case and expletives revisited: On Greed and other human failings. *Linguistic Inquiry* 26, 615 633.

Longobardi, Giuseppe. 1994. Reference and proper names. *Linguistic Inquiry* 25, 609 665.

Longobardi, Giuseppe. 1996. The syntax of N-raising: A minimalist theory. Ms., University of Venice.

Manczak,* Witold. 1956. Ile rodzajów jest w polskim? (How many genders are there in Polish?) Jezyk* *Polski* 36, 116 121.

Picallo, Carme. 1991. Nominals and nominalizations in Catalan. *Probus* 3, 279 316.

Ritter, Elizabeth. 1992. Cross-linguistic evidence for Number Phrase. *Canadian Journal of Linguistics* 37, 197 218.

Ritter, Elizabeth. 1993. Where's gender? *Linguistic Inquiry* 24, 795 803.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.

Rouveret, Alain. 1991. Functional categories and agreement. *The Linguistic Review* 8, 353 387.

Rozwadowska, Bozena.* 1995. The duality of *-nie/-cie* nominals. In Gussmann 1995, 85 106.

Siewierska, Anna. 1993. Syntactic weight vs. information structure and word order variation in Polish. *Journal of Linguistics* 29, 233 265.

Siloni, Tali. 1991. Noun raising and the structure of noun phrases. In Jonathan David Bobaljik and Tony Bures, eds., *Papers from the Third Student Conference*

on Linguistics, 255 270. (MIT Working Papers in Linguistics 14.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Speas, Margaret. 1991. Functional heads and inflectional morphemes. *The Linguistic Review* 8, 389 417.

Szober, Stanislaw. 1963. *Gramatyka jezyka * polskiego* (A grammar of Polish). Warsaw: Panstwowe* Wydawnictwo Naukowe.

Szwedek, Aleksander. 1976. *Word order, sentence stress and reference in English and Polish*. Edmonton, Canada: Linguistic Research, Inc.

Uriagereka, Juan. 1994. A conjecture on the form of parameters. In Jun Abe, Lisa Ferro, Laurel LaPorte-Grimes, Daiko Takahashi, and Miyuki Yamashina, eds., *Papers in honor of David Michaels*, 247 270. (UConn Working Papers in Linguistics 4.) Department of Linguistics, University of Connecticut, Storrs. [Distributed by MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.]

Vergnaud, Jean-Roger, and Maria Luisa Zubizarreta. 1992. The definite determiner and the inalienable constructions in French and in English. *Linguistic Inquiry* 23, 595 652.

Willim, Eva. 1995. In defence of the subject. In Gussmann 1995, 147 164.

Witkos,* Jacek. 1993. Some aspects of phrasal movement in English and Polish. Doctoral dissertation, UAM Poznan,* Poland.

Index

[Note: "Chomsky," "English," and "Minimalism" have not been indexed as these appear on roughly every page]

N