Cambridge Studies in Linguistics 143

The Minimalist Program

The Nature and Plausibility of Chomsky's Biolinguistics

Fahad Rashed Al-Mutairi

THE MINIMALIST PROGRAM

The development of the minimalist program (MP), Noam Chomsky's most recent generative model of linguistics, has been highly influential over the last twenty years. It has had significant implications not only for the conduct of linguistic analysis itself, but also for our understanding of the status of linguistics as a science.

The reflections and analyses in this book contain insights into the strengths and the weaknesses of the MP. Among these are: a clarification of the content of the strong minimalist thesis (SMT); a synthesis of Chomsky's linguistic and interdisciplinary discourses; and an analysis of the notion of optimal computation from conceptual, empirical, and philosophical perspectives.

This book will encourage graduate students and researchers in linguistics to reflect on the foundations of their discipline, and the interdisciplinary nature of the topics explored will appeal to those studying biolinguistics, neurolinguistics, the philosophy of language, and other related disciplines.

FAHAD RASHED AL-MUTAIRI is a PhD graduate from the University of Essex (UK). His academic interests are broad and include syntactic theory, the evolution of language, and the philosophy of science and mind.

CAMBRIDGE STUDIES IN LINGUISTICS

The Minimalist Program
The Nature and Plausibility of Chomsky's Biolinguistics

In this series

111. SAMUEL D. EPSTEIN AND T. DANIEL SEELY: *Derivations in Minimalism*
112. PAUL DE LACY: *Markedness: Reduction and Preservation in Phonology*
113. YEHUDA N. FALK: *Subjects and their Properties*
114. P. H. MATTHEWS: *Syntactic Relations: a Critical Survey*
115. MARK C. BAKER: *The Syntax of Agreement and Concord*
116. GILLIAN CATRIONA RAMCHAND: *Verb Meaning and the Lexicon: a First Phase Syntax*
117. PIETER MUYSKEN: *Functional Categories*
118. JUAN URIAGEREKA: *Syntactic Anchors: on Semantic Structuring*
119. D. ROBERT LADD: *Intonational Phonology* (Second Edition)
120. LEONARD H. BABBY: *The Syntax of Argument Structure*
121. B. ELAN DRESHER: *The Contrastive Hierarchy in Phonology*
122. DAVID ADGER, DANIEL HARBOUR AND LAUREL J. WATKINS: *Mirrors and Microparameters: Phrase Structure Beyond Free Word Order*
123. NIINA NING ZHANG: *Coordination in Syntax*
124. NEIL SMITH: *Acquiring Phonology*
125. NINA TOPINTZI: *Onsets: Suprasegmental and Prosodic Behaviour*
126. CEDRIC BOECKX, NORBERT HORNSTEIN AND JAIRO NUNES: *Control as Movement*
127. MICHAEL ISRAEL: *The Grammar of Polarity: Pragmatics, Sensitivity, and the Logic of Scales*
128. M. RITA MANZINI AND LEONARDO M. SAVOIA: *Grammatical Categories: Variation in Romance Languages*
129. BARBARA CITKO: *Symmetry in Syntax: Merge, Move and Labels*
130. RACHEL WALKER: *Vowel Patterns in Language*
131. MARY DALRYMPLE AND IRINA NIKOLAEVA: *Objects and Information Structure*
132. JERROLD M. SADOCK: *The Modular Architecture of Grammar*
133. DUNSTAN BROWN AND ANDREW HIPPISLEY: *Network Morphology: a Defaults-based Theory of Word Structure*
134. BETTELOU LOS, CORRIEN BLOM, GEERT BOOIJ, MARION ELENBAAS AND ANS VAN KEMENADE: *Morphosyntactic Change: a Comparative Study of Particles and Prefixes*
135. STEPHEN CRAIN: *The Emergence of Meaning*
136. HUBERT HAIDER: *Symmetry Breaking in Syntax*
137. JOSÉ A. CAMACHO: *Null Subjects*
138. GREGORY STUMP AND RAPHAEL A. FINKEL: *Morphological Typology: From Word to Paradigm*
139. BRUCE TESAR: *Output-driven Phonology: Theory and Learning*
140. ASIER ALCÁZAR AND MARIO SALTARELLI: *The Syntax of Imperatives*
141. MISHA BECKER: *The Acquisition of Syntactic Structure: Animacy and Thematic Alignment*
142. MARTINA WILTSCHKO: *The Universal Structure of Categories: Towards a Formal Typology*
143. FAHAD RASHED AL-MUTAIRI: *The Minimalist Program: The Nature and Plausibility of Chomsky's Biolinguistics*

Earlier issues not listed are also available

THE MINIMALIST PROGRAM

The Nature and Plausibility of Chomsky's Biolinguistics

FAHAD RASHED AL-MUTAIRI

The University of Essex (UK)

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107041349

First published 2014

Printed in the United Kingdom by Clays, St Ives plc

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data
Al-Mutairi, Fahad Rashed, 1973–
The Minimalist program : the nature and plausibility of Chomsky's biolinguistics / Fahad Rashed Al-Mutairi.
pages cm – (Cambridge studies in linguistics ; 143)
ISBN 978-1-107-04134-9 (hardback)
1. Minimalist theory (Linguistics) 2. Biolinguistics. 3. Generative grammar. I. Title.
P158.28.A5 2014
415′.0182–dc23

2014016831

ISBN 978-1-107-04134-9 Hardback

To the memory of my mother, Muneera, who, despite never having had the opportunity to learn how to read and write, worked very hard and sacrificed so much of her life to make sure all of her twelve children received the best education possible.

Contents

Acknowledgements *page* xi
List of abbreviations xii

1 Introduction **1**

2 The minimalist program **6**
2.1 Introduction 6
2.2 Chomskyan linguistics: refutation of some misconceptions 6
2.2.1 The combinatoric stage 7
2.2.2 The cognitive stage 10
2.2.3 The minimalist stage 18
2.2.4 Concluding remarks 21
2.3 The shift to minimalism 22
2.4 UG: from an *explanans* to an *explanandum* 31
2.5 Minimalism and the design of language 35
2.5.1 The model of grammar: from GB to minimalism 36
2.5.2 Merge, Move, and Agree 38
2.5.3 Economy 41
2.6 Why minimalism? 43

3 The strong minimalist thesis (SMT) **51**
3.1 Introduction 51
3.2 Conceptual necessity: a first encounter 52
3.3 SMT as a strict generalization 55
3.4 The imperfection strategy 61
3.5 Conceptual necessity: a second encounter 65
3.6 Some shortcomings 69
3.7 The three factors framework 73
3.8 Two discourses, one thesis? 78

4	**The SMT in an evolutionary context**	**82**
4.1	Introduction	82
4.2	The recursion-only hypothesis	83
4.3	The content of FLN: two extreme views	92
4.4	The where and how of recursion	97
4.5	The Merge-only hypothesis	102
5	**The SMT as an explanatory thesis**	**109**
5.1	Introduction	109
5.2	Minimalist explanation: interface conditions	109
5.3	Tautology	113
5.4	Teleology	118
5.5	Minimalist explanation: optimal computation	121
5.6	Loose correlations	126
5.7	Minimalism and teleological physics	133
6	**Optimal computation and multiple realization**	**143**
6.1	Introduction	143
6.2	Chomskyan naturalism	144
6.3	Optimal computation and non-genomic nativism	147
6.4	Functionalism	152
6.5	The multiple realization argument	158
6.6	Functionalism and naturalism: uneasy bed partners	161
6.7	Chomsky's case against functionalism	164
6.8	Optimal computation versus multiple realization	170
6.9	Implications for the biolinguistic approach	173
7	**Conclusion**	**178**
	Notes	185
	References	197
	Index	210

Acknowledgements

My interest in the study of language in general, and in Chomsky's work in particular, began some fifteen years ago when, as an undergraduate student at the University of Salamanca (Spain), I came under the happy influence of two of my professors, Emilio Prieto de los Mozos and José Jesús Gómez Asencio. That interest culminated in the present book, which is a revised and extended version of my 2011 Essex University doctoral thesis, "The Strong Minimalist Thesis: its Nature and Plausibility."

At Essex I was truly honoured and fortunate to have had Martin Atkinson as my supervisor, whose superb lectures on minimalism have been the source of my inspiration for this work. Martin was also a valued friend whose encouragement and support have never wavered, and for which I am very grateful. My thanks also go to Noam Chomsky for taking time out of a very busy schedule to answer my many inquiries about his work, and to my thesis examiners, Neil Smith and Roger Hawkins, for their much valued encouragement and guidance.

I must also thank Cambridge University Press for publishing this book. In particular, I am indebted to Helen Barton for her faith and interest in this project, and for her support and advice throughout the publishing process. Thanks are also due to several anonymous referees for their insightful comments and useful suggestions on the manuscript. Of course, whatever shortcomings remain in this book are all my responsibility.

In addition to my family, I am grateful to my friends Bader Al-Dehani and Pedro Arias for their continuous support and encouragement, and to my colleague, Abbas Al-Haddad, for generously providing a peaceful environment in which to work during the academic year 2012–13. My thanks also go to two colleagues at Essex, Antonis Polentas and Ibrahim Al-Mahboob, for the many discussions I had with them over the years.

Above all, I am most heartily indebted to my wife, Katarina, for tolerating an often-distracted husband, and for her never-ending love and support throughout the process of writing this book. Without her enormous help and many sacrifices I could not have possibly finished this work.

Abbreviations

B&H	Boeckx and Hornstein (2010)
C	complementizer
C_{HL}	computational system of human language
CI	conceptual-intentional (system)
cp-	*ceteris paribus* (clause)
CP	complementizer phrase
D	determiner
D-	deep-(structure)
DP	determiner phrase
FI	full interpretation
FL	faculty of language
FLB	faculty of language in the broad sense
FLN	faculty of language in the narrow sense
GB	government and binding (theory)
LF	logical form
MLC	minimal link condition
MP	minimalist program
MRA	multiple realization argument
N	noun
NP	noun phrase
PF	phonological form
PLD	primary linguistic data
P&P	principles and parameters (framework)
S	sentence
S-	surface-(structure)
SL	statistical learning
SM	sensory-motor (system)
SMT	strong minimalist thesis
Spec	specifier
T	tense

TP	tense phrase
UG	universal grammar
V	verb
v	light causative verb
VCN	virtual conceptual necessity
VLSI	very large-scale integrated (microcircuit)
VP	verb phrase

1 *Introduction*

Linguistics is an old subject, and since 1957 it has been a great one.[1] In that year *Syntactic Structures* was published, initiating officially what we would now call "Chomskyan linguistics." The enormous and successive stimuli to the discipline due to Chomsky's work over the past fifty years have significantly changed the landscape of linguistics, and its vast influence is palpable in the mainstream of linguistic thinking and research. The most recent book-length manifestation of these developments is *The Minimalist Program* (Chomsky 1995a).

The Minimalist Program (henceforth, MP) can be characterized from a number of perspectives which give it an interesting, and potentially far-reaching, interdisciplinary character. It can be considered simply as a linguistic framework involving a substantial revision of many of the technical assumptions and theoretical proposals which have been developed within the Chomskyan paradigm prior to the early 1990s. It can also be viewed as an extension and reconstruction of the biolinguistic approach to language which was initiated by Chomsky (1955, 1965) and Lenneberg (1967), an approach which views language as a biological capacity rooted in evolution.[2] Moreover, from ontological and methodological perspectives, the MP embraces a "naturalistic" approach to language as a "mental organ" of the brain, an approach based on the assumption that the mind is part of the natural world and, as such, it should be studied in the same way as any other aspect of nature (cf., for instance, Chomsky 2000a: 75).

These perspectives are closely related in that the substantial revision of pre-minimalist assumptions and theories is committed to the goal of achieving a principled explanation of linguistic phenomena, an explanation that is intended to go beyond the sphere of influence of genetic endowment to general principles which relate not just to language, but to general cognition or to the natural world as a whole. The standard for this explanation is set by the central thesis of the MP, the so-called Strong Minimalist Thesis (SMT), which suggests that language is "well designed" to satisfy certain legibility conditions for its interaction with other cognitive systems (cf. Chomsky 2001: 2).

In fact, proponents of the minimalist program believe that there is even more to the SMT than just setting a standard for true explanation in linguistics. It has been argued, for instance, that if this thesis is on the right track, one should expect a number of positive consequences, not just for linguistics, but for science in general. Two of these stand out. One concerns the study of the evolution of language, and the other relates to the prospects of unification in science. As to the former, Chomsky (2007b: 2–3) argues that "the less attributed to genetic information . . . for determining the development of an organism, the more feasible the study of its evolution." Similarly, though in a somewhat different context, Hauser *et al*. (2002) and Fitch *et al*. (2005) suggest that if most of the properties of language can be found in species other than humans, then a comparative empirical approach to the evolution of language becomes more feasible. In regard to the prospects of unification in science, some minimalists believe that if language turns out to indeed be "well-designed" or "optimal" in the sense of the SMT, that is, if aspects of the general principles that determine the "design" of language can be regarded as direct consequences of the workings of the laws of physics, then this outcome should be celebrated as a step towards this goal (cf., for instance, Boeckx and Piattelli-Palmarini 2005).

Thus, if the SMT turns out to be correct, it will have implications not only for the conduct of linguistic analysis itself, but also for our understanding of the place of language in the world. In fact, and as I hope to show in this book, different, but no less significant, implications will follow if the SMT turns out to be false in specific respects; minimalism promises to yield significant implications whether its central claim turns out to be true or false, and this, I take it, is the definition of a good research program. From such perspectives, it is hard to see how any linguist can fail to be interested in it.

The core questions that this interest raises, and that this book seeks to answer, are the following: (1) what is the nature of the transition to Minimalism? (2) How should the SMT be interpreted? (3) How plausible is the SMT from an evolutionary perspective? (4) To what extent does the SMT provide an appropriate standard for true explanation of linguistic phenomena? (5) Are there, as some minimalists would have us believe, genuine connections between the principles of language and the laws of physics? (6) Is the "bio" in "biolinguistics" really significant or does it merely reflect an implicit belief that the scientific merit of linguistics is proportional to the strength of its relation with the more advanced sciences?

In attempting to answer question (1), I have found it necessary to first take a broad view of the general development of Chomskyan linguistics, with the primary aim of clarifying some of the misconceptions that have been

expressed by (ironically enough) some well-known popularisers of Chomsky's work. For only when these misconceptions are dispelled will we be in a better position to understand the nature of the shift to minimalism. This shift, as I argue in this book, is neither a matter of taking account of methodological considerations, as some advocates of the MP seem to believe, nor is it as dramatic and unexpected as some critics have suggested. By examining the defining features of the pre-minimalistic conception of language, and by identifying their fate in the context of the MP, I seek to demonstrate that the shift to minimalism is merely one of emphasis among the factors that govern the nature of the language faculty.

The answer to question (2) calls for an exploration of the nature and content of the SMT, and this is one of the two main tasks that this book undertakes. I argue that Chomsky's work over the past fifteen years suggests three different formulations of the SMT, some of which seem to be incompatible with his own views on language. I also show that by making clear the differences between the three formulations of the SMT, and by submitting the phrase "virtual conceptual necessity" to critical analysis and examination, it is possible to avoid much of the confusion surrounding its interpretation. Each of the three formulations of the SMT will be discussed in turn, and after pointing out the shortcomings of the first two, I argue that the last formulation provides the most transparent reflection of the content of the thesis. This content involves two distinct claims and the evaluation of each of these constitutes the other main task of the book.

The first claim is an ontological one. It asserts that universal grammar contains nothing beyond the combinatorial operation Merge; i.e. the genetic component of the language faculty is confined to this recursive operation. This is what I will call the merge-only hypothesis, and to evaluate it is to evaluate the SMT from an evolutionary perspective and thereby provide an answer to question (3). Before this can be done, however, it is necessary to contrast Chomsky's work with his contributions to Hauser *et al*. (2002) and Fitch *et al*. (2005). For it is my contention that a careful analysis of the similarities and differences between Chomsky's linguistic and interdisciplinary discourses should caution us against an assumption that is widespread in the literature, namely that the notion of "recursion" as employed by Hauser *et al*. (2002) is identical to "merge" in the minimalist vocabulary. Through such an analysis I develop a conceptual and empirical assessment of the merge-only hypothesis, and I clarify its relation to the recursion-only hypothesis of Hauser *et al*., concluding that the two hypotheses are not equivalent and have different empirical content. I also argue that not only are there conceptual and empirical difficulties surrounding the merge-only hypothesis, but that there is also an

uncomfortable ambiguity in Chomsky's position regarding the ontological status of merge, an ambiguity that is not easy to resolve.

The second claim which the SMT involves is epistemological. As mentioned above, this minimalist thesis is intended to set a standard for true explanation of linguistic phenomena, and through its (implicit) connection to the notion of "physical law" it promises a better understanding of the place of language in the world. This brings to the fore questions (4) and (5): does the SMT offer a principled explanation of linguistic phenomena? Are there non-trivial connections between the principles of language and the laws of physics? As will be seen through this book, my own answer to both questions is negative, and while I suggest a way to ameliorate certain aspects of minimalist explanation, I make it clear that some of the attempts which have been made to ground optimal computation in physical law do more harm to the MP than good. I also consider the explanatory status of the kind of physics which some minimalists take to be germane to the MP, and I demonstrate that it is no longer acceptable in modern physics.

Of course, it is one thing to say that there is little empirical support for the deduction of optimal computation from physical "neatness" and another to say that such a deduction is untenable in principle. Indeed, an attitude of "let's wait and see" is sometimes expressed by advocates of the biolinguistic approach to language in defence of their speculations on the connections between linguistics and physics (see, among others, Freidin and Vergnaud 2001; Uriagereka 2002). In fact, Chomsky himself expresses this same attitude in his criticism of Fodorian functionalism, the philosophical doctrine that asserts, among other things, that the level of abstraction at which our explanatory theories of the mental are made is principled. He suggests that this level of abstraction should be conceived of as a "temporary convenience" which may not resist further examination at a more fundamental level, say that of neurology (Chomsky in Cela-Conde and Marty 1998). Clearly, this optimism on the prospects of unifying cognitive science with brain science underlies Chomsky's biolinguistics and its quest for a principled explanation in linguistics. But is this unification really necessary for linguistic theory or does it merely reflect Chomsky's (2000a: 77) belief that "the place to look for answers is where they are likely to be found: in the hard sciences"? This is another way of expressing the remaining question on the list above: is the "bio" in "biolinguistics" really significant or does it merely reflect an implicit belief that the scientific merit of linguistics is proportional to the strength of its relation with the more advanced sciences? As a necessary step on the way towards an answer to this question, I assess the explanatory status of optimal computation in the

context of the philosophy of mind. By bringing Chomsky's naturalism face-to-face with Fodorian functionalism I seek to pin down some tensions that arise between the two. I argue that some of these have significant implications for the explanatory role of optimal computation in particular, and for the status of the biolinguistic approach to language in general, leading to a revised version of minimalism in which optimal computation plays its familiar role but is now regarded as primitive rather than triggering a search for a deeper level of explanation.

In short, this book explores Chomsky's biolinguistics in general and its fundamental thesis in particular. I seek to shed some light on the content of the SMT and evaluate it from a number of perspectives. In this endeavour, I identify gaps in current minimalist theorizing and, whenever possible, I search for ways to fill such gaps.

Unlike many who are excited by the prospects opened up by taking minimalism seriously, I adopt an open-minded view on its tenets and their formulation; I do not hesitate to be critical when criticism is necessary, especially when unsubstantiated claims are apt to lead either to erroneous conclusions or pretentious proclamations. To use the words of Whitehead (1997 [1925]: 18): "If science is not to degenerate into a medley of *ad hoc* hypotheses, it must become philosophical and must enter upon a thorough criticism of its own foundations." Indeed, given the fact that minimalism involves a variety of theoretical issues from a wide range of scientific as well as philosophical inquiries, this book is in a sense philosophical in that it encourages a critical examination of the very foundations of the MP and its relation to other fields of inquiry.

2 *The minimalist program*

2.1 Introduction

This chapter takes a broad view of the minimalist program (MP), providing background for the chapters that follow. It has five main sections. Section 2.2 concerns the historical development of the generative program from its inception in the early 1950s to the emergence of minimalism in the 1990s. It follows the structure of the useful review of Boeckx and Hornstein (2010), while maintaining that this review is misleading in important respects. Section 2.3 is an attempt to uncover the nature of the shift to minimalism and to explain how this program differs from its predecessors. A description of how this shift affected the theoretical role of universal grammar (UG) is given in Section 2.4, while in Section 2.5 an illustration of the impact of minimalism on the question of the design of language is presented. Finally, Section 2.6 asks a simple question, "Why minimalism?", and attempts to provide a tentative answer to it. This last section also prepares the ground for the discussion of the strong minimalist thesis in the next chapter.

2.2 Chomskyan linguistics: refutation of some misconceptions

When asked whether the history of his work on linguistics is misconceived, Chomsky (p.c.) replied by saying that "[t]he history of [generative grammar] is hopelessly misconceived, sometimes ludicrously so," and he referred, as an example, to overtly hostile critics such as Boden (2006).[1] However, as the present section purports to show, there seems to be no reason to believe that certified members of the Chomskyan school are immune from historical misconceptions, albeit of a different nature to those displayed by Boden. A case in point is Boeckx and Hornstein's (2010) goal-directed approach to the development of Chomsky's work on linguistic theory.

The authors (henceforth B&H) advocate a three-period distinction in connection with the generative enterprise. They attribute to each historical period

its own theoretical goal, core text(s), and related field of inquiry. The first period is called the *combinatoric stage*, the primary goal of which is discovering "the right axioms." Its core text is Chomsky (1957), and its related fields of inquiry are engineering and logic. The second period is the *cognitive stage*, where the chief aim is solving the problem of how language is acquired, i.e. so-called Plato's problem (Chomsky 1986). The core texts are Chomsky (1965) and Chomsky (1981), and the related field is biology. Finally, the third period is the *minimalist stage*, concerned with finding the *best* solution to Plato's problem. Chomsky (1995a) is its core text, and physics is its related field of study.

2.2.1 The combinatoric stage

As observed, B&H relate the early period of Chomsky's linguistic theory to the fields of engineering and logic. It is not clear how Chomsky's work on linguistic theory bears on engineering, and since B&H do not provide even a hint as to how this might be the case, I shall not seek to evaluate this claim. The case of logic as a field of inquiry, however, is different, for here we are given an explicit analogy between generative grammars and logical systems, and we should therefore be able to assess the proposal that Chomsky's early work on syntactic theory can be seen as related to work in logic.

To begin with, there is no reason to deny that early work on generative grammar was influenced, to a certain extent, by modern logic. For instance, Chomsky's "rules of formation" and "rules of transformation" are two expressions which were adopted from Carnap (1937), and Chomsky (1965: 9) himself acknowledges the apparent similarity between his phrase structure rules and Post's (1943) production rules. Moreover, it is perfectly legitimate to draw a tentative analogy between early generative grammar and logical systems, where the notions "initial string," "rewriting rule," and "grammaticality" in a generative grammar might be associated with their respective counterparts "axiom," "inference rule," and "theoremhood" in a logical system.[2] Indeed, in his description of a phrase structure grammar of the form [Σ, F], where Σ denotes the set of initial strings and F represents the set of rewriting rules of the form $X \rightarrow Y$, Chomsky (1956: 117) considers a derivation in such a grammar as "roughly analogous to a proof, with Σ taken as the axiom system and F as the rules of inference." However, we should be careful not to stretch the analogy too far. B&H (p. 120) appear to do this when they identify the goal of linguistic theory in early generative grammar as "finding the right axioms." In particular, the authors claim that this alleged goal of linguistic theory parallels that of theories in logic – "to find a set of axioms from which it was possible to derive

all and only the valid inferences" (p. 119). The analogy is far from convincing, however. To see why, let us examine briefly the notion of "axiom" and how it might be related to early linguistic theory.

There are at least two issues to consider when speaking about axioms. First, given a formal deductive system, the axioms of the system represent a *starting point* from which certain statements or theorems can be derived by rules of inference. It is by virtue of this property that the tentative analogy made earlier between phrase structure grammars and logical systems may be justified, for the string S in a phrase structure grammar is considered as the starting point from which all subsequent strings are derived by formation rules. To give a simple example, a phrase structure grammar G can be formally defined as a quadruple {N, Σ, S, P}, where N, Σ are two finite and disjoint sets of non-terminal and terminal symbols, respectively, S ∈ N, and P is a finite set of production rules of the form X → Y, where X and Y are strings of symbols from N ∪ Σ.[3] Without going into all the details of the grammar, it should be clear how the sentence in (1) can be derived using such a grammar from the start symbol S by repeated applications of the rules in (2) as shown in (3):

(1) The postman rang the bell.

(2) Rule 1: S → NP + VP
Rule 2: NP → D + N
Rule 3: VP → V + NP
Rule 4: D → the
Rule 5: N → postman, bell, etc.
Rule 6: V → rang, saw, etc.

(3)

S	
NP + VP	[by rule 1]
D + N + VP	[by rule 2]
D + N + V + NP	[by rule 3]
the + N + V + NP	[by rule 4]
the + postman + V + NP	[by rule 5]
the + postman + rang + NP	[by rule 6]
the + postman + rang + D + N	[by rule 2]
the + postman + rang + the + N	[by rule 4]
the + postman + rang + the + bell	[by rule 5]

As should be clear from this example, the initial string S represents the point from which the derivation of the sentence in (1) begins, and in this sense S may be said to be comparable to axioms in a logical calculus. However, there is more to axioms than just initiating a set of derivations, which brings us to the second important property of this logical notion, namely that an axiom is standardly

viewed as expressing a *proposition* that has, or can be assigned, a *truth value*. It is by virtue of this property that axioms in logic differ from initial strings in phrase structure grammars. Notice that this is a non-trivial distinction, for the notion of "finding the right axioms" can only be meaningful insofar as an axiom has a propositional content. Thus, given an axiomatic system and a domain of objects, it is reasonable to inquire as to whether the chosen axioms are true of these objects and whether they provide a basis for deriving *any* true proposition in the domain. These inquiries about the truth of axioms are neither here nor there when it comes to the derivation in (3), and this is because the initial strings are simply not propositions. Thus, it makes no sense to ask whether the initial string "S" is true. Consequently, and contrary to what B&H believe, there never was a notion analogous to "finding the right axioms" in generative grammar, nor could there have been.

As already noted, Chomsky (1956) draws a tentative analogy between his theory of generative grammar and proof theory, an analogy also raised in Chomsky (1955: 729) when he says: "A derivation is roughly analogous to a proof, with Sentence playing the role of the single axiom, and the conversions corresponding roughly to rules of inference" (underlining in original). Thus if S is considered to be "the single axiom" as is the case in (3) above, then it would be absurd to suggest, as B&H do, that the goal of linguistic theory is to "find the right axioms," for there is only one axiom and it is known beforehand. In short, "finding the right axioms" was never an issue for Chomsky.

B&H identify the aim of Chomsky's early work with an alleged computational goal. They say:

> The primary aim is computational or combinatoric [footnote omitted]. The problem is framed by two observations. First, the set of well-formed sentences of a natural language is infinite. Second, natural language sentences naturally partition into two sets: the well-formed and the ill-formed. Given this partition into two infinite sets the grammarian's goal is to characterize them by *finding a set of rules (a grammar) that will generate all the well-formed sentences and not generate any ill-formed ones*. If successful, such grammars would constitute comprehensive theories of language comparable to the kinds of theories that chemists and biologists construct in their respective areas (this sentiment is especially made explicit in Lees' 1957 review of *Syntactic Structures*). (B&H, p. 116, my italics)

Setting aside this unwarranted movement from "finding the right axioms" to "finding the right rules," which may be taken as indicative of more confusion on

the part of the authors, it is important to remember that Chomsky (1955, 1957) was well aware of the fact that two or more sets of rules (i.e. grammars) may be compatible with whatever data are available, and it was precisely because of this that he introduced the notion of "evaluation measure" as a special case of relating particular grammars to the general theory of linguistic structure. Thus the task facing the linguist cannot be merely computational, for it is not limited to finding an adequate grammar. Rather, the task facing the linguist is understood to be ultimately explanatory. This is what Chomsky (1955: 11) calls "the problem of justification," that is, "the problem of choosing among the vast number of different grammars, each giving a different structure, and all meeting ... external criteria."

More importantly, and contrary to what B&H assert in the passage above, the goal of generating all and only the grammatical sentences was merely a *descriptive* one, and meeting it in no way leads to "comprehensive theories of language," at least not in Chomsky's sense.[4] Since B&H cite Lees' famous review of Chomsky (1957) favorably, it is instructive to note that in that review the generation of all and only grammatical sentences is held to be a descriptive requirement, for Lees (1957: 382) maintains that a grammar "must permit us to generate automatically all and only the grammatical sentences of the language, else it could not be called a description at all." Now, it is clear that B&H's views on the early period of Chomskyan linguistics involve at least one problematic implication. For if the primary goal of linguistic theory were essentially computational, in the sense of generating all and only the grammatical sentences of a language, and if this goal amounts to an adequate description of the object of inquiry as Lees seems to suggest, it follows that the primary goal of Chomsky's *Syntactic Structures*, a text which, no doubt, B&H regard as revolutionary (and rightly so), was merely descriptive in nature. But nothing can be further from the truth. For the requirement of separating grammatical sequences from ungrammatical ones constitutes only a first step towards what Chomsky was trying to achieve in the mid-fifties, namely an explanation for the linguistic intuition of native speakers. In fact, B&H do not seem to appreciate the role of cognition in the early writings of Chomsky, for otherwise they would hardly have divided Chomsky's pre-minimalist conjectures into combinatoric and cognitive stages. To show that this is indeed the case, let us now turn to their second stage.

2.2.2 The cognitive stage

The alleged transition from a combinatoric stage to a cognitive stage is described by B&H as a shift "from finding the right axioms ... to solving

what came to be known as Plato's problem" (p. 120). But how do the authors justify the existence of such a shift? Their rationale (p. 122) runs something like this: the aim of the combinatoric stage was to find a set of rules from which all and only the grammatical sentences of a language may be derived. This aim, they suppose, arose in the context of an underlying assumption, namely that native speakers had direct access to the underlying grammatical rules of their language. However, with the advent of the cognitive stage, this assumption was revised in terms of a distinction between *grammaticality* and *acceptability*. Such a distinction affected the central aim of linguistic theory, for now "grammaticality" denotes a theoretical notion underlying the speaker's intuitive ability to judge whether or not a given sentence is acceptable. Consequently, according to the authors, the combinatoric aim of generating all and only the grammatical sentences of a language can now be conceived, in the cognitive stage, as providing a description of what needs to be explained, namely the underlying cognitive ability of native speakers to know and acquire their language.

Now, we have already stressed that the requirement on a grammar to generate all and only the grammatical sentences of a language was viewed by Chomsky as a means, rather than as an end in itself. Interestingly, B&H (p. 122) concede this, but they suggest that it was only with the advent of the cognitive stage that this requirement came to be considered as a first step towards providing an explanation of linguistic competence. Clearly, however, the record is not on their side, as the following passage from Chomsky (1955: 39–40) demonstrates:

> [A] speaker has an "intuitive sense of grammaticalness." But to say this is simply to state the problem. Suppose that we (i) construct an abstract linguistic theory in which grammaticalness is defined, (ii) apply this linguistic theory in a rigorous way to a finite sample of linguistic behavior thus generating a set of "grammatical" sentences, and (iii) demonstrate that the set of grammatical sentences thus generated, in the case of language after language, corresponds to the "intuitive sense of grammaticalness" of the native speaker. In this case, we will have succeeded in giving a rational and general account of this behavior, i.e., a theory of the speaker's linguistic intuition. This is the goal of linguistic theory.

Thus, right from the outset, the goal of linguistic theory was far from simply a matter of finding a set of rules that would generate all and only the grammatical sentences of a language; rather, the goal was to show that such a set of rules (1) follows from a general theory of linguistic structure and (2) reflects

the native speaker's "intuitive sense of grammaticalness."[5] Of course, Chomsky (1955, 1957) recognizes the difficulty of defining the notion of "grammaticalness" in terms of a general theory of linguistic structure, and in response to this he adopts a strategy of assuming a partial intuitive knowledge of the concept, i.e., there are certain clear cases in which native speakers can reliably decide whether or not a given sentence is grammatical. In unclear cases, Chomsky (1957: 14) says that "we shall be prepared to let the grammar itself decide." That is, lacking a general theory of linguistic structure, Chomsky thought it necessary to bypass the difficulty in formulating a criterion for "grammaticality" by relying on the fact that speakers can provide reliable data over a certain range concerning which sentences are grammatical and which are not. As we will see shortly, this was considered a necessary step towards the construction of a general theory of linguistic form.

B&H seem to imply that the distinction between "grammaticalness" and "acceptability" was an innovation introduced in Chomsky (1965). Now, it is true that in Chomsky (1965) (henceforth *Aspects*), we find an explicit distinction between "grammaticalness" and "acceptability," with the former notion belonging to native speakers' tacit knowledge of their language (i.e. competence) and the latter belonging to the actual use of language (i.e. performance). But are B&H correct in implying that such a distinction played no role before the publication of *Aspects*? An affirmative answer to this question seems warranted when we consider the following statement from Chomsky (1957: 13): "One way to test the adequacy of a grammar ... is to determine whether or not the sequences that it generates are actually grammatical, i.e., acceptable to a native speaker." At first sight, this seems to indicate that, in contrast to what we find in *Aspects*, Chomsky did not distinguish between "grammatical" and "acceptable." However, further inspection reveals otherwise. For instance, consider Chomsky's strategy referred to in the previous paragraph. In order to avoid the unreliability of some judgements involved in defining the notion of "grammaticalness," the strategy starts by acknowledging that there are clear cases in which native speakers' judgement about a sentence's grammaticality is reliable enough for the construction of a grammar. It is in this context that Chomsky's statement quoted above is to be understood; i.e. the notion of "grammatical" may be equated with "acceptable" *in certain clear cases*. Now, the important point that B&H have missed is this: the reliance on the speaker's intuition to determine "grammaticality" constitutes a necessary step towards the construction of a general theory of linguistic structure, but as soon as this theory is constructed, its success will be determined by providing (1) an explanation

of the native speaker's intuitions, and (2) a formal definition of the notion of "grammaticality." Chomsky makes this point clear when he says:

> We begin by recognizing the existence of an intuition ... We end, if successful, by giving an objective theory which, in a certain sense, explains this intuition. Before linguistic theory is constructed, the subject matter for description is determined ... by reference to the speaker's intuitions about which forms are grammatical ... After a successful theory has been constructed, the subject matter for linguistic description is determined by the theory itself. (Chomsky 1955: 40)

B&H (p. 122) misrepresent what Chomsky's theory is about when they claim that, in the combinatoric stage, it was tacitly assumed that native speakers can judge a sentence's grammaticality, and that only with the advent of the cognitive stage was this tacit assumption deemed to be unfounded. There is no such assumption, tacit or otherwise – indeed, that native speakers can, in certain cases, judge a sentence's grammaticality is a fact, not an assumption. This, according to Chomsky (1955: 93–4), makes linguistic intuition both (1) a convenient tool for the investigation of linguistic structure, and (2) an *explanandum* for which a theory is needed.

Moreover, B&H (ibid.) assert that, in the cognitive stage, the notion of "grammaticality" has come to be conceived as "a theoretical assessment made by the linguist, not an eye-witness report that a speaker makes by introspective examination of his intuitions," and that, as a consequence, the central aim of linguistic theory has shifted from being merely "combinatoric" to being "a project, first in cognitive psychology and ultimately in biology more generally." Once again, the record is simply not on their side.

First, the whole of Chapter IV in Chomsky (1955) is devoted to "grammaticalness," the primary task being to offer some proposals as to how to go about reconstructing this notion which should ultimately be defined by a general theory of linguistic form. Second, we have already seen that Chomsky's early generative grammar could not have been purely "computational or combinatoric" in B&H's sense. Third, we have also argued that the distinction between "grammaticality" and "acceptability" was present and, indeed, crucial for Chomsky's theorizing already in the mid-fifties. Thus, there is no reason to believe that the cognitive theme marks a shift in the aim of linguistic theory, for such a theme was never absent in Chomsky's early work.[6] Let us elaborate more on this last point by turning to Plato's problem because this is central to B&H's argument about the advent of cognition in *Aspects*.

The significance of Plato's problem in linguistics can be appreciated by contrasting the results of linguistic analysis with a striking empirical observation. On the one hand, such analysis reveals a complex structure of rules and relations underlying linguistic knowledge. On the other, children acquire their language within a relatively short period of time and on the basis of limited linguistic experience. Given these boundary conditions of time and experience, how could the child's knowledge of an intricate linguistic system ever be possible? Implicit in this question is an assumption without which the problem of language acquisition is hardly genuine. Put briefly, the assumption asserts that linguistic experience *alone* cannot account for the child's knowledge of language. This constitutes the focus of a well-known argument in the literature, the so-called poverty-of-stimulus argument, which purports to show that much of what children know about their language goes far beyond what their linguistic environment actually justifies.[7]

Now, insofar as the argument from poverty of the stimulus was intended to establish the discrepancy between knowledge and experience, it was by no means absent in Chomsky's early work. Thus Chomsky writes:[8]

> A speaker of a language has observed a certain limited set of utterances in his language. On the basis of this finite linguistic experience he can produce an indefinite number of new utterances which are immediately acceptable to other members of his speech community. He can also distinguish a certain set of "grammatical" utterances, among utterances that he has never heard and would never produce. Can we reconstruct this ability in a general way? I.e., can we construct a formal model, a definition of "grammatical sentence" in terms of "observed sentence," thus, in one sense, providing an explanation for this ability? Can we show that there are deep underlying regularities in observed utterances that lead to these results? (Chomsky 1955: 715)

Clearly, some notion of poverty of stimulus resonates here. To be sure, a native speaker "has observed a certain limited set of utterances in his language," and despite "this finite linguistic experience," he nonetheless exercises his linguistic capacity far beyond what his linguistic experience could have provided a basis for, in the sense that his knowledge of what constitutes an acceptable sentence in his language could not have possibly been derived inductively from his limited linguistic experience. A familiar example from Chomsky's early work will make this clear. The sentence in (4) and its mirror image in (5) are from Chomsky (1955: 38–9) [his (3) and (5), respectively]:

(4) Colorless green ideas sleep furiously.
(5) Furiously sleep ideas green colorless.

Chomsky (1955) argues that, although both (4) and (5) are nonsensical, a native speaker of English will most likely consider the former, but not the latter, to be grammatical. This intuitive sense of grammaticalness is manifested, for instance, by the fact that the native speaker will most likely read (4) with a different pattern of intonation than that of (5). In the former case, the intonation pattern would reflect the standard intonation of an English sentence. In the latter, however, the intonation would be falling with each word, as in the case of reading a list of isolated words. "Yet [the speaker]," Chomsky (1955: 39) argues, "has presumably never heard either [(4)] or [(5)]." In other words, there is nothing in speakers' linguistic experience that would indicate to them that (4) and (5) should have different intonation patterns. The important point to be stressed here is that, as far as limited linguistic experience is concerned, the above characterization of the linguistic intuition of the native speaker is clearly parallel to that of the child acquiring her language.[9] In short, there is nothing that B&H attribute uniquely to the cognitive stage that cannot be found in the "earlier" stage. This applies also to what they say here:

> In the *Aspects* era, grammars are empirically motivated in two ways: *internally*, in that they respect a speaker's intuitions about the grammar and *externally* by being acquirable by a child in the circumstances that characterize language acquisition. (p. 121)

But this is true of *Syntactic Structures* just as much as it is true of *Aspects*.[10] To see that this is indeed the case, consider the following passage from *Syntactic Structures*:[11]

> Clearly, every grammar will have to meet certain *external conditions of adequacy*; e.g., the sentences generated will have to be acceptable to the native speaker ... In addition, we pose a *condition of generality* on grammars; we require that the grammar of a given language be constructed in accordance with a specific theory of linguistic structure. (Chomsky 1957: 49–50, italics in original)

This strongly echoes what we find in *Aspects*:

> [T]here are two respects in which one can speak of "justifying grammar." On one level (that of descriptive adequacy), the grammar is justified to the extent that it correctly describes its object, namely the linguistic intuition – the tacit competence – of the native speaker. In

> this sense, the grammar is justified on *external* grounds . . . On a much deeper and hence much more rarely attainable level (that of explanatory adequacy), a grammar is justified to the extent that it is a *principled* descriptively adequate system, in that the theory with which it is associated selects this grammar over others, given primary linguistic data. In this sense, the grammar is justified on *internal* grounds. (Chomsky 1965: 26–7, italics in original)

Indeed, the relationship between a grammar and the general theory with which it is associated is characterized in the same way in Chomsky (1965: 32) as it is in Chomsky (1955: 78, 1957: 52). I am referring here to the reliance on *evaluation procedures*, i.e. the weakest form by which the relationship between particular grammars and the theory of linguistic structure may be conceived. Chomsky (1955, 1957) advocates the development of evaluation procedures as the most ambitious realistic goal for linguistic theory and takes some tentative steps towards formulating partial versions of such procedures with the fundamental notion being simplicity. Thus given two descriptively adequate grammars G_1 and G_2, the former is preferred to the latter if it is the simpler of the two.[12]

This conception of an evaluation measure remained essentially unchanged in *Aspects*. However, in order to make acquisition explicable in principle while taking account of limited linguistic experience and the relatively short period of time over which it occurs, there has been a continuous effort to reduce the number of possible hypotheses available to the child. To achieve this, Chomsky (1965: 46) suggested two different approaches: to refine the evaluation measure for grammars, and/or to augment the constraining power of the theory of linguistic form (i.e. universal grammar, or UG) in such a way "that it becomes more difficult to find a highly valued hypothesis compatible with" the meagre linguistic data that the child receives from the environment (i.e. primary linguistic data, or PLD). This latter approach seemed more promising to Chomsky, leading eventually to his framework of *principles and parameters* (Chomsky 1981).

B&H seem to suggest that, with the advent of the principles and parameters approach (henceforth P&P), an evaluation measure becomes unnecessary. This is explicit in Boeckx (2006: 55), who asserts that "with the development of the P&P framework during the past decade and a half an evaluation measure for grammars has become essentially superfluous" (cf. Freidin 2007: 288, who makes a similar claim). But this assertion does not seem to be obviously correct.

Consider, first, how the rule-based acquisition model was supposed to work. It is supposed that "the child composed a grammar by writing rules in the rule writing system, under the constraints that the rules must be compatible with the

data, and that the grammar must be the one most highly valued by the evaluation metric" (Williams 1987: vii). Consider now how a P&P acquisition model is supposed to work.[13] It is assumed that the child comes equipped with a finite set of universal principles, together with a set of open parameters that are sensitive to PLD. To acquire a specific language, the child sets each parameter one way or another on the basis of PLD. Thus different parametric settings lead to different possible languages. At some intermediate stage of the acquisition process, we suppose, certain parameters have been set, whereas others await their appropriate "triggering" data. The child, we continue to suppose, is revealing something at this stage about her grammatical system. But it is not all clear what this system could be, with some parameters set and others unset. It seems implausible to suggest that the unset parameters can just be ignored, for a P&P model is considered to be massively "deductive," in the sense that any "gap" in the system will render it unworkable. One way to get around this uncertainty is by adopting a theory of "markedness" in the sense of Chomsky (1981). Thus we may suppose that, assuming binary parameters, the child chooses one parameter value (the unmarked one) by default, while he chooses the other value (the marked one) only if the evidence warrants such a choice. Accordingly, the initial choice of unmarked parameter values represents the initial hypothesis that the child makes about the language in her environment, whereas the acquisition process amounts to resetting at a marked value each default parameter value when the evidence warrants such a resetting. Now, it is not unreasonable to regard this notion of "markedness" as comprising something like an evaluation measure, with hypotheses containing unmarked parameter values being more highly valued than those with corresponding marked values. Consequently, the assertion that an evaluation measure becomes "essentially superfluous" with the advent of the P&P framework is questionable at best, and misleading at worst.

It is a curious characteristic of some of the proponents of the Chomskyan school to think highly (perhaps, too highly) of the P&P framework. For instance, there have been repeated assertions that the P&P approach has "solved" Plato's problem (e.g. Boeckx 2006: 61; Hornstein *et al.* 2005: 5; Boeckx and Hornstein 2010: 134).[14] It is interesting to notice that Chomsky himself has been more circumspect, merely observing that the approach "offered the hope of overcoming the tension between descriptive and explanatory adequacy" (Chomsky 2005: 8), or that it "suggested ways to reduce the tension between descriptive and explanatory adequacy" (Chomsky 2007b: 2). Thus, one is left to ask on what basis the claim that the P&P approach has solved the acquisition problem is made, especially when the literature suggests otherwise. Indeed, consider

what a genuine solution to Plato's problem would look like. This would consist of proposing and justifying a (fairly small) number of parameters that, on the basis of the specific properties of input data, interact in such a way as to yield a particular set of parameter settings (i.e. a grammar). However, and despite many efforts, a solution of this sort has never been even approximated. And to make matters worse, there seems to be a lack of consensus regarding the number (and nature) of parameters ascribed to UG.[15] For instance, while Cowper (1992: 17) suggests that "[w]hat is mysterious about parameters in syntax is that there seem to be few of them," Newmeyer (2006: 7) maintains that "[i]f fewer than 1000 parameters can be found in the literature, [he] would be very surprised." Moreover, in the words of a fairly recent book on the subject, "[e]ven the most studied and well-known parameter, the *pro*-drop or null subject parameter, is still being debated" (Ayoun 2003: 8). This is not the place to review the vast literature on language acquisition, but if the aforementioned examples are indicative, we might urge caution with respect to the achievements of the P&P framework with regard to this field. Of course, this is not to deny the significance of the P&P framework, especially its role in giving fresh momentum to cross-linguistic comparative research, and in breaking with a long tradition of conceptualising the structure of language in terms of rules and constructions. The point is rather that we should guard against wishful thinking by not rushing to claim victory where there is none.

But as soon as we examine the rationale that B&H (and others) use to explain the emergence of minimalism, we begin to see why they were quick in claiming success for the P&P framework.

2.2.3 The minimalist stage

Consider first how B&H (p. 134) justify the shift from the cognitive to the minimalist phase:

> Because the principles-and-parameters approach "solves" Plato's problem, more methodological criteria of theory evaluation revolving around simplicity, elegance, and other notions that are hard to quantify but are omnipresent in science can become more prominent. Until [Chomsky's (1981) *Lectures on Government and Binding*], solving the acquisition problem was the paramount measure of theoretical success. Once, however, this problem is taken as essentially understood, then the question is *not* how to solve it but how *best* to do so. (B&H, p. 134, italics in original)

Hornstein *et al*. (2005: 5–6) make the same point by suggesting that "the consensus that P&P-style theories offer a solution to Plato's problem necessarily affects how one will rank competing proposals," where "the issue becomes which of the conceivable P&P-models is best," with the consequence of paving the way "for simplicity, elegance, and naturalness to emerge."

It is not our task here to go into the question of what has triggered the shift to minimalism – we will return to this interesting question in detail later (Section 2.6). For now, we will continue our (indirect) review of the development of Chomskyan linguistics with an eye to clearing up some misconceptions, of which we have just cited two examples. What these examples appear to suggest is that the need to solve the problem of language acquisition had kept methodological notions such as simplicity and elegance off the scene, and that only when this problem was "solved" did such notions acquire a dominant role in minimalism, especially in connection with finding the "best" solution to the acquisition problem. There seem to be at least three difficulties with this suggestion, apart from the fact that it is fanciful to suppose that Plato's problem has been solved.

First, it falsely implies a discontinuity in the role of methodological considerations in Chomskyan linguistics. To be sure, methodological considerations of elegance and simplicity may be relevant to choosing the most appropriate P&P model, but as mentioned in the previous section, the same considerations have been present in evaluating competing grammars. Moreover, the very significance of the P&P framework can be appreciated by an appeal to methodological considerations. For instance, insofar as we grant that the P&P approach allows only a small number of parameters in a deductively rich system, we may argue for the superiority of a P&P model over a rule-based model. Thus, not only were methodological considerations present throughout the generative enterprise, but they were also helpful in evaluating its progress.

Second, the suggestion that the urgency of Plato's problem kept methodological considerations off the agenda makes the notion of simplicity contrast inappropriately with the notion of explanatoriness. But the truth of the matter is that not only has simplicity been a decisive factor in the effort to investigate the structure of language, but it has also been invoked as part of the explanation for the acquisition of language.[16] As noted earlier (Section 2.2.2), considerations of simplicity constituted the basis on which the internal justification of grammars was advanced, and, as such, they formed an integral part of explanatory adequacy. Moreover, one might be impressed by the *rate* at which the child

acquires his language – a factor that might be seen as implying that insofar as acquisition involves choices, these choices are relatively restricted in number. Thus there is room here for the application of the notion of simplicity to the acquisition process.

Third, and perhaps most importantly, to suggest that, with the advent of minimalism, "the issue becomes which of the conceivable P&P-models is best" is to suggest that the methodological aspect of simplicity is what makes it relevant to minimalism. But surely there is more to minimalism than Occam's razor. It is a misconception of the minimalist program to regard it as simply an exercise in searching for the "best solution" to Plato's problem. For if it were indeed a matter of finding the "best" P&P-model, one would expect to find systematic comparison of different parametric models, but the truth is that there is little contentful discussion of parameters in Chomsky's minimalist work.[17] Moreover, even if we grant that the essence of the MP is to act as an arbiter of competing syntactic models, one would expect that this should be in relation to the problem of connecting sound and meaning, and not in relation to Plato's problem, for it is a basic tenet of minimalism, encapsulated in the strong minimalist thesis (SMT, Chapter 3) that language constitutes an optimal solution to the problem of relating sound and meaning.

It might further be observed that by suggesting that the MP reduces to finding the best P&P model the authors are in fact implying that empirical evidence from language acquisition is relevant to the validity of minimalism. But Chomsky (2000b: 96) makes it clear that, in assessing the empirical validity of the SMT, information about acquisition "is irrelevant in principle." Now, if this is so, then it cannot be true that the goal of minimalism is to find the best solution to Plato's problem.

One last point is in order before we conclude the present discussion. As mentioned at the beginning of the present chapter, B&H claim a relationship between minimalism and physics (see Chapter 5 for a critical discussion of this claim). However, they offer little by way of detail regarding what this relationship amounts to, merely quoting eminent physicists such as Albert Einstein and Richard Feynman and proposing that minimalists do not differ from physicists in their search for an ultimate explanation of natural phenomena. But it is not at all clear why the search for an ultimate explanation should be viewed as establishing a noteworthy relationship between minimalists and physicists, for every self-respecting scientist in any field would strive for such an explanation. As Section 2.3 will make clear, the intended relation to physics in Chomsky's work is supposed to be much deeper than this.

2.2.4 Concluding remarks

It is a major defect of B&H's paper that their account of the development of linguistic theory fails to recognize continuities in Chomsky's work. They consider *Aspects* as the starting point of cognitive themes in Chomsky's linguistic theory, but we have seen that a cognitive dimension was present from the outset of generative grammar. And they claim that methodological considerations of simplicity and elegance were revived with the advent of minimalism, but we have argued that such considerations were never absent from Chomsky's theorizing.

B&H's failure to offer an accurate and insightful account of the development of Chomskyan thought can perhaps be attributed to their goal-directed approach, where they identify different stages with distinct goals. Consider, for example, the question of why the goals of linguistic theory shift from one historical stage to another – a question that can be regarded as a direct consequence of B&H's goal-directed approach. As is evident from the preceding review, this is a question that has led to rather gratuitous answers, such as that linguistic theory shifts from one goal to another when it has been discovered that the goal in question was misguided (the goal of the combinatoric stage), or else it has been accomplished (the goal of the cognitive stage).

As an alternative to the goal-directed approach, it might be maintained that the development of Chomskyan linguistics is better conceived of in terms of the *problems* it has been faced with. We do not pursue this approach here, but merely remark that there are good reasons why such an approach may be more enlightening. First, by dealing with the development of linguistic theory in terms of the problems it has faced, there is no need to force an analogy between the study of language and other fields of inquiry, for a single scientific problem need not be classified under a single field of research; indeed, it is often the case that a single scientific problem requires a variety of tools derived from distinct scientific fields.[18] As Popper (1963: 88) puts it: "*We are not students of some subject matter but students of problems*. And problems may cut right across the borders of any subject matter or discipline" (italics in original). Second, scientific problems have a crucial role in shaping the development of scientific theories, since these are often motivated, whether on empirical or theoretical grounds, by certain problematic observations, and also lead to new questions and different types of problems. From this perspective, a problem-directed approach appears attractive, since it promises to uncover the basis on which the development of linguistic theory rests, namely the give-and-take relationship between problematic observation and theory construction.[19] Finally, unlike the goal-directed approach, in which the justification of the shift from one goal

to another suggests a lack of continuity in the advance of linguistic theory, a problem-directed approach gives unity to the general development of Chomsky's work, a positive outcome since the fundamental problems in linguistic theory exhibit a striking resemblance in terms of the ways in which they have been posed and pursued.[20]

2.3 The shift to minimalism

If the transition to minimalism from its predecessors is not viewed in terms of a focus on what are essentially methodological grounds (see above) how should it be understood? Let us examine this question.

One of the major characteristics of the MP is the speculation that the faculty of language is simple, operates on the basis of capacities that are largely identifiable in other aspects of human cognition and indeed in other species, and is (largely) not determined by aspects of the human genome. This emphasis is seen by many as an acute departure from the pre-minimalist view on the nature of language, where the prevailing position seemed to be that language was an intricate system, substantially determined by genetic endowment, and fundamentally autonomous from other cognitive domains. It is my intention here to show that the shift to minimalism is not as dramatic as some would believe it to be (cf. Pinker and Jackendoff 2005).

For the purpose of the discussion to follow, let us enumerate what may be regarded as the defining features of the pre-minimalistic conception of the language faculty. These are that it is: (a) innate; (b) rich or complex; (c) genetically determined; (d) cognitively autonomous; and (e) species-specific. One may be tempted to question whether it is sensible to distinguish between (c) and both (a) and (e), but as will become clear later, the distinctions are warranted, especially if we are to appreciate the nature of the shift we are concerned with. The present discussion will deal with these attributes and seek to identify their fate in the context of the minimalist program. But before we begin, a general observation may be useful.

As we will see in the next section, the MP attempts to reduce the complexity of UG by shifting the burden of explanation of core aspects of language from genetic constraints to general principles that are not specific to language.[21] This may be regarded as a shift in emphasis from genetic to *non-genetic nativism*, with the assumption being that what is innate in language, as opposed to what is learned, is either a manifestation of our genes or a consequence of non-biological, physical law. How this assumption is justified will be discussed later, but here suffice it to say that, although non-genetic nativism was given

explicit recognition only in minimalism, there is reason to believe that it was an important factor behind Chomsky's pre-minimalist attitude of open-mindedness and caution towards the five attributes referred to above. That some notion of non-genetic nativism has been in the background of Chomsky's thinking prior to the emergence of minimalism is an assumption that I shall attempt to justify later (Section 2.6). What is important for our present purposes is to remark that such an undogmatic and cautious attitude, amply illustrated in the discussion that follows, seems to always leave the door open for non-genetic nativism. However, in pre-minimalism the evidence was not available to bring it to the fore, and so long as this was the case, the five attributes listed above came together to yield the "classic" view of the language faculty. But as soon as Chomsky found a way to relate his linguistic framework to the notion of "physical law," the classic view was replaced by the minimalist view *but without entailing a fundamental shift regarding the nature of language*. This was indeed the nature of the change – a re-balancing of emphases rather than a fundamental re-orientation, as we now demonstrate.

One line of reasoning that has been constant throughout Chomsky's pre-minimalist work is this: if it can be shown that a property P cannot possibly be derived, *via* standard inductive procedures, from linguistic experience, then P must be *innate* (as opposed to "learned"). This is essentially the force of the argument from poverty of the stimulus, to which we have already alluded in p. 14. In addition to postulating an innate structure for the language faculty, Chomsky sought to show that this structure must be *rich* or *complex*. This is another way of asserting that the poverty-of-the-stimulus argument applies widely. Thus Chomsky maintains that the moment one takes seriously the linguistic knowledge attained by the child at a very young age, one is "led to the conclusion that intrinsic structure is rich (by the argument from poverty of stimulus)" (Chomsky 1980a: 41).[22]

Notice, however, that the poverty-of-the-stimulus argument is essentially negative; the ascription of innateness is justified because learning, in the sense of induction from experience, *cannot* be solely responsible for the apparent complexity of linguistic structure. What this indicates is that the argument is neutral as to whether or not the innate structure is *genetically determined*. As a consequence of this neutrality the argument survives the transition to minimalism. Indeed, if what is *innate* need not be *genetically innate*, then the poverty-of-the-stimulus argument is consistent with the minimalist emphasis on a non-genetic nativism. Before this emphasis became official, however, Chomsky's appeal to *genetic* endowment to account for the discrepancy between what is innate and what is learned seems appropriate in a period where the *available* evidence directed him to the genetic factor. It also seems

perfectly in tune with an approach that emphasizes the biological nature of language. Thus Chomsky writes:

> The argument from poverty of the stimulus leaves us no reasonable alternative but to suppose that [language] properties are somehow determined in universal grammar, *as part of the genotype.* (Chomsky 1980a: 66, my italics)

And again:

> If, say, we find extensive evidence that the principles that underlie [a given linguistic] constraint belong to universal grammar and are available to the language learner without relevant experience, *then it would only be rational to suppose that these mechanisms are genetically determined* and to search for a further account in terms of biological development. (Chomsky 1980a: 209, my italics)

Indeed, in this connection, Chomsky (1975a: 29) refers to what he terms "biological necessity," a concept that he assimilates into his definition of "universal grammar."[23] It is well to bear in mind, however, that while the argument from poverty of the stimulus (even when couched in terms of "biological necessity," thereby setting aside the non-genetic possibility) is intended to rule out the environment as the primary source from which the complexity of linguistic structure is derived, it does not exclude the possibility that much of this complexity might be neither linguistically autonomous nor absent in non-human species. In other words, the ascription of genetic endowment, while confining linguistic experience to a secondary role with respect to acquisition, neither entails that the properties thereby accounted for do not exist in other cognitive systems nor that they are specific to humans. In the case of non-human species, it would not be logically absurd to suggest that language has a rich genetic basis and that, at the same time, virtually every property of language can be identified in some non-human domain. As to human cognitive faculties, one can consistently maintain that the mind has a rich innate structure and that, at the same time, mental structure is homogeneous across cognitive domains (including language). In other words, that the structure of the mind is innate and complex is a position that is neutral with regard to whether or not the various cognitive systems of the human mind are determined by different principles. In Chomsky's (1980a: 40) terms, "[o]ne might hold that there is rich innate structure, but little or no modularity."

Yet, no doubt, many would argue that Chomsky's position has clearly favoured a modular approach to cognition in general. But this assertion must

be examined carefully. We should first be clear that the sense of "modularity" here is distinct from and less articulated than that developed in Fodor (1983).[24] Modularity for Chomsky (ibid.) is simply the assumption "that various systems of the mind are organized along quite different principles." In the case of language, this translates into the assumption that language (*qua* cognitive system) is specific in terms of the principles that underlie its structure and properties. As we will see shortly, the notion of "language specificity" receives a new interpretation in minimalism. But let us first consider Chomsky's pre-minimalist views on this issue. To the question of whether the properties of language are specific to it or whether they are shared with other cognitive capacities, Chomsky proposed a clear, albeit cautious, answer:

> There seems little reason to suppose, for the moment, that there are general principles of cognitive structure, or even of human cognition, expressible at some higher level, from which the particular properties of particular "mental organs," such as the language faculty, can be deduced, *or even that there are illuminating analogies among these various systems*. (Chomsky 1980a: 215, my italics)

This was also Chomsky's position in his famous debate with Jean Piaget, which took place at the Abbaye de Royaumont (Paris) in 1975, reported in Piattelli-Palmarini (1980). A central issue that was extensively discussed in that debate concerns human linguistic capacities and their foundation. Counter to Piaget's views, in which human language is regarded as the product of progressively constructed processes of general intelligence, Chomsky (1980b) was at pains to argue for the view that the faculty of language is an intricate system with specific properties which are genetically fixed and he remained skeptical about the prospects of finding UG-like properties in non-linguistic domains. This skepticism is reflected in the following interchange with one of the participants of the debate, the eminent psychologist David Premack:

> PREMACK: ... you said that there is no hope for the possibility that one will find in non-linguistic domains the kind of formal properties one finds in language ...
> CHOMSKY: I didn't say that; I said that I didn't *see* any hope for that.
> PREMACK: That seems to me a very premature judgment ... I think it is premature to conclude that the formal structures one knows to exist in language will not be found elsewhere, in another species or perhaps even in other human domains. Let's wait and see.

> CHOMSKY: . . . I don't see any particular reason to expect the same result, but if that happens, I would be very pleased.
> (Piattelli-Palmarini 1980: 179–80)

Observe, however, that Chomsky's skepticism does not extend to dogmatism, as he has repeatedly and explicitly considered the issues to be empirical, their outcome to be determined *a posteriori* by further research and experimentation.[25] It is noteworthy that in the 1980s the skepticism that we have been highlighting is apparent not just with respect to cognition in general, but also with regard to systems with which the language faculty might be assumed to interact closely, *viz*. systems of sound and meaning. Note, however, that the caution that we have seen above is also manifest here with the use of "seems" and "tends" in the following passage from Chomsky (1980a: 246):

> The belief in the "simplicity" of mental structures is related to the doctrine of uniformity. In the case of language, it is commonly argued by linguists and others that the principles of grammar cannot be "too complex" or "too abstract" but must reflect properties of sound and meaning, or must be directly determined in some manner by "functional considerations," aspects of language use. Evidently, there can be no *a priori* argument to this effect. To me, it seems that recent work tends to support a rather different view: that rules of syntax and phonology, at least, are organized in terms of "autonomous" principles of mental computation and do not reflect in any simple way the properties of the phonetic or semantic "substance" or contingencies of language use.

Now, apparently, this passage contrasts sharply with Chomsky's minimalist views, most notably with his "strong minimalist thesis" (SMT), in which the claim is not only that the language faculty cannot be "usable" unless it engages with speech and thought systems, but also that such engagement with these performance systems determines to a large extent the properties of language.[26] Moreover, it would seem that this shift in perspective involves a retreat from the claim of modularity, but does it really?

Boeckx (2006) seems to imply that it does, and he provides a rather unsatisfactory justification. In what is supposed to be a defence of Chomsky's position, Boeckx (2006: 148) asserts that modularity "has . . . been assumed and emphasized as a reaction to the Piagetian view that language learning is an expression of intelligence."[27] But this trivialises Chomsky's position on modularity, portraying it simply as a tactical move to secure a nativist view of

language. More importantly, insofar as modularity is viewed as an empirical matter, the crucial assumption is not so much whether or not modularity should be "assumed," but whether there is enough evidence for modularity to refute the hypothesis of uniform development across cognitive domains, and, therefore, undermine the Piagetian view that innate mental structure need not be complex.

As to whether the shift to minimalism involves a retreat from the claim of modularity, consider the following passage from Chomsky (1995a: 221):

> Suppose that [the minimalist] approach proves to be more or less correct. What could we then conclude about the specificity of the language faculty (modularity)? Not much. The language faculty might be unique among cognitive systems, or even in the organic world, in that it satisfies minimalist assumptions. Furthermore, the morphological parameters could be unique in character, and the computational system C_{HL} biologically isolated. Another source of possible specificity of language lies in the conditions imposed "from the outside" at the interface, what we may call bare output conditions.

There are two points to consider here. First, Chomsky seems to suggest that there is not much to be concluded about modularity from the minimalist postulation of a simple genetically determined component in the structure of the language faculty. This is clearly consistent with his earlier view, to which we have referred above, namely that the extent to which language may be genetically determined entails nothing about modularity. Second, there appears to be a reinterpretation of the notion of "specificity." As mentioned earlier, from a pre-minimalist perspective, the specificity of language is interpreted in terms of the principles that underlie the structure of the language faculty. Given the passage just cited, we now seem to have an interpretation of specificity in terms of what is considered to be the core function of language, namely satisfying legibility conditions at the interfaces of sound and meaning. Consequently, it is not so much that the assumption of modularity has been abandoned, but rather that the domain in which it is investigated has changed. While the above discussion has not excluded considerations of non-human systems of cognition, it has been largely framed with human cognition in mind. We now turn to explicit discussion of specificity where the comparisons with the language faculty involve non-human cognitive systems.

In his (relatively) recent collaborations with the biologists Marc Hauser and Tecumseh Fitch (see Hauser *et al*. 2002; Fitch *et al*. 2005), Chomsky has defended the view that many of the properties of language may be identified

in other species, hence suggesting that the extent to which language is special may be very restricted. Some have inferred from this that the pre-minimalist assumption of species-specificity has been *largely* renounced (Pinker and Jackendoff 2005; Kinsella 2009). However, a closer inspection of Chomsky's work reveals otherwise. Notice, first, that even if we grant that there is a shift involved here, we must nevertheless be aware of the fact that it is entirely quantitative and not qualitative; that is, there is no denial here of the view that language is a human prerogative, but only a speculation as to the extent that it is such. Indeed, the fact that language has evolved *only* in humans would appear to clearly indicate that there is *at least something special* about it.[28] We turn then to consider briefly whether there has been a change in Chomsky's conception of language with respect to the notion of species-specificity.

What is the main claim that has been made by Chomsky and his collaborators? The substance of the ideas involved in this question will be discussed in the next chapter, and in greater detail in Chapter 4. Here suffice it to say that Hauser *et al*. (2002) distinguish between the faculty of language in the narrow sense (FLN) and the faculty of language in the broad sense (FLB). The former contains properties that are unique to humans and unique to language, and the latter comprises all other properties which (a) are somehow involved in language and (b) may be shared with other species or have a role in other human cognitive domains. Given this distinction, the authors advance the hypothesis that *recursion* is the only aspect of the language faculty that is unique to humans and specific to language. The general notion of "recursion" is instantiated by the syntactic operation Merge, which "takes a finite set of elements and yields a potentially infinite array of discrete expressions" (Hauser *et al*. 2002: 1571).[29] This constitutes the major claim of this collaborative work, and it does not immediately appear to signal a major departure from Chomsky's pre-minimalist views on the nature of the language faculty. Consider, for instance, the following two passages from one of Chomsky's early works, in which he draws "striking similarities between the seventeenth-century climate of opinion and that of contemporary cognitive psychology and linguistics" (Chomsky 1968: 15). The first passage is a commentary on the definition of human intelligence according to the Spanish physician Juan Huarte:

> Huarte came to wonder at the fact that the word for "intelligence," *ingenio*, seems to have the same Latin root as various words meaning "engender" or "generate." This, he argued, gives a clue to the nature of mind. Thus, "one may discern two generative powers in man, one common with the beasts and the plants, and the other participating of

> spiritual substance. Wit (Ingenio) is a generative power. The understanding is a generative faculty." Huarte's etymology is actually not very good; the insight, however, is quite substantial. (Chomsky 1968: 9)

Chomsky then turns to an observation made by Descartes concerning the distinctive feature of human language:

> In fact, as Descartes himself quite correctly observed, language is a species-specific possession, and even at low levels of intelligence, at pathological levels, we find a command of language that is totally unattainable by an ape that may, in other respects, surpass a human imbecile in problem-solving ability and other adaptive behavior ... There is a basic element lacking in animals, Descartes argued ... namely Huarte's second type of wit, the generative ability that is revealed in the normal human use of language as a free instrument of thought. (Chomsky 1968: 10–11)

Clearly, these two passages are much in line with the hypothesis that the recursive property of language is what distinguishes it from non-human communicative systems. In fact, the continuity between Chomsky's pre-minimalist work and Hauser *et al*. (2002) goes even further. For instance, Chomsky's (1980a: 54–5) description of the distinction between the computational and conceptual systems is fairly similar to that between FLN and FLB:[30]

> Suppose that what we call "knowing a language" is not a unitary phenomenon, but must be resolved into several interacting but distinct systems. One involves the "computational" aspects of language – that is, the rules that form syntactic constructions or phonological or semantic patterns of varied sorts, and that provide the rich expressive power of human language. A second component involves the system of object-reference and also such relations as "agent," "goal," "instrument," and the like ... For want of a better term, let us call the latter a "conceptual system." ... Supposing all of this, let us distinguish a system of "computational" rules and representations that constitute the language faculty, strictly speaking, and a system of conceptual structure organized along the lines just described. The two systems interact. Thus certain expressions of the linguistic system are linked to elements of the conceptual system and perhaps rules of the linguistic system refer to thematic relations.

On the basis of this distinction, Chomsky (1980a: 57) goes on to speculate that higher apes might share with humans certain parts of the conceptual system, although they lack the recursive property of the computational system.[31]

In view of the passages cited above, there seems to be no substance to Pinker and Jackendoff's (2005) conviction that Chomsky's contribution to Hauser *et al.* (2002) represents a "major recantation." It is of interest to notice that critics who believe in Chomsky's "major recantation" often refer to the pre-minimalist conception of language as a genetically complex system to support their position.[32] But we have already argued that this conception was based on the argument from poverty of the stimulus, an argument that is compatible with the non-genetic nativism underlying the minimalist view. Once this point is taken into consideration, it is easy to see that the argument offered to substantiate Chomsky's "major recantation" cannot be sustained. To show that this is indeed the case, let us consider what is at issue here.

Put simply, the argument seems to be this: since pre-minimalist approaches considered UG to be complex, and since minimalism counters this by attempting to demonstrate that much of the complexity that has been assigned to UG falls outside its domain, it follows that a subscription to minimalism entails that much of the complexity of language is shared with other species. There are at least two implicit assumptions underlying this argument which, when made explicit, demonstrate its untenability: first, that what falls inside UG must be both genetically determined and unique to language; second, that what falls outside UG must be *ipso facto* shared with other species. The former assumption does not necessarily reflect the definition of UG in Chomsky's pre-minimalist work, although it does seem to be in tune with his later work.[33] For the sake of argument, I will assume that it is a valid assumption. But in this case, a property X that falls outside UG may belong to one of the following three categories: (I) X is genetically determined, but not unique to language; (II) X is not genetically determined, but is unique to language; and, (III) X is neither genetically determined nor unique to language. By considering category (I) as the *only alternative* to what falls inside UG, proponents of the argument above are clearly guilty of a false dichotomy – a fallacy which renders their argument invalid. Category (II) exemplifies the specificity of language from a minimalist perspective as indicated earlier, and category (III) represents the minimalist bet on non-genetic nativism, namely that many of the properties of UG may not form part of the genetic component of language (or they are at least not unique to it). We can now see, I hope, that what the above argument has failed to take into account is that reducing the "size" of UG does not necessarily amount to an increase in the number of shared properties, for there exists the

possibility that what falls outside UG may not be genetically determined, in which case the distinction "unique versus shared" does not even arise.

To conclude, I hope to have shown that the shift to minimalism is probably not as dramatic as some critics would suggest. We have tried to discern the nature of this shift in terms of the five pre-minimalist attributes mentioned above, and we have sought to identify the fate of each in the context of minimalism. It should be clear from the above discussion that the key inference in pre-minimalist thinking was from innateness to genetic endowment (i.e. (a) → (c)), but more explicit recognition of the role of non-genetic nativism has questioned the general applicability of this inference. This recognition has also opened the way for Chomsky's caution and open-mindedness on the questions of cognitive autonomy (d) and species-specificity (e) to be retrospectively justified. This seems to confirm the general observation with which we began our discussion, namely that Chomsky's non-genetic nativism was an important factor behind his pre-minimalist attitude of open-mindedness and caution towards the nature of language. The shift in emphasis from genetic to non-genetic nativism has consequently led to a shift in the theoretical role of UG, the topic to which we now turn.

2.4 UG: from an *explanans* to an *explanandum*

Two different, but closely related, aspects of language have dominated Chomskyan linguistics for many decades, namely acquisition and typology. There are many languages across the world and their apparent diversity is undeniable. Yet, *ceteris paribus*, children acquire any language with apparent ease, "so it must be that the basic structure of language is essentially uniform and is coming from inside, not from outside" (Chomsky 2002: 93).

Diversity and uniformity are two seemingly incompatible linguistic phenomena, and together they form the *explananda* of UG. It is not clear what form UG should have, especially when being "pulled" in two different directions by these two phenomena. On the one hand, acquisition requires UG to be restrictive enough to account for the discrepancy between experience and knowledge. On the other, typological variation requires UG to be unrestrictive enough to allow for the observed variation between natural languages. "We must postulate," writes Chomsky (1968: 79), "an innate structure that is rich enough to account for the disparity between experience and knowledge," but it "must not be so rich and restrictive as to exclude certain known languages." This statement encapsulates what is known in the literature as the tension between descriptive and explanatory adequacy, an issue that has dominated the research agenda for many years

since the introduction of the so-called standard theory model in *Aspects*.[34] As Jackendoff (1983: 10) has pointed out, every "revision of the [standard] theory has been motivated by the desire to constrain the possible variation among particular grammars, so as to limit the choices the language learner must make." As it turned out, the expressive power of rules and transformations in terms of data coverage was impressive, but so too was their number and complexity. Such complexity resulted from a focus on descriptive adequacy in terms of specific rules and constructions, and it eventually led to "the question of how language learners go about finding their correct formulation" (Webelhuth 1995: 8).

The decade 1970–80 saw vigorous efforts to give due weight to explanatory adequacy in addressing the tension between it and descriptive adequacy. An outstanding exemplification is Chomsky (1973), in which the primary goal was to restrict the descriptive power of transformational rules by imposing general constraints on transformations. Details aside, a steady change from rule-based to principle-based frameworks culminated in Chomsky's (1981) *Lectures on Government and Binding*.

As noted earlier, the P&P approach suggested a way to explain how children acquire their language and why natural languages appear to differ from each other at a superficial level, thus directly addressing the tension between descriptive and explanatory adequacy. But Chomsky (1995a, 2000b, 2002, and in other works) takes this claim a step further by suggesting that the P&P approach has paved the way for the emergence of the minimalist program. From Chomsky's (2004a) perspective, the parametric approach to language opens the way to move linguistic inquiry "beyond explanatory adequacy" to a deeper level of explanation. It should be noted, however, that, *pace* Chomsky, it is *not* the resort to parameter setting itself that has this consequence, but rather that the recognition of non-genetic factors – which are supposed to lead to the "deeper level of explanation" – opens the way for UG to have a reduced role in accounting for acquisition. If this is true, then whatever properties have been ascribed to UG in the course of explaining language acquisition should now be justified on a non-genetic basis; thus, UG *itself* calls for a reassessment.[35] In short, the Chomskyan appeal to proceed "beyond explanatory adequacy" embodies a shift in the theoretical role of UG from an *explanans* to an *explanandum*. This is essentially the major topic of the minimalist program.

Notice now that the statement that UG constitutes the *explanandum* of minimalism must be located in a broader context. First, there is an important distinction, which becomes more marked in Chomsky's later work, between the faculty of language (FL) and universal grammar (UG); the former denotes a biological/cognitive system, and the latter refers to the theory of the initial state

of this system (i.e. those aspects of the system that are *genetically* determined). Second, there is the "three factors framework" as first outlined in Chomsky (2005); factor one refers to *genetic endowment* (the topic of UG), factor two concerns *linguistic experience*, and factor three denotes *general principles* that are not specific to language.[36] Now, one might correctly infer from the "three-factor framework" that UG remains a source of explanation (i.e. an *explanans*), for it constitutes one factor among others in determining the true nature of the language faculty. However, from a different perspective, it is also correct to say that UG constitutes an object of explanation with respect to factor three, in the sense that the various properties that have been regarded as falling within the explanatory domain of UG may now be deduced from the general principles under factor three.[37] In short, the shift in the theoretical role of UG is relative and in no way denies UG an explanatory role with respect to the faculty of language and its properties.

With this point clear, we now turn to the following question: if UG is the *explanandum* of minimalism, what are its *explanantia*? In other words, what are the tools of minimalist explanation? To find out, we need first to introduce some of the core assumptions of Chomsky's enterprise. One assumption is that the human mind/brain comprises certain faculties, one of which is a faculty dedicated exclusively to language, i.e. the FL.[38] In addition to this, the MP adopts the assumption that the FL interacts with its "neighbouring" cognitive systems, namely speech and thought systems, or, to use the conventional terminology, the articulatory-perceptual system and the conceptual-intentional system, respectively. Given this assumption, for the FL to be *usable*, it *must* possess properties that make the interaction with these systems possible.

The MP defines this interaction in terms of its *function* and *quality*; to wit, the purpose of the interaction is to satisfy the *interface conditions* imposed on the FL by the two systems, while the way in which it is achieved is assumed to be *optimal*. The interface conditions comprise so-called *legibility conditions*, which require that, whatever properties the FL "presents" to them, these must be *legible* to speech and thought systems. Put differently, these performance systems should be able to "read" the information provided by the FL (cf. Chomsky 2000b: 94). As to the assumption of optimality, the term "optimal" has a relativistic tinge; for when we describe something as being optimal, we imply that other relevant things are less optimal. In the case of language, the computational system is optimal in the sense that it opts for the best solution, among other less favorable alternatives, to the problem of satisfying legibility conditions at the interfaces.

We can now identify the *explanantia* of minimalism. The tools of minimalist explanation are interface conditions and optimal computation; both are explanatory in the sense that they constitute the basis from which the apparent complexities of UG are to be deduced.

From a minimalist perspective, language epitomises an "optimal design," for it is governed by elegant, economical, and simple laws of nature, an instance of which are the principles of computational efficiency. We will see later what these principles involve (Section 2.5.3). Suffice it to say here that the conception of language as exhibiting optimal design should not be confused with the idea of discovering the best theory of language (cf. Chomsky 2002: 97). Whereas the former refers to an intrinsic quality of the object of inquiry, the latter refers to an aesthetic feature of the theory under consideration. To put it in general terms, the contemplation of a beautiful natural world and the construction of beautiful scientific theories are two separate matters, although the former might sometimes serve as a hint as to how to proceed with the latter.

On the other hand, and despite this distinction between conceiving and approaching an object of inquiry, the MP seems to derive both its conception of language and its approach to it from a common source, namely *naturalism*. We shall have more to say about this topic in Chapter 6. Here suffice it to say that two minimalist theses originate from the idea that language is part of the natural world: on the one hand, there is an *ontological thesis* which holds that language is inherently as optimal as the objects comprehended by the laws of nature; and, on the other hand, there is a *procedural thesis* to the effect that mental aspects, including language, should be studied in the same way as any other aspects of the natural world (cf. Martin and Uriagereka 2000: 1; Atkinson 2005a: 23, and, to a lesser extent, Epstein and Hornstein 1999: xi; Hornstein *et al*. 2005: 7).

Some minimalists, through their emphasis on the idea that the field of linguistics is part of the natural sciences, are keen to define the minimalist notion of "optimality" in terms of the "tidiness" resulting from the laws of physics (see Fukui 1996; Uriagereka 1998; Freidin and Vergnaud 2001; Boeckx and Piattelli-Palmarini 2005, among others). More explicitly, what makes language optimal is precisely its compliance with the neat laws of nature. Bold as it might seem, this assumption is taken very seriously by the advocates of the MP, sometimes to extravagant excess (see Chapter 5 for a critical discussion of this assumption). But one thing is sure: minimalism has led to a substantial revision of many of the assumptions and theories which have been developed within the generative enterprise prior to its emergence. One aspect of this is the minimalist impact on the "design of language," to which we now turn.

2.5 Minimalism and the design of language

Chomsky (1971) gives an analogy of how the problem of language acquisition was conceived. He asks us to imagine an engineer who is faced with the problem of constructing a theory of the internal states of a device by studying its input–output relations. Thus the engineer is the linguist, and the device is a "language acquisition device," the input and output to which are the primary linguistic data and the postulated grammars, respectively. The analogy continues; the linguist does not quite understand the *function* of the device, but from the information about its input–output relations, he constructs a theory of the internal *structure* of the device. The linguist – to complete the analogy – observes the disparity between the input and output of the device, and concludes that, in order for the device to perform its function, it must have a rich internal structure (cf. the poverty of stimulus argument alluded to above). Thus the task for the linguist is to determine how rich the internal structure must be in order for the device to perform its function. This is what Chomsky calls retrospectively a "top-down" approach to UG; i.e. "how much must be attributed to UG to account for language acquisition" (Chomsky 2007b: 3).

In minimalism, however, the picture is conceptually different. We now have a "computational device" (i.e. a language faculty), the input and output of which are sets of lexical items and (legitimate) pairings of sound and meaning, respectively. From a minimalist perspective, that is, from a perspective which considers language to be a cognitive system, it is assumed that the *sole* function of the language faculty is to satisfy interface conditions imposed by other cognitive systems. Also from a minimalist perspective, that is, from a perspective which views language as a "natural object," it is further assumed that the faculty of language is "perfect" in performing its sole function. In testing these two assumptions, two design questions arise. First, what is the minimal structure required for the language faculty to perform its function? This question encapsulates what Chomsky (2007b) calls a "bottom-up" approach to UG; i.e. "How little can be attributed to UG?" The second question asks to what extent the language faculty is optimally designed to perform its function. This sort of question is familiar from optimization theory and has its philosophical roots in Leibnizian optimism.[39]

Now, minimalism does not start from scratch but attempts to move from the "top-down" to the "bottom-up" approach to UG.[40] To see what this task might involve, let us sketch the government and binding model of grammar as outlined in Chomsky (1981), and then proceed to show briefly how this model is affected by minimalist assumptions about language design. Needless

to say, there are many ongoing technical discussions on the nature of such a design, but the sketch which follows should be sufficient for the purposes of the present book.

2.5.1 The model of grammar: from GB to minimalism

The theory of government and binding (GB) includes a lexicon and four levels of syntactic representation. The lexicon provides the building blocks for sentence structures; it is the repository of lexical items and features. The four syntactic levels are: D-structure, S-structure, logical form (LF), and phonological form (PF). The first two levels are internal to the computational system. The system is regulated by various principles and modules, including the projection principle, case theory, theta theory, X-bar theory, etc. The GB approach embraces a wide range of theoretical notions, but as these are well known, it is not my intention to review their details here.[41] Rather, I shall turn immediately to consider how this framework differs from what we find in minimalism.

From a minimalist perspective, recalling that the FL is assumed to interact with two neighbouring cognitive systems, the only necessary levels are those that serve these interactions, *viz*. phonological form and logical form. Such justification does not attach to D-structure and S-structure, however; these, minimalists argue, are theory-internal and may be dispensable in principle. Now, since levels of representation are largely defined in terms of the principles that hold of them, there is clearly more at stake here than just a reduction of the number of postulated levels. What this indicates is that the minimalist program is committed to the elimination of these two levels without loss of either empirical coverage or explanatory power. As Hornstein (1995: 63) notes, this commitment involves a re-allocation of explanations framed in terms of at least some of the GB principles from D- and S-structure to the remaining two levels (i.e. PF and LF). One example is binding theory, which applies only at S-structure in some variants of the GB model. With S-structure eliminated from the system, Chomsky (1995a) has sought to show that some of the phenomena explained by binding theory can be comprehended by LF conditions.

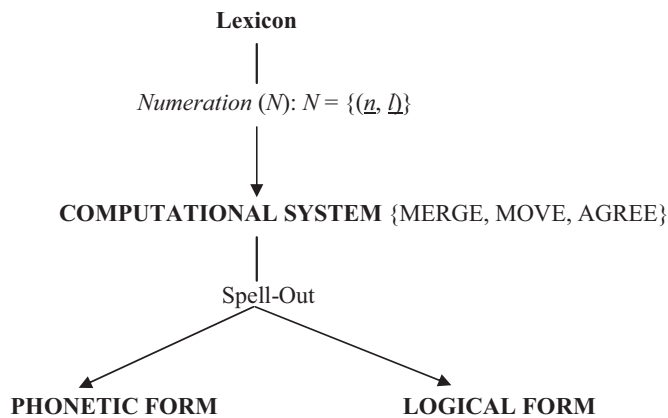
We have already observed in the previous section that the *explanantia* of minimalism include interface conditions and optimal computation. Supposing that these exhaust the minimalist's explanatory repertoire, Chomsky (2000b: 112–13) proposes that the faculty of language "provides no machinery beyond what is needed to satisfy minimal requirements of legibility and that it functions in as simple a way as possible." In other words, all conditions on the language system are either legibility conditions or conditions geared to achieve optimality somehow understood.

Generalizing with respect to legibility conditions, we might expect the computational system of language to operate only with lexical items the features of which are legible at the interfaces, i.e. we might anticipate that the "interpretability condition" suggested by Chomsky (2000b: 113) obtains. It is of considerable interest that this condition, as Chomsky himself (2000b) points out, is apparently false, since the computational system seems to rely on both interpretable and uninterpretable features (see below).

Supposing that interface representations are determinate functions of the lexical items from which they are derived suggests another condition, the "inclusiveness condition," which requires that no new features should be introduced in a computation mapping a set of lexical choices to these representations. Clearly, the consequences of this latter condition for the GB model are enormous: for instance, it involves the elimination of X-bar theory with all its references to phrasal categories and bar levels (Chomsky 2000b: 114).

Given the above assumptions about the design of language, and in contrast to the GB model of grammar, minimalists propose the model in (6), where the operation of the computational system is governed by principles such as interpretability (insofar as it can be maintained) and inclusiveness:

(6)



Additional comments on this sketch are in order. First, the lexicon is regarded as an indispensable component of the language faculty; it provides the "atoms" of computation. These are "features" of sound and meaning and the lexical items that are assembled from them. These features can be interpretable or uninterpretable at the interfaces (cf. the interpretability condition). Interpretable features comprise two obvious sets, phonetic and semantic features that are legible at the phonetic and semantic interfaces, respectively. Some formal, syntactic features

are also interpretable, but syntactic features include some that are uninterpretable, formal features that are illegible at either interface but are required to carry out computational operations. An example of the former is provided by agreement features, or φ-features (person, number, and gender) of nominals, with the same features of verbal elements being regarded as uninterpretable.

There are different proposals in the literature regarding how lexical items exit the lexicon and enter the computational system of human language (C_{HL}), but we need not go into these here. Suffice it to say, for the present purpose, that it is assumed that C_{HL} has access to the lexicon via some array of lexical choices or a *numeration* (*N*): $N = \{(\underline{n}, \underline{l})\}$, where $\underline{l}$ is a lexical item and $\underline{n}$ is an integer indicating the number of instances of $\underline{l}$ that have been selected from the lexicon. Then, given a numeration *N* of lexical items (LIs), the computational system C_{HL} maps *N* to a pair comprising a phonetic form and a logical form, each legible at the appropriate interface in the case of what are referred to as *convergent* derivations.

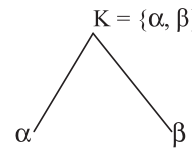
The computational system contains a number of operations that we will briefly consider in a moment. However, here it is appropriate to mention the operation of Spell-Out that must apply somewhere in the computation and have the effect of effectively partitioning the features into those that are interpretable at the two interfaces. This is often described as an operation that "strips off" the phonetically interpretable features leaving the semantically interpretable features (including those that are formal) for transmission to the LF-interface.

2.5.2 Merge, Move, and Agree

There is no consensus in the literature regarding the number of indispensable operations inside C_{HL}, but for our purposes we briefly consider here the three computational operations: Merge, Move, and Agree. Let us look at each in turn.

From a minimalist perspective, the ontological status of Merge is claimed to be justified on conceptual grounds alone (see Chapter 4, where the status of Merge is discussed in detail). It is suggested that, since it is indispensable in any language-like system, Merge "comes free" (Chomsky 2008a: 137). Merge combines two syntactic objects α and β to form a complex syntactic object K. This is illustrated in (7), with the proviso that the tree diagram is merely conventional and has no theoretical significance from a minimalist perspective:

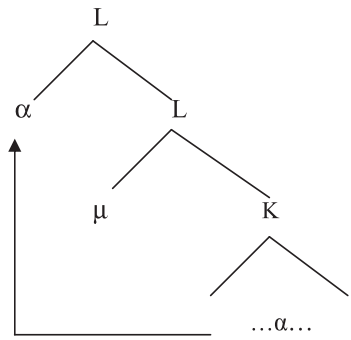
(7) *Merge* (α, β) → K = {α, β}:



Chomsky (1995a: 243) maintains that the value of K must reflect the fact that "verbal and nominal elements are interpreted differently at LF and behave differently in the phonological component." Consequently, he suggests that the value of K must at least include a *label* indicating the type to which K belongs. However, given the "inclusiveness condition" referred to earlier, Chomsky (1995a: 244) adds that the label of K "must be constructed from the two constituents α and β" and logical considerations lead him to conclude that the label of K "is either α or β; one or the other *projects* and is the *head* of K" (italics in original). Accordingly, the value of *Merge* (α, β) is K, which is either {α, {α, β}} or {β, {α, β}} (see Chomsky 2013 for his recent views on projection).

The second computational operation to consider is Move, an operation that displaces a lexical item from one structural position to another. This syntactic operation is schematised in (8), where a constituent α is shown to "move" from K-internal position to the specifier position of L:

(8) Move α



To give just one concrete example, ignoring the possibility that a chain of movements is involved and supposing that a copy of the moved item remains *in situ*, the derivation in (9) involves *what* moving from its direct object position to the specifier of the higher C:

(9) [[$_{CP}$What [$_{C}$ will][$_{TP}$ the cat eat what]]]

Movement is a ubiquitous property in natural language. Unlike Merge, however, the operation Move has not enjoyed ontological stability within minimalism. Initially, Move was regarded by Chomsky (1995a) as an "imperfection" because

of its absence in special-purpose symbolic systems. Additionally, it was seen as a "composite operation" involving a token of Merge and a token of Agree (see below), and this provided the basis for an argument that these latter operations pre-empt Move, except when movement is unavoidable:

> Plainly Move is more complex than its subcomponents Merge and Agree, or even the combination of the two ... 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. (Chomsky 2000b: 101–2)

However, several pages later, Chomsky provides an argument to the effect that (a) Move (as an operation expressing the dislocation property of natural languages) is motivated by interface conditions, and (b) uninterpretable features (regarded as another source of imperfection) are the mechanism which allows Move to operate. Thus Chomsky (2000b: 121) concludes that Move is no longer a straightforward imperfection.

The re-branding of Move is completed with its later elevation to become a "virtual conceptual necessary," and thus on a par with the operation Merge (Chomsky 2005). This ontological shift is justified by the proposal that Merge and Move are two sides of the same coin; the former merges α to β from the outside of β (i.e. External Merge), and the latter merges α to β from within β (i.e. Internal Merge). To have one without the other would require stipulation and prejudice the good design of the system.[42]

The last syntactic operation to be considered here is Agree. This relation is an abstract analogue of familiar agreement patterns and comprises an asymmetric relation between a "probe" and its c-commanded "goal," a terminology introduced by Chomsky (2000b, 2001, 2004a). At a certain point in a given derivation, a head-constituent (e.g. the functional heads v, T or C) serves as a "probe" which initiates a search for a goal (a DP) within its c-command-domain. While the head or probe enters the derivation with unvalued φ-features (e.g. uninterpretable person and number features), the DP enters the derivation with its φ-features already valued. Agree applies to such a (probe, goal) pair under an abstract notion of φ-feature identity with the consequence that φ-feature values from the goal are copied onto the probe.[43] The goal DP in turn has an unvalued case feature which is valued as nominative if the probe is T (or C) and accusative if it is v. These various suboperations together constitute the operation Agree.[44]

As to the ontology of the operation Agree, Chomsky (2000b: 101) speculates that it is a consequence of the need to satisfy "the design conditions for human language," because it "is language-specific, never built into special-purpose symbolic systems and apparently without significant analogue elsewhere." The route to establishing content for this speculation consists, at least partially, in recognising that a token of Agree is presupposed in any token of Move, i.e. the operation constitutes a necessary condition on movement and is part of the implementation of this process.[45] Supposing, then, that movement can be linked to interface requirements as hinted at above, then Agree too can be viewed as having credentials rooted in the fundamental *explanantia* recognized in the minimalist program.

2.5.3 Economy

So far we have said nothing about optimal computation, highlighted in Section 2.4 as one of the two major factors underlying minimalist explanation. Central to the concept of optimality is the notion of "economy," understood in terms of general constraints on representations and the derivation of syntactic objects. Economy of representation is illustrated by the principle of full interpretation (FI), which narrows the class of symbols appearing at PF and LF to those interpretable at the interfaces: there should be no redundant symbols in representations (cf. Chomsky 1995a: 151). As to economy of derivation, the operative constraints have a "least effort" and "last resort" flavour, such as "Procrastinate" and "Greed" (Chomsky 1995a). Procrastinate requires that whenever an overt application of an operation O is possible, then *covert* application of O is preferred. In other words, the principle has a preference for derivations which delay movements of items until after Spell-Out, in order that transformational effects do not reach the PF level. Greed is described as a "self-serving last resort," where an "operation cannot apply to α to enable some different element β to satisfy *its* properties" (Chomsky 1995a: 201).[46] In short, the intuitive idea behind the notion of "economy" is that language is highly non-redundant. "Insofar as that is true," Chomsky (1996: 30) says, "language seems unlike other objects of the biological world, which are typically a rather messy solution to some class of problems." In later chapters we will have more to say about the implications of minimalism for language evolution. Here, for concreteness, we shall offer a brief description of just one condition on derivations that has played a fairly central and consistent role in the development of the MP. This is the condition known as Shortest Move (Chomsky 1995a; Marantz 1995), which requires, at its name suggests, that movement of syntactic objects should be over

as short a "distance" as possible. To illustrate, consider the following example (cf. Chomsky 1995a: 181):

(10) (a) whom$_1$ did you expect t_1 [to feed whom$_2$]
(b) *whom$_2$ did you expect whom$_1$ [to feed t_2]

As (10) illustrates, movement of *whom$_1$* to [Spec, CP] is "shorter" than movement of *whom$_2$* to this position. As a result, Shortest Move licenses (10a) and blocks (10b). Of course, what is lacking here is a precise definition of syntactic "distance" – a technical matter that we shall ignore. What is important for present purposes is that minimalists see Shortest Move as subsuming various traditional constraints on movement. Thus Marantz (1995: 355) asserts that Shortest Move "takes over much of the work performed by Relativized Minimality (Rizzi 1990), Subjacency, and the Head Movement Constraint in earlier versions of the P&P theory." This view is in line with Chomsky's (1995a: 202) suggestion that, given Shortest Move, "we can incorporate aspects of Subjacency."

It is important to appreciate the sense in which these traditional constraints on movement may be incorporated into Shortest Move, for here we have an implicit indication of what it means for UG to be an object of minimalist explanation (cf. Section 2.4). Take, for instance, the superiority condition, proposed in Chomsky (1973) to constrain transformations. Put simply, Superiority is a condition which specifies that, when two *wh*-words appear in the same CP-domain, the structurally higher (i.e. superior) *wh*-word should be the one to move into the specifier position of CP. Taking the example in (10) as an illustration, since *whom$_1$* is superior to *whom$_2$*, the former must move into the specifier position of CP. Thus the superiority condition accounts for the well-formedness of (10a) and the ill-formedness of (10b). This demonstrates how the economy principle Shortest Move takes over the explanatory role of a traditional constraint on movement.

It may be felt that the above argument is trivial. However, once we take into consideration the ontological difference between economy conditions and traditional constraints on movement, the issue takes on a substantial air. Recall that the latter have been thought to be genetically determined, in the sense of being part of the initial state of the genetic component of FL. Indeed, in his debate with Piaget (to which we referred earlier), Chomsky (1980b) refers explicitly to traditional constraints such as subjacency and the specified subject condition as examples of the genetic content of UG.[47] By contrast, Shortest Move is assumed to be a general constraint on language whose effects extend beyond the sphere of the organic world. As Uriagereka (1998: 403) would have

us believe, economy conditions in language "might in the end follow from deeper properties of the universe at large." We shall return to this claim later (Chapter 5). Here let us merely observe that, if this is true, then it makes sense to seek to incorporate or derive the traditional constraints on movement from more fundamental principles of economy. It might be argued, for instance, that they are actually epiphenomena deriving from deeper principles of nature. This is certainly what Hauser *et al*. have in mind when they suggest that

> the generative processes of the language system may provide a near-optimal solution that satisfies the interface conditions to FLB. Many of the details of language that are the traditional focus of linguistic study [e.g., subjacency, Wh-movement, the existence of garden-path sentences . . .] may represent by-products of this solution, generated automatically by neural/computational constraints. (Hauser *et al*. 2002: 1574)

We will discuss this matter further in later chapters when we come to consider the empirical and conceptual plausibility of the strong minimalist thesis. Now it is time to turn to the last section of the present chapter and apply the *why*-question to the minimalist program itself.

2.6 Why minimalism?

The question in the title of this section has received little attention in the literature and, to my knowledge, has not been satisfactorily answered. As seen in Section 2.2, minimalism has been regarded by some of its proponents as an attempt at finding the best solution to Plato's problem. But we have already presented several arguments to the effect that minimalism cannot be reduced to an exercise in discovering the best explanatory model of language acquisition. In fact, even if we grant that the MP reduces to such an exercise, this will still not facilitate a satisfactory answer to the question of what has actually triggered the shift to minimalism. For it is a historical fact that almost a decade and a half separates the emergence of the P&P framework and the advent of minimalism, and thus one may ask why it took so long for the issue of finding the best P&P model to be raised.

It is interesting to note that, in their rejection of the MP, some critics have also appealed to a connection (or, rather, to a lack of a connection) between minimalism and its predecessors. For instance, some authors maintain that Chomsky's early frameworks provide no motivation whatsoever for the shift to minimalism, and argue that the program is motivated instead by Chomsky's "erroneous"

intuitions about the nature of language (Johnson and Lappin 1997, 1999; Johnson *et al*. 2000). Unfortunately, this criticism does not seem very helpful in providing an answer to the question that concerns us here. Indeed, it merely begs the question, for we shall have to ask what the motive behind Chomsky's intuitions might be.

A relatively recent argument from ideology has emerged as an explanation of the shift in Chomsky's thought. In their criticism of Hauser *et al*. (2002), Pinker and Jackendoff (2005: 30) attempt to explain the emergence of the MP in terms of Chomsky's conception of human nature, "in which people are innately equipped with an ability for spontaneous, creative, free expression, which is neither trained by society nor utilized in the service of some practical end." The authors seem to suggest that this conception underlies the apparent similarity between Chomsky's views on politics and his views on linguistics. For instance, they claim that Chomsky's "anarcho-syndicalism," with its emphasis on the tendency of people "to cooperate and to engage in productive, creative work for its own sake," is comparable to his generative system "which allows for the expression of thought for its own sake but is not designed for ... the practical function of communication" (Pinker and Jackendoff 2005: 31). They further argue that failing to take this connection into account will make Chomsky's minimalist views, and especially his views on the evolution of language, appear "capricious." To see how this connection is supposed to explain the shift in question, consider what these authors have to say:

> In the first decades of the cognitive revolution, a vague notion of innateness was sufficient to distinguish Chomsky's ideas from those of the behaviorists and other empiricists. He could point to a set of properties that distinguished language from generic learned behavior, such as its complexity, modularity, expressive power, and uniqueness among species. But with the rise of evolutionary psychology in the 1980s and 1990s, the *origin* of innate abilities began to be scrutinized. According to modern biology, complex innate traits arise because they were useful to the organism's ancestors. This focus reveals a tension between a vision of human nature in which innate traits are exercised for their own sake and a Darwinian explanation in which innate traits evolved for their fitness benefits. Chomsky apparently has responded to this tension by emphasizing the recursive generative capacity that is at the heart of his vision of human nature and distancing himself from the features of language that call for a Darwinian explanation, namely, adaptive complexity in the service of communication. Thus language,

> for him, is not designed for communication, and the parts of language that had to evolve in humans are so minimal that invoking selection is unnecessary. (Pinker and Jackendoff 2005: 31)

What is suggested here is that the rise of evolutionary psychology has provoked a tension between the Chomskyan conception of human nature and the Darwinian conception of trait evolution; the former conceived of innate traits as being "exercised for their own sake" (whatever the authors might mean by this), the latter as being constrained in their evolution by their usefulness for the species. In his attempt to avoid a Darwinian explanation of how language might have evolved – so the argument seems to run – Chomsky has sought to deprive natural selection of its force by shifting his view of language from a complex system that has evolved for communication to a simple system that has evolved for the expression of thought.

How are we to respond to this argument? To begin with, notice that Pinker and Jackendoff view the shift in Chomsky's views as involving, *inter alia*, (a) an emphasis on the recursive power of language, and (b) a departure from the view that human language is a complex adaptation that has evolved for communication purposes. Observe further that they consider "the rise of evolutionary psychology in the 1980s and 1990s" as the starting point for the shift in Chomsky's views; indeed, this is a crucial proposition on which their argument is based. Now if these changes in Chomsky's thinking were properly ascribed, we would expect: (i) prior to the 1980s, there was no emphasis in Chomsky's work on the recursive power of language; (ii) prior to the 1980s, Chomsky did not distance himself from the position that language constitutes a complex adaptation for the purposes of communication. However, the former proposition is clearly false, as has been shown in Section 2.3. As to proposition (ii), Chomsky has never, to my mind, aligned himself with an adaptationist position. In fact, the contrary is the case, for as early as the 1960s Chomsky affirmed "the hopelessness of the attempt to relate human language to animal communication" and saw "no substance to the view that human language is simply a more complex instance of something to be found elsewhere in the animal world" (Chomsky 1968: 69–70). Consequently, the claim that the rise of evolutionary psychology has anything to do with the shift in Chomsky's views is unfounded.

However, rejection of such an argument should not prevent us from examining the possibility that other considerations about the evolution of language may be the real drivers of the shift to minimalism. For instance, Chomsky has been explicit in adopting a saltational view on language evolution, according to

which the emergence of language represents a sudden event with no intermediate stages. It may not be unreasonable to suspect that such a saltational view underlies the shift to minimalism. Indeed, in one of his fairly recent public talks, Chomsky (2008b) presented an argument which seemed to suggest that evolutionary considerations (such as saltationism) had exactly this role:

> [The P&P approach] raised another question: What about the principles [of UG]? Where do they come from? ... If they're in universal grammar, if they are part of the genetic endowment, then they had to evolve somehow. But not a lot could have evolved because it's too recent ... what evolved in that short period of time cannot have been very complex ... Therefore, what you predict is that some other principle external to language, maybe some principle of nature, principle of computational efficiency ... interacted with a small mutation which just gave rise to the universal grammar. *Well, that sets forth a new goal of research ... to see if you can determine that the principles [of UG] do not have the intricacy that they appeared to have*, but are actually the result of application of non-linguistic, in fact, maybe non-human, like general principles of computational efficiency, to whatever small change took place [in the brain]. And the small change was probably the capacity to carry out recursive enumeration. (Chomsky 2008b, my italics)

As is made clear, the perspective favoured in this passage gives rise to a research program (i.e. minimalism) the aim of which is to exhibit the extent to which the apparent complexity of UG is an epiphenomenon of the interaction between the genetically determined capacity of recursion and general principles of nature.

It is interesting to observe that this argument (from saltationism) for the emergence of minimalism fits nicely with Chomsky's epistemological nativism, there being an apparent isomorphism between the problem of language acquisition and the problem of language evolution. First, the time factor is essential to both problems; the time during which the child acquires her language, and the time through which language has evolved are both judged to be "short." Second, these time factors militate against the importance of external factors, *viz*. linguistic experience for the former, and natural selection for the latter. Finally, both problems have led to a reduction in the "size" of UG; one in terms of the removal of specific rules, and the other in terms of the shedding of specific principles.

Despite the above, I believe that it would be rash to conclude that the attractiveness of saltational evolution holds the key to the shift to minimalism.

One reason for caution here is that Chomsky appears to have been attracted by a saltational view on evolution as far back as the late 1960s, and if this is so it is not clear why it has taken such a long time for the minimalist program to emerge. Consider, for instance, the following passage from one of Chomsky's early writings, in which the "true emergence" of language "at a specific stage" is seen as constituting a problem for biology:

> There seems to be no substance to the view that human language is simply a more complex instance of something to be found elsewhere in the animal world. This poses a problem for the biologist, since, if true, it is an example of true "emergence" – the appearance of a qualitatively different phenomenon *at a specific stage* of complexity of organization. (Chomsky 1968: 70, my italics)

The saltationist perspective in this passage is unmistakable, and, therefore, as far as the shift to minimalism is concerned, there is more at stake than just an adoption of saltationism. To further substantiate this point let us consider what this passage is saying in a broader context. If language exhibits unique properties that are not found in non-human communicative systems, the question arises as how these "emergent" properties came into existence. Here, this constitutes a problem for biology from Chomsky's perspective. But notice that such a problem is further aggravated when we take into consideration the pre-minimalist standpoint – a standpoint which saw language as comprising a *significant number* of emergent properties. Chomsky seems to have reacted to this by adopting a saltational view in which "random mutations" might have been responsible for the emergence of a complex and unique system such as human language. For instance, in his debate with Piaget he says:

> Although it is quite true that we have no idea how or why random mutations have endowed humans with the specific capacity to learn a human language, it is also true that we have no better idea how or why random mutations have led to the development of the particular structures of the mammalian eye or the cerebral cortex. (Chomsky 1980b: 36)

Notice that the concern here is not whether random mutations have led to the development of language – presumably this was taken for granted – but rather how and why these mutations have led to such a development. We see, then, that Chomsky's old commitment to saltationism could not have been responsible for the more recent shift to minimalism.

Observe further how this passage differs from the one cited earlier (p. 46), where Chomsky argued that language (*qua* complex biological system) could not possibly have evolved in a relatively short period of time, and concluded that much of the apparent complexity might not have been the result of evolution. In short, we witness here a significant shift, one that moves away from explaining the apparent complexity of human language in terms of random mutations toward an explanation in terms of an interaction between a single mutation and laws of nature.[48] It might be worth pointing out that this latter position is more readily consistent with saltationism; if we have a (largish) set of random mutations, we are presumably looking for something miraculous if they all occur near-simultaneously. Such a position could amount to Chomsky having always favored saltationism but, in the absence of better alternatives, having to subscribe to a rather outlandish view involving random mutations (we will return to this point shortly).

But if adoption of saltationism is not the answer to the question "Why minimalism?", what is the answer? There is reason to believe that the emergence of minimalism was motivated by developments that have taken place during the 1980s in the fields of biology and neuroscience. This is not the place to go into historical details on this issue, and I shall limit myself here to giving some examples as an illustration of the possible connection between the development of the mentioned fields and the shift to minimalism.

The rise of developmental biology in the late 1970s has emphasized the role of developmental constraints in evolution against the traditional emphasis on natural selection. One example is provided by Maynard Smith *et al*. (1985), a well-known paper in the field of evolutionary theory, in which the authors argue that evolutionary constraints are not limited to selective constraints but also include developmental constraints that follow from the laws of physics. One such constraint, which is regarded as a consequence of the physical law of the lever, is that "any uncompensated change in the shape of a skeleton that increases the speed with which some member can be moved will reduce the force which that member can exert" (Maynard Smith *et al*. 1985: 267, quoted in Jenkins 2000: 191).

In the case of neuroscience, the 1980s witnessed the emergence of the field of computational neuroscience, which focuses on the study of neural networks of the brain *qua* biological-computational system.[49] A key notion in this field is that of "optimal wiring" which, in the view of some authors, illustrates a possible connection between the physics of the brain and its anatomy (see, for instance, Ringo 1991; Cherniak 1994; Cherniak *et al*. 2002). The basic assumption here is that the brain's neural structure is optimal with respect to the total length of its

neural "wire" connections; i.e. the shorter wiring a neural structure has, the more optimal it is. The computational neuroscientist Christopher Cherniak, whose work is cited by Chomsky, argues that the internal organization of the brain is largely constrained by the principle of "save-wire" – a principle that he regards as a direct consequence of the laws of physics – which ensures efficient connectivity in neural networks (Cherniak 1990, 1994, 2005).[50]

It is not unreasonable to suspect that the shift to minimalism was largely motivated by developments such as those outlined above. This, if true, will not be the first instance in which Chomsky's work has been influenced by developments in other fields. Two examples come to mind. The first, mentioned earlier in this chapter, is that early transformational approaches to syntactic theory were influenced to a limited extent by Carnap's logical analysis of language and Post's formalization of proof theory. Another example is Chomsky's P&P approach to language acquisition, which he himself (1980a: 67) maintains was suggested by François Jacob's ideas on the diversification of organisms.

I should perhaps make clear that the point I am trying to make here is that certain developments in the fields of computational neuroscience and developmental biology may have suggested a way of *how* physical constraints might play a significant role in the evolution and development of organisms. By this I am by no means suggesting that, before these developments took place, Chomsky had never alluded to a possible role of physical constraints on evolution. In fact, there is ample evidence in Chomsky's early work indicating that he was indeed explicit in referring to (unknown) physical laws underlying the development of biological systems. For instance, in the famous debate to which we referred earlier, and in response to Piaget's remark that the evolutionary development of language in terms of random mutations was "biologically inexplicable," Chomsky (1980b: 36) says:

> Little is known concerning evolutionary development, but from ignorance, it is impossible to draw any conclusions. In particular, it is rash to conclude either (A) that known physical laws do not suffice in principle to account for the development of particular structures, or (B) that physical laws, known or unknown, do not suffice in principle. Either (A) or (B) would seem to be entailed by the contention that evolutionary development is literally "inexplicable" on biological grounds. But there seems to be no present justification for taking (B) seriously, and (A), though conceivably true, is mere speculation.

This is by no means the only example which could be quoted in this context; other relevant passages can be found in Chomsky (1965: 59, 1968: 97–8, 1975a:

9–10, 1980a: 6, 1988: 97). Thus, we may fairly say that the possible relevance of physical law to the evolutionary development of the language faculty had been acknowledged by Chomsky long before the advent of the MP. To put it differently, although the two factors of genetic endowment and environment have been in the forefront of Chomsky's theory of language, a third factor involving, *inter alia*, the workings of physical laws has been in the background of his thinking for many years. Interestingly, when asked why, given his early recognition of the notion of physical law, he waited so long before explicitly introducing the third factor in his linguistic theorizing, Chomsky (p.c.) replied: "There wasn't much to do with the third factor then, or until the minimalist program began to take off in the mid-1990s." It would be reasonable to infer from this that, with the advent of minimalism, a theoretically workable way to exploit the third factor became available. If this is true, then it is consistent with our suggestion that recent developments in biology and neuroscience have facilitated the incorporation of general physical principles into Chomsky's theoretical framework.

Here, perhaps, lies the answer to our "why" question. Chomsky probably never really believed in random mutations being miraculously responsible for the emergence of an intricate system of linguistic knowledge. Yet, to return to a point made earlier, he had no choice but to bite the bullet and insist that the evolutionary development of such an intricate system might be "biologically unexplained," though not "biologically inexplicable" (Chomsky 1980b: 36). Now, drawing on some insights from other fields, he has introduced what he believes offers the prospect of "true explanation" of language and its origin. I am referring here to his strong minimalist thesis (SMT), a thesis which is ultimately concerned with providing an answer to the question of how it is possible for a biological system such as human language to have both an apparent structural complexity and a recent evolutionary history. But what does this thesis amount to? And how does it relate to the evolution of language? These questions lead us into the subject of Chapter 3.

3 *The strong minimalist thesis (SMT)*

3.1 Introduction

As observed in the previous chapter, minimalism can be regarded as an exploration of the view that the complexity of language is only apparent, and that there are in fact deeper principles from which much of this complexity can be derived. This perspective is embodied in the strong minimalist thesis (SMT), which, in one formulation, states that language is an optimal solution to the problem of satisfying interface conditions (cf. Chomsky 2001: 1). This chapter takes a closer look at the content of the SMT, and how it is supposed to account for the apparent complexity of language. This is necessary before we embark on a detailed assessment of the thesis – a task that we undertake in subsequent chapters.

Some may be inclined to question whether the effort we are about to make is worthwhile; after all, to the extent that the notions of "optimal solution" and "interface conditions" can be given specific content, there should be no problem in principle in understanding what the content of the SMT amounts to. However, there are at least three reasons to suggest that this is not the case.

First, and as this chapter will seek to demonstrate, the question of how the SMT should be interpreted is not straightforward. This is because Chomsky's approach to the apparent complexity of language over the last fifteen years has not been uniform; indeed, a careful examination of his work will reveal three different sorts of emphases, linked to different formulations of the SMT, at different points in his writings – a fact that, to my knowledge, has been explicitly recognized by neither Chomsky nor anyone else in the field.[1] It is thus desirable to clarify what otherwise might be a source of confusion and even inconsistency. This will be our major task in the present chapter.

Second, the notion of "virtual conceptual necessity," despite being central to minimalism in general and to the SMT in particular, has not received careful consideration in the minimalist literature – a fact evidenced by the different and even contradictory views expressed on this notion, as we shall see. It is,

therefore, necessary to subject this fundamental notion to closer examination to see whether it throws light on the nature of the SMT.

Third, for a proper understanding of the SMT, one has to take into consideration Chomsky's extralinguistic discourse, namely his contributions to the "evolution papers" (Hauser *et al*. 2002; Fitch *et al*. 2005). For whatever interpretation one may attach to the SMT, it must be consistent with the hypothesis that recursion is the only property that is unique to language and to humans, a hypothesis that constitutes the central claim of these papers. A secondary aim of the chapter, therefore, is to bring together Chomsky's two discourses in order to see how one discourse might inform the other. However, a full discussion of this topic must wait until the next chapter.

It is suggested here that the three formulations of the SMT are perhaps best viewed in terms of three different approaches that Chomsky has probably contemplated at different times over the past fifteen years. The first approach gives rise to an interpretation of the SMT as a strict generalization – an interpretation which I shall argue is difficult to justify. The second approach is dominated by what I will call here the "imperfection strategy," a strategy which I will argue can be misleading, especially when it comes to the cogency of the SMT. The third approach may be regarded as a shift in strategy in Chomsky's take on the SMT, embedding it in what will be termed the "three factors framework."

The chapter is organized as follows. We begin with a demonstration that various views on the notion of "virtual conceptual necessity" are both inconsistent with each other and incongruent with Chomsky's (early) usage of this notion (Section 3.2). We then proceed to discuss one formulation of the SMT, namely that in which it is viewed as a strict generalization (Section 3.3). After having dealt with the first approach to the SMT, we turn to the second one by introducing the "imperfection strategy" and showing how this approach differs from the previous one (Section 3.4). This will allow us to re-examine the notion of "conceptual necessity" (Section 3.5), and to identify certain shortcomings in the second approach (Section 3.6). Finally, the last two sections of the chapter deal with the third approach to the SMT. Section 3.7 presents the three factors framework and shows how it differs from the earlier approaches, and Section 3.8 examines the extent to which Chomsky's linguistic and interdisciplinary discourses agree with, or differ from, each other, and sets the stage for the next chapter.

3.2 Conceptual necessity: a first encounter

The notion of "virtual conceptual necessity" (henceforth VCN), despite being central to minimalism, has not been treated with an appropriate standard of care,

and the interpretations it has received in the literature contrast starkly with each other. Let me substantiate this claim.

Grohmann (2003: 10) identifies VCN with interface conditions by suggesting that what this notion "dictates is that all conditions on the computation follow from Bare Output Conditions," that is, "conditions that relate directly to the conceptual-intentional and articulatory-perceptual interfaces." Langendoen (2003: 307), however, associates VCN with the notion of optimal computation, or, more specifically, with "general considerations of simplicity, elegance and economy." Yet others define it in terms that refer neither to optimal computation nor to interface conditions. Thus, Hornstein *et al*. (2005: 6) define VCN in terms of what they regard as "big facts," that is, "those facts about language that any theory worthy of consideration must address." One such fact, the authors argue, is that sentences are composed of smaller units like words and phrases, a fact that is manifested by the conceptually necessary operation Merge. Thus the authors maintain that "Merge is conceptually necessary given the obvious fact that sentences are composed of words and phrases" (Hornstein *et al*. 2005: 207).

In his foreword to Uriagereka (1998), Piattelli-Palmarini (p. xxxiv) asserts that the minimalist program "vastly expands the bounds of 'virtual conceptual necessity' (i.e. of what about the basic design of human languages must be as it is because it could not possibly, conceivably, be otherwise)." And Smith (2004) makes a contrast between what is "conceptually necessary" and what is "empirically unavoidable"; he defines "conceptual necessity" as that which it "is impossible to do without," and as examples of this, he cites, *inter alia*, the lexicon and the two interface levels, PF and LF (Smith 2004: 84). In his foreword to Chomsky (2000a) Smith comments on the minimalist question "How 'perfect' is language?" by saying "that any deviations from conceptual necessity manifest by the language faculty ... are motivated by conditions imposed from the outside" (Smith 2000: xii). Given the contrast Smith makes between what is conceptually necessary and what is empirically unavoidable, this comment amounts to saying that any deviation from conceptual necessity is motivated by empirical necessity.

In sharp contrast to all of the above, Boeckx understands VCN in terms of the contingent state of current inquiry, for he says that this notion "refers to what appears to be necessary at the present stage of understanding," adding that "everything we now know is subject to change" (Boeckx 2006: 75). This conception of VCN seems to be extremely implausible; I take it that for Boeckx, any rewriting rule from the early period of generative grammar must have been a conceptual necessity in that period, only to shed this necessity a few years

later! More importantly, none of the above conceptions of VCN seems to be in tune with the following passage from Chomsky (1995b [1994]: 385–6), in which the notion in question appears for the first time:

> [(i)] What conditions on the human language faculty are imposed by considerations of virtual conceptual necessity? [(ii)] To what extent is the language faculty determined by these conditions, that is, how much special structure does it have beyond them? The first question in turn has two aspects: what conditions are imposed on the language faculty by virtue of (A) its place within the array of cognitive systems of the mind/brain, and (B) general considerations of simplicity, elegance and economy that have some independent plausibility.

The modifier "'virtual" has received various interpretations, but since it has no direct impact on the present discussion, I propose to leave it aside and focus just on "conceptual necessity."[2] This necessity, according to Atkinson's (2005a) interpretation of this passage, stems from two ways of conceptualizing language: one as a cognitive system embedded in other such systems, and another as a natural object.[3] More specifically, if the language faculty is viewed as a system that generates linguistic expressions, and if the information encapsulated in these expressions is available to speech and thought systems, it follows that it is a conceptual necessity that these two systems have access to the information provided by the language faculty (cf. Chomsky 2002: 108). On the other hand, if human language is viewed as a natural object, and if the natural world *in toto* is governed by universal laws (a premise without which science could hardly be possible), then it is a conceptual necessity that the laws which govern the natural world also govern human language; hence the general considerations of simplicity, elegance, etc., to which the passage refers. Recall from the previous chapter (Section 2.4) that the naturalism advocated by Chomsky is not merely methodological but also ontological, so one is justified in linking simplicity considerations with the notion of "natural object" despite the fact that only the former is explicitly mentioned in the passage above.

Adopting Atkinson's interpretation, we now can see why the above conceptions of VCN differ from what we have in the passage quoted above. Here, we have two routes to conceptual necessity: interface conditions and optimal computational. Grohmann fails to refer to the latter, Langendoen fails to refer to the former, and Hornstein *et al*., Piatelli-Palmarini, and Smith all refer neither to the former nor to the latter. Presumably, in providing a definition of "conceptual necessity," these latter authors are not referring to a certain way

of conceptualizing human language; rather, they seem to refer instead to the concept of language as such.

I will argue later (Section 3.5) that this latter conception of language as such should be viewed as a third route to conceptual necessity, one which becomes more apparent at a later stage in Chomsky's work. Here, I argue that although Atkinson's two-route interpretation of VCN is based on an understandable reading of the passage cited above (p. 54), it has implications that are not compatible with some fundamental minimalist assumptions. As we shall see in Section 3.5, these implications disappear once we recognize the third route just mentioned. Before we proceed further, however, let me say a little more about the passage in question.

Consider the second question this passage raises. It seems that what we have here are two complementary questions. The more positive the answer to the question "To what extent is language determined by conditions that arise from considerations of conceptual necessity?" the more negative the answer to the question "How much special structure does language have beyond these conditions?" and *vice versa*. In other words, we have here a contrast between what is *special* to language as opposed to what follows from general considerations of conceptual necessity – in the former case by virtue of language being a species-specific capacity, and in the latter case either by virtue of its place among other cognitive systems, or by virtue of its being an "object" in the natural world. The reader will recall from the previous chapter (Section 2.4) that minimalism attempts to reduce the complexity of universal grammar (UG) by shifting the burden of explanation of core aspects of language from *genetic* constraints to general principles that are not specific to language. Accordingly, I will assume – throughout the discussion to which we now turn – the contrast between what is special to language and what follows from general considerations of conceptual necessity to be a contrast between what is genetically determined in language and what follows from general principles that are not specific to language (i.e. principles that determine its place in both cognition and the natural world at large).

3.3 SMT as a strict generalization

On the basis of the two-route interpretation of VCN, Atkinson (2005a: 17) suggests that if it turns out that the faculty of language "exhibits a large range of properties that are not 'virtually conceptually necessary,' . . . we are going to have to conclude that minimalism is incorrect." He believes that this view is supported by the following formulation of the SMT, which appears in

Chomsky (2001: 1): "The strongest minimalist thesis would hold that language is an optimal solution to [legibility] conditions." Since this formulation makes no reference to the contribution of genetic endowment to the determination of the structure of the language faculty, and since reference is made only to the two aspects of conceptual necessity as described above, Atkinson (2005a) concludes that the SMT, as formulated here by Chomsky, suggests that *all* language properties are fully determined by legibility conditions at the interfaces and/or general principles of optimal computation; i.e. beyond these general considerations of conceptual necessity, the language faculty has no special structure. If this is correct, then it is easy to see what problematic implications such an interpretation might have for some fundamental minimalist assumptions.

To begin with, it plainly goes counter to the claim that there is *at least something special* about language – a claim that, as argued in the previous chapter (Section 2.3), Chomsky never ceases to defend. This recognition of there being something special immediately gives rise to another problematic implication, namely that the SMT is *untenable a priori*, and, therefore, cannot be regarded as an empirical thesis; i.e. if we accept the fact that there must be something special about language, then the SMT is, *ipso facto*, false and no empirical research is required so as to find out whether all language properties derive from general considerations of conceptual necessity (as defined above). Evidently, this stands in direct conflict with how the SMT is viewed by Chomsky and his followers. From a minimalist perspective, the SMT is considered to be an explanatory-empirical thesis; it is explanatory because it purports to shift the burden of explanation from genetic endowment to general principles that are not specific to language, and it is empirical because its rejection must be based on evidence indicating, *inter alia*, that the general principles fail to account for the apparent complexity of UG.

Notice further that even if we set aside the issue that something must be special to language, the interpretation of the SMT as a strict generalization, i.e. as a thesis which asserts that all language properties are determined by external considerations (in the sense described above), seems to contradict what Chomsky says here:

> SMT, or a weaker version, becomes an empirical thesis insofar as we are able to determine interface conditions and to clarify notions of "good design." While SMT cannot be seriously entertained, there is now reason to believe that in nontrivial respects some such thesis holds. (Chomsky 1995b [1994]: 386)

Clearly, Chomsky allows here for a weaker thesis than the unqualified SMT, a version that need not be interpreted as a strict generalization. In fact, if one insists on viewing the SMT as a strict generalization, then one is forced to conclude that the thesis is vacuous; it amounts to the suggestion that *all* language properties are determined by external considerations *except when some aren't so determined* (cf. Atkinson 2005b: 203).

One might try to mitigate the force of this conclusion by observing that exceptionless universal generalizations are rare outside the boundaries of (an ideally completed) physics and that most (if not all) law-like statements in special sciences are hedged with a *ceteris paribus* clause. Thus the expression "*all else being equal*," when present in a law-like statement, indicates a generalization weaker than that expressed by an unqualified law-like statement. In view of this, and accepting the respectability of such qualified laws in the special sciences, it might be argued that there is no reason to be alarmed about what the passage quoted above says, for what we have here is simply an acknowledgement that a thesis weaker than the SMT might be correct, but this in itself does not undermine the generalization contained in the SMT, any more than *ceteris paribus* clauses undermine the status of the law-like statements to which they are attached in the special sciences. However, closer inspection will reveal that this analogy is not only misplaced, but also leads to a further implication that is highly counterintuitive. To see this, we need first to say a little more about *ceteris paribus* clauses (henceforth cp-clauses) in the special sciences.

The quest for exceptionless regularities in nature lies at the heart of scientific inquiry in general. But to come up with a strict law with no exceptions outside an ideally completed physics is both rare and difficult. Consider, for instance, the following often-cited example from the field of economics: at a constant rate of supply, if demand increases, prices rise. At first glance, we might think of this statement as expressing a strict law. Nevertheless, it is not difficult to see that this is not the case, for there are various easily conceivable circumstances in which this economic generalization fails. Think, for instance, of the case where a central government controls prices, or greed-free people manage all kinds of trade, etc. It appears, then, that a cp-clause is needed to supplement the "law" of supply and demand. But since cp-clauses allow exceptions, there will always be the risk of having no law at all. In other words, we are led to the conclusion that the law of supply and demand is correct, except when it isn't.

Encapsulating the concerns of the previous paragraph, Fodor (1987: 5) remarks: "Nothing that happens will disconfirm [a cp-law]; nothing that happens could." However, he goes on to defend the content of the generalizations of special sciences, thus there is no charge of vacuity coming from him. That laws

in the special sciences can have genuine content (and, indeed, be used to formulate predictions) while containing cp-clauses is regarded as a major dilemma by Fodor, a dilemma he seeks to resolve relying on ideas that we shall shortly advert to (see Fodor 1975: 9–26, 1987: 1–26).

Despite the obvious similarities that we see between the import of the passage quoted above and laws in the special sciences, we shall now argue that there are also important differences here, differences that illustrate why the suggested analogy between a weaker version of the SMT and laws including cp-clauses is not only flawed, but also counterintuitive.[4]

As Fodor (1987) has pointed out, one of the reasons why a special science is special lies in the fact that the exceptions to its generalizations are inexplicable in the vocabulary of that science. Instead, we have to rely on the vocabulary of some other, more basic, science to spell out the content of these exceptions. As an example of this, Fodor (1987: 5) gives the following generalization from the field of *geology*: "*all else being equal*, a meandering river erodes its outside bank." He then points out that what is covered by the cp-clause here does not normally refer to *geological* events:

> Try it and see: "A meandering river erodes its outside banks unless, for example, the weather changes and the river dries up." But "weather" isn't a term in *geology*; nor are "the world comes to an end," "somebody builds a dam," and indefinitely many other descriptors required to specify the sorts of things that can go wrong. (Fodor 1987: 6)

But the number of "things that can go wrong" can be indefinitely large; hence the charge of vacuity that might be levelled against this *geological* generalization. Yet *geology*, like all other special sciences, relies upon its cp-clauses in formulating its generalizations, and, in fact, the role of these clauses in psychology and cognitive science, key areas of interest for Fodor, appears to be even more crucial. Accordingly, his intention is to seek to save statements containing cp-clauses from triviality by developing a view on the truth conditions of *ceteris paribus* laws.[5] Whether Fodor's proposals to this end are on the right track will not concern us here. What is of importance for the purposes of the present discussion is to find out how far this conception of cp-clauses and their roles in special sciences differs from what we find in the SMT.

Let us go back to our earlier example from the field of economics: at a constant rate of supply, if demand increases, prices rise. We have already seen that this statement cannot be taken as expressing a strict law, the reason being that one can think of factors that can ensure that the statement is false. As examples of such factors, we mentioned "government control," and "greed-free

traders." Now, the reader would agree that these expressions are most plausibly assigned to the vocabularies of *politics* and *psychology*, respectively. Given our example, then, we see that, although the two falsifying factors are covered by a cp-clause, they do not constitute a natural class, since it is clear that, in order to properly describe them, we would have to rely on the vocabulary of a range of disparate disciplines. Indeed, this is also the case in many (if not all) cp-clauses in special sciences – think back to Fodor's own example of the meandering river. Now, it is immediately apparent that what might appear in a qualified version of the SMT is different from what we find covered by cp-clauses in special sciences; the language properties that might not be determined by external considerations of interface conditions and optimal computation, i.e. the putative exceptions to the SMT, constitute some sort of a natural set in the sense that they all are manifestations of biological necessity, that is, we need only resort to the *vocabulary of genetics* in order to properly state them.[6] In a nutshell, one reason why the suggested analogy is misguided is this. The exceptions to the SMT constitute a natural set, whereas those covered by cp-clauses in special sciences do not. There is a further reason for cleaving a distinction here.

As already noted, Fodor (1987) suggests that one of the reasons why a special science is special has to do with the fact that the exceptions to its generalizations are not describable in the vocabulary of that science; any description of the content of these exceptions relies on the vocabulary of some other science. Moreover, if we are to make *causal* sense of the role of the cp-clause, it will be necessary to move to a more basic science in which events from both the special science and the science of the cp-clause can be given descriptions. Let us assume, for the sake of argument, that this is possible, at least in principle. Now it is not difficult to see that this way of dealing with exceptions is quite different from what we find in the SMT. For one thing, reliance on general principles that are not specific to language is considered by minimalists to provide the "deepest" level of explanation, at which the laws of physics might be operative; indeed, any language property that can be explained in terms of these principles is regarded by Chomsky and his followers as having a *principled explanation* (see Section 3.7 below). If we now grant that the SMT might have exceptions, and if all these exceptions can be explained in terms of *genetics* (recall the assumption at the end of the previous section), we would then end up with an awkward and counterintuitive situation, namely that the exceptions to the laws of *physics* would be being explained in terms of a special science such as *biology*! Notice that any attempt to go beyond *genetics* and restate the exceptions in terms of *physics* would only take us back to square one,

that is, the SMT is a strict generalization without any exceptions. Undesirable, indeed, perhaps incoherent, consequences such as these indicate that the suggested analogy between a weaker version of the SMT and what we find covered by cp-clauses in special sciences cannot be maintained.

Before we conclude the present discussion, there is one further point that needs clarification. Why should Chomsky think that "SMT cannot be seriously entertained"? One may be inclined to suggest that the answer to this question lies in the fact that something must be special to language. But this presupposes that the correct interpretation of the SMT is that there is nothing special to language. However, we have already argued that this interpretation commits us to an *a priori* rejection of what is supposed to be an empirical thesis (i.e. the SMT). A more plausible answer seems to lie in the following remark by Chomsky (p.c.):

> [SMT] is a very bold hypothesis, and while it might be true, my own expectation is that the world is unlikely to be as elegant as that. Others disagree – appropriately for an empirical hypothesis. My own view is that for now at least it should be understood as providing guidelines for research, seeking to determine how closely it can be approximated, and sharpening it along the way.

Thus, "SMT cannot be seriously entertained," not because at least one language property must be "special" in that it is determined by genetic endowment, but because it would be too extraordinary for a biological system such as language to be completely efficient in using its resources to link sound and meaning. To put it concisely, the reason why Chomsky believes that a thesis weaker than the SMT might be true does not relate to the fact that language must be special, but rather it follows from the expectation that language is unlikely to be as perfect as the SMT prescribes.

The preceding discussion has sought to show that Atkinson's interpretation of the SMT as formulated by Chomsky has implications that are not compatible with some fundamental minimalist assumptions. A natural question to ask at this point is: should we then abandon such an interpretation? That this interpretation might be inconsistent with Chomsky's position is hardly a good reason for ruling it out. What is important is to guard against the tacit assumption that Chomsky has always been consistent in expressing his views. Indeed, it may turn out that the phrase "virtual conceptual necessity" should never have appeared in this context, i.e. it should not have been introduced in such a way that it could be interpreted in terms of *only* legibility conditions and optimal computation. One hint that this might be the case is that the passage quoted

in p. 54, which appears originally in Chomsky (1995b [1994]), reappears with a slight, but very suggestive, modification in Chomsky (1995a: 1), where the first sentence reads: "[(i)] What are the general conditions that the human language faculty should be expected to satisfy?"[7] As the reader can easily verify, the phrase "virtual conceptual necessity" disappears altogether, and when Chomsky employs it again in his writings, he does so in a completely different fashion – a fashion that signals in fact a different approach to the SMT, as we will now see.

3.4 The imperfection strategy

In the previous section, we were led to view the SMT as suggesting that nothing is special to language, and we have discussed the difficulties that this view creates. In this section, we focus on what might be regarded as a different approach to the SMT, namely what I will call here the "imperfection strategy." Under this approach, the SMT seems to suggest, not that nothing is special to language, but rather that nothing is "imperfect" in language. Let us consider how this approach differs from the previous one.

In his effort to substantiate the thesis that language exhibits "perfect design," Chomsky has adopted a research strategy which (i) specifies what he takes to be the core function of language, and (ii) asks how perfect language is at performing this function. To put it differently, one starts with the SMT as defined in the previous section, and then proceeds to try to discover where this thesis fails. Thus Chomsky (2000b: 97–8) says:

> 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 ... If all such efforts fail, then add "imperfections" as required.

Two clarificatory comments on this passage are in order. First, a property P of language constitutes a departure from "perfect design" (and, therefore, a departure from the SMT) whenever it can be established that P is neither motivated by the need to satisfy some legibility condition or other nor follows from the efficiency with which the computational system is supposed to operate. That this characterization is plausible is suggested by how Chomsky understands compliance with the SMT. For instance, taking the SMT as his point of departure, he assumes that the faculty of language "provides no machinery beyond what is needed to satisfy minimal requirements of legibility and that it functions in as simple a way as possible" (Chomsky 2000b: 112–13). We may,

therefore, understand the SMT as suggesting that *nothing is imperfect in language*. This amounts to an empirical claim which, as we shall see later, does not necessarily imply that *nothing is special to language* (in the sense of the previous section).

Second, the passage suggests that whenever the SMT fails, we must "add 'imperfections' as required." Setting aside, for the moment, what the phrase "add imperfections" might mean, it should be noted that Chomsky refers here to real, as opposed to apparent, imperfections. Thus, just because a property P *seems* to challenge the SMT does not make it *ipso facto* an instance of an imperfection; one must first ascertain whether P constitutes a "real" property, an issue that is only going to be resolved in the context of serious linguistic analysis. "The research strategy," says Chomsky, "is to seek 'imperfections' of language, properties that language should not have," and he continues to suggest that, given some property P of language, there may be three possible outcomes of such a strategy:

(i) P is real, and an imperfection
(ii) P is not real, contrary to what has been supposed
(iii) P is real, but not an imperfection; it is part of a best way to meet design specifications (Chomsky 2000b: 112).

As one would expect, of these three outcomes, only the first can be regarded as a genuine falsification of the SMT. The second outcome, when achievable, is responsible for the elimination of much of the machinery of pre-minimalist models of language. We have already seen examples of this "excess baggage," as Chomsky (2000a: 11) calls it, in the previous chapter (Section 2.5), including the elimination of S-structure and D-structure. As to the third outcome, it is perhaps the most interesting from a minimalist perspective, for it substantiates the SMT in areas where, at first sight, it appears to be under threat.

So far nothing in this "imperfection strategy" suggests that it differs in any significant way from that discussed in the previous section, especially with respect to how the SMT is understood. Specifically, this strategy appears to assert extensional equivalence between those properties that are indicative of "perfection" and those that are conceptually necessary. A difference begins to emerge, however, when we observe that Chomsky employs another criterion for recognizing candidates as instances of "imperfection" in language design. This criterion relies on the intuition that if a property P is present in natural languages and is absent from formal languages, then P may constitute a deviation from "perfect design," i.e. an imperfection. As we shall see in the next section, this

criterion opens up a third route to conceptual necessity, but arguably conceptual necessity of a subtly different kind to that discussed in Section 3.2. For now, let us explore this new criterion.

Chomsky (2002: 109) says:

> A good guiding intuition about imperfection is to compare natural languages with invented "languages," invented symbolic systems. When you see differences, you have a suspicion that you are looking at something that is a *prima facie* imperfection.

We thus have two criteria for suspecting that a property P constitutes an imperfection in language: (i) *prima facie* indications that P cannot be "grounded" in either legibility conditions or optimal computation and (ii) indications that P does not occur in artificial systems. To see these two criteria at work, we need only look at what Chomsky (2002) says about the status of morphology. There, he considers morphology "as a striking imperfection" by virtue of its absence in formal languages. However, he also notes some exceptions to this generalization, suggesting that "plurality on nouns is not really an imperfection. You want to distinguish singular from plural, the outside systems want to know about that" (Chomsky 2002: 111). Clearly, the generalization and its exception here rely on different criteria; the former is based on a comparison between natural and formal languages, and the latter is suggested by one aspect of the SMT, namely interface requirements.

At this point, it is instructive to see how this new criterion, which is based on a comparison between natural and formal languages, fits in with Chomsky's research strategy as described above. Employing the new criterion, if property P shows up (or does not show up) in formal languages, then P may constitute a compliance with (or a departure from) "perfect design." Call this first-order perfection (or, correspondingly, first-order imperfection). The next step will be to attempt to show that a first-order imperfection is only apparent and that P is in fact an instance of the third of the three outcomes cited above (p. 62); that is, P is motivated by the external systems or follows from the efficiency with which the computational system operates. Depending on whether such an attempt succeeds or fails, the validity of the SMT will be assessed accordingly. Call this second-order perfection (or, correspondingly, second-order imperfection). This seems to be at least a component of the strategy that Chomsky has pursued in his attempt to account for the apparent complexity of language. To take a concrete example, consider the way he approaches two *prima facie* imperfections of language, namely uninterpretable morphosyntactic features and dislocation:

> 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. (Chomsky 2000b: 120)

Chomsky (2000b: 121) then goes on to seek to bolster this suspicion by speculating that the presence of the dislocation property may be motivated by the interface requirement to express scope and discourse properties, and that the presence of uninterpretable features may provide an effective way of implementing the dislocation property. Thus, setting many details aside, two *prima facie* imperfections turn out to be merely apparent, i.e. examples of the third outcome.

Two questions arise at this stage: (i) why should the absence of properties in formal language systems lead to them being regarded as *prima facie* imperfections in natural languages?; (ii) why would such an absence lead to the inference that interface conditions may be responsible for the presence of these properties in human language? Seeking answers to these questions, we observe that inflection and dislocation were deemed by Chomsky to be "special properties of human language, among the many that are ignored when symbolic systems are designed for other purposes," purposes "which may disregard the legibility conditions imposed on human language by the architecture of the mind/brain" (Chomsky 2000a: 12).

In the light of this, a plausible answer to (ii) may be that, whenever natural and formal languages differ from each other in terms of their properties, we have reason to suspect that the difference in properties is due to the different *functions* which these two types of language are designed to perform. If this is true, then it is reasonable to infer from the absence of certain properties in formal systems that legibility conditions are responsible for why these properties are present in human language, since satisfying these conditions constitutes a central function that human language, and only human language, is assumed to perform.

Turning now to (i), notice, first, that this question arises from what Chomsky says in a passage we quoted earlier (p. 63), namely that when we encounter differences between natural and formal languages, we "have a suspicion that [we] are looking at something that is a *prima facie* imperfection." Notice further that the notion of "imperfection" as invoked here cannot be defined or understood in terms of violation of legibility conditions, for we have just agreed that these conditions are not relevant to formal or symbolic systems. We are thus led to ask about the grounds on which this notion is based. To see these, consider what Chomsky says about the ontology of two computational operations, Merge and Agree:

> One [operation] is indispensable in some form for any language-like system: the operation *Merge* ... A second is an operation we can call *Agree* ... 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. (Chomsky 2000b: 101)

Taken together, this and the previous passage quoted (p. 64) make it plausible to infer that Chomsky conceives of the contrast between perfection and imperfection in this context in terms of the contrast between indispensability and dispensability. Thus Merge is considered to be an aspect of the "perfection" displayed by natural language, not because its existence is motivated by legibility conditions, but because its existence is founded in the very notion of a language as a combinatorial system; it is a *necessary* property of any conceivable language (cf., however, Postal 2003). By contrast, the operation Agree cannot be said to enjoy the same status; rather, Agree is merely a consequence of the need to satisfy legibility conditions in an optimal way.

If this is correct, then the answer to our first question above is straightforward; the reason why the absence of a given property in symbolic or language-like systems leads to the suspicion of its being an imperfection has to do with the fact that such a property is *dispensable* in these systems, in the sense that it is not necessary for them to perform their function, let alone to be identified as "language-like" systems.

3.5 Conceptual necessity: a second encounter

We have just observed that, from a minimalist perspective, Merge is a necessary property of any conceivable language, but what kind of *necessity* is this? We recall from Section 3.2 that some minimalists define "conceptual necessity" in terms of the concept of *language as such*. To see what this means, consider the following passage from Chomsky (1980a), in which we find an emphasis on the distinction between two usages of "universal grammar." Having offered a familiar characterization in terms of human biological properties, he goes on:

> It is important to distinguish this usage from a different one, which takes "universal grammar" to be a characterization not of human language but of "language as such." In this sense, universal grammar attempts to capture *those properties of language that are logically or conceptually necessary, properties such that if a system failed to have*

> *them we would simply not call it a language: perhaps the properties of having sentences and words, for example*. The study of biologically necessary properties of language is a part of natural science: its concern is to determine one aspect of human genetics, namely, the nature of the language faculty. Perhaps the effort is misguided … The criteria of success or failure are those of the sciences. In contrast, the study of logically necessary properties of language is an inquiry into the concept of "language." I should add at once that I am skeptical about the enterprise. It seems to me unlikely to prove more interesting than an inquiry into the concept of "vision" or "locomotion." But in any event, it is not an empirical investigation, except insofar as lexicography is an empirical investigation, and must be judged by quite different standards. (Chomsky 1980a: 28–9, my italics)

I quote this passage at length, not just because it illustrates the concept of *language as such* with which we are concerned here, but also for another reason that will become apparent in the next section. Now, setting aside the "logical necessity" to which the passage refers, it should be clear that the notion of "universal grammar" as understood here in terms of the conceptually necessary properties of "language as such" does not seem to differ from the notion of "conceptual necessity" as understood by some minimalists. This is clear from the fact that Chomsky refers here to "sentences and words" as illustrative of conceptual necessity in the same sense in which Hornstein *et al*. (2005: 6) cite the relationship between "sentences" and "words and phrases" as an example of conceptual necessity.

We can now see what kind of necessity is involved in the operation Merge. To put it in terms of the italicized portion of the passage just cited from Chomsky, Merge is a *conceptual necessity* in the sense that if a system failed to have it we would simply not call it a language.[8] This is consistent with how Merge was distinguished from Agree above; namely, Merge was considered to be an aspect of the "perfection" displayed by natural language, in the sense that it is an *indispensable* property of any conceivable language. We see, then, how the imperfection strategy indicates that there are two ways in which a property can manifest "perfection," by being a property of any conceivable language or, failing that, by satisfying interface conditions in an optimal way. Now, we can simultaneously achieve consistency and clarity if we make a similar distinction in terms of "conceptual necessity." This suggestion can be readily realized if we (i) adopt Atkinson's two-route interpretation of "conceptual necessity" (see Section 3.2), and (ii) recognize that Merge's conceptual necessity extends

beyond natural language to "language as such." It is *this* latter kind of necessity that Chomsky fails to make explicit in the passage I quoted in Section 3.2 (p. 54). There he encourages thinking of conceptual necessity *only* in terms of legibility conditions and optimal computation, but later (cf. the imperfection strategy discussed above) he acknowledges a third way of "grounding" conceptual necessity, *viz*. *via* the very idea of a "language." To see how the two kinds of conceptual necessity differ from each other, let us briefly reconsider the distinction between Merge and Agree.

In the case of Merge, if the presence of this operation can be justified in terms of the concept of language as such, then Merge inherits its conceptual necessity in terms of this justification; that is, Merge is conceptually necessary by virtue of conceptualizing any language as a combinatorial system. Given this necessity, it makes no sense to ask *whether* language has Merge; if Merge is present by virtue of a conceptual necessity in the sense just specified, then we can question its presence only by questioning the presence of language itself. Turning to the operation Agree, the situation is different in that if its presence can be justified in terms of either legibility conditions or optimal computation, then Agree inherits its conceptual necessity in terms of this *empirical* justification. Since legibility conditions and optimal computation are both a matter of empirical inquiry, it follows that the necessity involved here can only be justified by empirical evidence regarding its presence in language. It is precisely for this reason that, unlike in the case of Merge, the *empirical* question of whether Agree is present in language makes sense.

Now, while Merge is widely conceived of as a conceptual necessity in the minimalist literature, almost no one would regard Agree as a having the same status. But there seems to be no reason to deprive the latter operation from this status so long as we distinguish between the types of conceptual necessity involved in each operation. Note, further, that this distinction is desirable if we are to reconcile what appear to be mutually contradictory interpretations of conceptual necessity. For instance, and as mentioned in Section 3.2, Smith (2000: xii) asserts that any deviations from conceptual necessity are motivated by legibility conditions. This is clearly inconsistent with how Atkinson interprets this notion. As observed, according to Atkinson (2005a), satisfying legibility conditions is regarded as conceptually necessary. But if, as I argue here, we recognize the distinction referred to above between the two species of conceptual necessity, the tension between the two interpretations can be resolved.

I suspect that both Atkinson and Smith recognize each other's interpretation of conceptual necessity, at least implicitly. For instance, Atkinson (2005a: 22, n. 37) defends the conceptual necessity of Merge against Postal's (2003) assault

by arguing that the conceptual necessity of this operation is justified so long as the language system is one which we assume to be derivational, and in which the lexicon and the computation system are separated from each other. He emphasizes that this necessity should *not* be understood as a *logical* necessity. If this is true, then he seems to recognize a *third* way of conceiving conceptual necessity, one which differs from the necessity associated with legibility conditions and optimal computation.[9] Now, to the extent that the notion of "language as such" is understood in Atkinson's terms – that is, in terms which do not ascribe logical necessity to Merge – I think that the way he conceives of the conceptual necessity of Merge matches the way Smith (2004: 84) thinks about this operation, namely as one which it "is impossible to do without."

Smith, in turn, also seems to recognize the kind of conceptual necessity associated with legibility conditions. For instance, on the same page in which he asserts that any deviations from conceptual necessity are motivated by legibility conditions, he comments on Chomsky's argument regarding the two apparent imperfections of uninterpretable morphosyntactic features and dislocation (referred to earlier in this section) by saying: "In fact, if the argument is correct, the imperfections are, indeed, only 'apparent.'" The reason for this, as he explains, is that "[g]iven the constraints that other systems of the mind/brain impose on solutions to linking sound and meaning, there may be no other alternatives, so conceptual necessity explains the form of the grammar overall" (Smith 2000: xii). Clearly, this is a conceptual necessity that relates to legibility conditions. More importantly, the suggestion that "there may be no other alternatives" indicates that Smith conceives of conceptual necessity in this case as one which rests its justification on empirical, *a posteriori* grounds. This seems to me to be in line with Atkinson's two-route interpretation of conceptual necessity.

A note of qualification is in order before we proceed further. We have been making an implicit assumption throughout our discussion of conceptual necessity, namely that the necessity associated with legibility conditions concerns both of the interface levels. This need not be the case, however, especially if we take into account Chomsky's views on the evolution of language. Chomsky (2002, and in many other works) claims the primacy of the syntax-semantics interface over the syntax-phonology interface. This claim is linked to one of his assumptions about the evolution of language, namely that language did not evolve for communication but rather for the expression of thought (for discussion of this latter point, see Chapters 4 and 5). The point that is important for our present purposes, however, is that Chomsky seems to be trying to substantiate this claim by arguing that "imperfections" in the design of language arise from

the need to satisfy the external requirements coming from the phonological component. Thus, Chomsky (1995a: 265) suggests that language "imperfections" are due to "the external requirement that the computational principles must adapt to the sensorimotor apparatus, which is in a certain sense 'extraneous' to the core systems of language as revealed in the N → λ [i.e., the mapping from the syntax to the semantics – FA]."[10] If this is the case, then the conceptual necessity that is linked to legibility conditions should be viewed as concerning only the conceptual-intentional system.

Before proceeding to point out what seem to me to be shortcomings in Chomsky's imperfection strategy, let us recapitulate briefly what has been said so far. We began by acknowledging two distinct approaches to the SMT in Chomsky's work. Having dealt with the first approach in Sections 3.2 and 3.3, our effort in this and the previous section was concerned with uncovering the major aspects of the second approach (i.e. what we have called the "imperfection strategy"). We noted that the notion of "imperfection" can be understood in two different ways, each of which represents one aspect of departure from "perfect design." At one level, imperfection arises where a property is *not* a *sine qua non* of the concept of language as such; at another, it arises where a property is *not* a *sine qua non* of the concept of language as a cognitive, natural system. This led us to suggest that Chomsky's imperfection strategy involves a third route to conceptual necessity, one which was not explicit, although almost certainly presupposed in his earlier approach to the SMT. We argued that the two notions of "conceptual necessity" and "perfection" are equivalent and that the distinction between the two levels of perfection should also be carried over to the notion of conceptual necessity. We also noted a tension between two conflicting interpretations of conceptual necessity, and we suggested that recognition of two distinct kinds of this necessity might resolve this tension. Lastly, we suggested that, given Chomsky's views on the evolution of language, the conceptual necessity associated with interface conditions refers only to the semantic interface.

3.6 Some shortcomings

In evaluating Chomsky's imperfection strategy, it is important not to lose sight of the problem to which this strategy is a response, namely the apparent complexity of UG. As mentioned at the end of the previous chapter, minimalism (or, more specifically, the SMT) is ultimately concerned with providing an explanation for how it is possible for an apparently complex language system to arise in humans in a relatively short period of time. With this in mind, we now turn to outline some shortcomings in the imperfection strategy.

To begin with, there seems to be a lack of consistency in Chomsky's position regarding his appeal to the analogy between natural and formal systems. For instance, and as the passage quoted at length in p. 65 makes clear, Chomsky was particularly keen to draw a sharp distinction between the *empirical* study of language as a biological system and the *non-empirical* study of language in a wider sense that encompasses formal systems; he was strongly positive about the former, and explicitly skeptical about the latter. It is therefore something of a surprise to read him saying (as quoted in p. 63) that one "good guiding intuition about imperfection is to compare natural languages with invented ... symbolic systems." Here, despite the fact that the notion of "imperfection" has formal (i.e. non-empirical) content, it is nevertheless called on to provide guidance on an empirical issue, namely the nature of the language faculty. This becomes particularly evident in the use to which "virtual conceptual necessity" is put in the imperfection strategy; we have here an *a priori* notion that dictates (partially, at least) what properties human language should have – a matter that should be resolved *a posteriori*.

One may object to this by saying that the appeal to formal systems in assessing the nature of human language (*qua* biological system) should not in itself be objectionable, any more than should the appeal to mathematical models in assessing the nature of physical objects. But this objection puts the matter in a false light. What is at issue here is not a matter of a logical entailment of properties whose existence awaits empirical confirmation. Rather, the issue is one of setting up an analogy on the basis of which one can determine *a priori* which properties of language are a matter of empirical discovery and which are not. When Chomsky (1995a: 378) says, for instance, that "Merge is inescapable in any languagelike system," or that it "is necessary on conceptual grounds alone" (Chomsky 1995a: 243), he is in effect saying that any empirical inquiry as to whether human language has something like Merge is not worth the effort it involves. This may well be the case. But if so, it follows that Merge cannot be taken as falling within what Chomsky refers to as the "biologically necessary properties of language." Yet this conclusion is clearly at odds with how Chomsky viewed – and still views – Merge, namely as a property unique to the genetic component of the language faculty.[11]

Perhaps here lies the reason why the imperfection strategy seems to me to leave much to be desired in terms of clarity. Chomsky (2001: 2) suggests that "[i]f empirical evidence requires mechanisms that are 'imperfections,' they call for some independent account: perhaps path-dependent evolutionary history, [etc.]." Now, it is not clear where this leaves the operation Merge.[12] As Atkinson (2009: 7) correctly observes, "Merge

comes out ... looking like both a 'perfection' ... and an 'imperfection.'" Indeed, it looks like a "perfection" because, as seen above, it is indispensable in any language-like system, and it looks like an "imperfection" because, as mentioned several times in this and the previous chapter, it is regarded by Chomsky as a genetically determined property of language and, therefore, has a "path-dependent evolutionary history." As a consequence, we may take the SMT as embodying the empirical claim that nothing is imperfect in language, but we are not allowed to infer from this that nothing is special to language, for the uncertain status of Merge restrains us from drawing such an inference.

More importantly – and this is a point that requires attention – Chomsky's approach to "apparent imperfections" makes the central claim of "perfect design" immune from falsification. This may not seem obvious at first, because the approach seems to be driven by attempts to falsify the SMT. However, this is not the case. Consider again Chomsky's research strategy as outlined in Section 3.4, which suggests three possible outcomes for some property P of language:

(i) P is real, and an imperfection
(ii) P is not real
(iii) P is real, but not an imperfection.

The first outcome entails the availability of two criteria as to when P is "real," and when it is an "imperfection." If our exposition of Chomsky's imperfection strategy is correct, these criteria are as follows. First, P is an imperfection if it both does not occur in artificial systems and cannot be "grounded" in either legibility conditions or optimal computation. Second, P is real if its existence is supported by empirical evidence in the context of convincing linguistic analysis. Stated in this form, the two criteria appear to be independent, in the sense that what is "real" need not be an "imperfection," and *vice versa*. Indeed, it is precisely because these two criteria are mutually independent that we can speak of the possibility of other outcomes (e.g. (iii)).

Now, consider the second outcome. In accordance with the second criterion above, one may be inclined to say that the reason why P is not real is because it lacks analytical and empirical support. But what kind of support might this be? We have already seen several examples where much of the "excess baggage" of language structure has been eliminated on the grounds that it neither has the sanction of "virtual conceptual necessity" nor "empirical necessity." Yet, we have also agreed that these two types of necessity together provide the foundation on which the notion of "imperfection" is defined. Therefore, we are forced

to conclude that P is not real *because* it is an imperfection. Since this conclusion is at variance with the independence of the two criteria – and since this independence seems necessary for the possibility of outcomes like (i) and (iii) – we therefore have reason to suspect the cogency of the set of outcomes above. To see that such suspicions are justified, we need only observe that the proposition "P is not real because it is an imperfection" is equivalent to: "If P is an imperfection, then P is not real." Stated as such, this conclusion is clearly inconsistent with (i). It also seems intuitively (though not logically) to run counter to (iii).

We can now see, I hope, that the tendency by which minimalists seek to justify the outcome in (ii) entails that the outcome in (i) will be most unlikely to arise. Grohmann (2006: 1) provides a good example of this tendency, for he maintains that "levels of representation that do not follow from either '(virtual) conceptual necessity' or 'bare output conditions' are rejected." A moment's reflection reveals that what Grohmann is actually saying is that there can never be a case in which levels of representation constitute a real imperfection.[13] Though he does not explain why this might be so, his reason is not hard to guess; if a level of representation is forced upon us by reasons that do not bear any relation to either conceptual necessity or the SMT, this level cannot be real and, therefore, must be "rejected." Another example of the same type comes from Hinzen (2006a: 166), who sees the minimalist approach as an attempt to demonstrate that imperfections in language are either "merely apparent" or "real": if real, they reflect an optimal way of meeting legibility conditions; if merely apparent, they are merely "an artefact of our description or theoretical perception." Clearly, Hinzen does not even contemplate the possibility that real imperfections may be found which cannot be explained in terms of the SMT (i.e. the first outcome as described above).

It should not, therefore, be surprising to find that, to the best of my knowledge, there are no examples in the minimalist literature that might be listed under outcome (i).[14] What this really means is that we are given no clear indication as to how the minimalist claims about "language design" might be falsified. Of course, we are told, in a passage quoted earlier, that if all efforts at satisfying the SMT fail, we must "add imperfections" (Chomsky 2000b: 98). But if this proposal is to be taken seriously, it has be to shown that it is possible to obtain "real imperfections" on *independent grounds*, that is, on grounds independent of the empirical necessities dictated by the SMT. Indeed, when Chomsky (2008a: 135) says that "[a]ny departure from SMT ... merits close examination, to see if it is really justified," he is in effect implying that real imperfections can be justified on grounds that lie beyond the scope of the SMT. But since there is no indication as to what these independent grounds might be,

we are left with the minimalist practice which has always been an immediate rejection of what might constitute a departure from the SMT.

One last point before we close this section. One would have thought that the notion of "virtual conceptual necessity," as understood by many minimalists, is no more difficult to grasp than the underlying concept of language as a combinatorial system, and, therefore, one would have expected that what is conceptually necessary in language is not – and should not be – a matter of discovery or dispute. Yet clearly this has not been the case, even in Chomsky's own work. As observed earlier, for many years Chomsky and his colleagues considered the operation Move as an (apparent) imperfection, on the grounds of its absence from formal language-like systems. Unlike Merge, the operation Move "is clearly not conceptually necessary," as Smith (2004: 85) has put it. Of course we now know that this is not the case and that Move is as virtually conceptually necessary as is Merge (cf. Chomsky 2005: 12). What are we to make of this perplexity in deciding on the scope of "virtual conceptual necessity," other than to see it as an indication that the imperfection strategy is seriously flawed in failing to make a clear distinction between conceptual necessity and empirical justification? Whatever the answer to this question may be, I hope that the above discussion may help in clarifying some of the confusion that has prevailed in the literature.

3.7 The three factors framework

The minimalist approach to the apparent complexity of language culminates in the "three factors framework" (Chomsky 2005, 2007b, 2008a, 2010). In this framework, the notion of "imperfection" receives little emphasis, and the phrase "virtual conceptual necessity" is barely mentioned. What's more, as we will see later in this section, the SMT receives a much more explicit formulation than has previously been the case. In what follows, I introduce the three factors framework and show how it differs from the earlier approaches that we have looked at. Here the discussion will focus on two closely related topics: the essence of the SMT and the content of UG.

In Chomsky (2005), we meet the suggestion that the growth of language in the individual is determined by the interaction of three factors: (a) genetic endowment; (b) experience; and (c) general principles not specific to the language faculty. The last of these factors falls into two subcategories: one concerns principles of data analysis, and the other refers to principles of structural constraints and efficient computation. For convenience, we will refer to these three factors as Factor I, Factor II, and Factor III, respectively.

As discussed in the previous chapter, the first two factors have been central to the problem of language acquisition for many years. As to the third factor, although we have argued that its significance to the problem of language acquisition and evolution has always been recognized by Chomsky (see Section 2.6), it did not emerge as a "factor" in its own right until the formulation of the three factors framework. Here, not only is it explicitly called a "factor," but it is also suggested to be the factor that should bear much of the burden of minimalist explanation. To put it in terms adopted from Chomsky (2004a), Factor III is the one responsible for carrying linguistic inquiry "beyond explanatory adequacy" to a deeper level of explanation. Before we proceed to discuss the significance of Factor III in accounting for the apparent complexity of language, let us say a few words about the three factors and their interaction.

Chomsky (2005) offers some examples of how the three factors might interact to determine the growth and development of language. One involves the concept of *canalization*. This term, coined by Waddington (1940), refers to "the robustness of developmental processes in the face of environmental and genetic variation" (Siegal and Bergman 2006: 235). The basic idea is that when canalization takes place, a phenotype assumes an "inertiatic state" with the consequence that phenotypic development is insensitive to certain genotypic or environmental differences.

However, while the concept of canalization might indicate an interaction among environmental, genetic, and physical factors in the development of a phenotype, it is not clear how this could be translated into a meaningful linguistic context in terms of the three factors framework. In fact, even if we assume, as Waddington (1942) does, that the outcome of canalization is an "optimal" phenotype, the vagueness here is not reduced by simply stating that "[a] core problem of the study of the faculty of language is to discover the mechanisms that limit outcomes to 'optimal types'" (Chomsky 2005: 5), as there appears to be no clear way by which one can relate the "optimal types" in the biological sense to optimal representations or derivations in the context of linguistic computations.

As another example, Chomsky refers to a language acquisition study carried out by Gambell and Yang (2003) in which the authors examine the extent to which human infants are able to use statistical learning (SL) to perform one of the fundamental tasks of language acquisition, namely the segmentation of words from fluent speech. Their results indicate that SL mechanisms fall short unless certain innate knowledge of phonological structure is assumed; for instance, the knowledge that a single primary stress accompanies each word. From this Chomsky (2005: 7) concludes that

> the early steps of compiling linguistic experience might be accounted for in terms of general principles of data analysis applied to representations preanalyzed in terms of principles specific to the language faculty, the kind of interaction one should expect among the three factors.

It should be noted that while this does indeed provide an appropriate example of the interaction among the three factors, it nevertheless appears to be at odds with some of the tenets of minimalism in at least two respects. In the first place, what it really shows is that Factor III is linked to issues of explanatory adequacy in such a way that its explanatory power becomes dependent on certain assumptions relating to Factor I; infants can use SL but they first need to know "where to look," knowledge which the authors assume to be genetically innate. This does not seem to be congruent with the advertised explanatory role of Factor III as described above. Second, and as we shall see below, Chomsky wishes to confine Factor I to the recursive operation Merge. Yet, the above example appears to extend the content of Factor I to include the innate knowledge of "where to look."[15]

At any rate, it is perhaps in light of language evolution, rather than language acquisition, that the significance of Factor III is to be best appreciated. For here we find an explicit proposal as to how the apparent complexity of language may be accounted for. For instance, we are told that a "principled explanation" of the language faculty and its properties may be achieved by "shifting the burden of explanation from the first factor ... to the third factor" (Chomsky 2005: 9). To the extent that this is feasible, the problem of how language, despite its apparent complexity, could have evolved in a relatively short period of time would be eased, since "the less attributed to genetic information (in our case, the topic of UG) for determining the development of an organism, the more feasible the study of its evolution" (Chomsky 2007b: 2–3).

Now, supposing that Factor III is simply the label under which the two aspects of the SMT fall, namely the satisfaction of legibility conditions and optimal computation, the notion of "principled explanation" expresses the view that any account of the language faculty is "principled" insofar as it can be derived from the SMT. This is what Chomsky (2007b: 3) seems to have in mind when he says:

> To the extent that third factor conditions function, the language will be efficiently designed to satisfy conditions imposed at the interface ... We can regard an account of some linguistic phenomena as *principled* insofar as it derives them by efficient computation satisfying interface

> conditions. We can therefore formulate SMT as the thesis that all phenomena of language have a principled account in this sense, that language is a perfect solution to interface conditions, the conditions it must at least partially satisfy if it is to be usable at all.

However, the suggestion "that all phenomena of language have a principled account" in terms of the SMT is misleading, for it may be taken to mean that *all properties of language are determined by Factor III*. But recall that, in Section 3.3, we have argued that such an interpretation cannot be true for a number of reasons, most notably because it leaves no room for the fact that *something* must be special to language. Indeed, Chomsky (p.c.) says that "Factor I must be non-empty, or it would be a miracle that my granddaughter acquired English though her pet kitten ... could not even get as far as identifying part of the environment as language related."[16] Notice in passing that the implicit assumption here – an assumption that reveals yet another way in which the three factors framework could be misleading – is that Factor I, though it refers to the genetic component of the language faculty, is in fact quite exclusive, in the sense of comprising *only* those properties that are *genetically unique* to humans and, one might add, to language. In other words, mere genetic determination of a language property does not guarantee membership of Factor I, for it may be the case that such a property is shared with other non-human species. Without this assumption, there is no reason why, as Chomsky asserts, Factor I should be non-empty. This is a point we will want to keep in mind when we later compare Chomsky's three factors framework with his contributions to Hauser *et al*. (2002) and Fitch *et al*. (2005).

Returning now to the passage quoted above (p. 75), we observe that, just a few lines earlier, Chomsky speaks of the SMT in such a way that does not preclude Factor I being non-empty. He says:

> [We seek] to close the gap between SMT and the true nature of FL. UG is what remains when the gap has been reduced to the minimum, when all third factor effects have been identified. UG consists of the mechanisms specific to FL, arising somehow in the course of evolution of language. (Chomsky 2007b: 3)

Two observations about this passage are in order. First, compared with the previous one, this passage is more accurate in that the generalization is expressed by "all third factor effects," rather than by "all phenomena of language." It is thus possible to entertain the suggestion – in my view somewhat implicit in the above passage – that the correctness of the SMT does *not* imply

that *nothing* should be regarded as special to language. Second, this passage also seems to suggest that the validity of the SMT does *not* provide us with the true nature of the language faculty; in order to come to terms with this latter, we need to have a view on whatever fills the gap (i.e. the content of UG). Presumably, the strategy here is to (i) start with the assumption that Factor I is heavily populated with specific properties and special mechanisms, (ii) move gradually to account for more and more of these in terms of Factor III, and (iii) identify what remains as the content of UG. Interestingly, this is quite the opposite of the "imperfection strategy" as outlined in Section 3.4, where one starts with the assumption that the SMT is correct, and then attempts to falsify it by seeking "imperfections" of language, adding thus to the content of UG.

At this point the question arises as to what fills the gap; that is, what is the content of UG? The answer that many minimalists would give is that UG consists only of the recursive operation Merge. On the face of it, this seems to be in line with Hauser *et al.*'s (2002) hypothesis that *recursion* is the only aspect of language that is uniquely human and uniquely linguistic.[17] But let us suspend judgement on this issue until we have considered the relation between Chomsky's linguistic and interdisciplinary discourses (Section 3.8). Here we end this exposition of the three factors framework with a brief look at how Chomsky advances a formulation of the SMT in which he makes explicit the suggestion that what is special to language is confined to Merge. He says:

> A very strong thesis, sometimes called "the strong minimalist thesis" SMT, is that language keeps to the simplest recursive operation, Merge, and is perfectly designed to satisfy interface conditions. (Chomsky 2010: 52)

When compared with the previous formulations of the SMT in earlier approaches (as described in Sections 3.3 and 3.4), the formulation contained in this passage stands out not only for its explicit reference to the operation Merge, but also for its specification of the ontological status of such an operation. For to say that language "keeps to" Merge is to say that this computational operation is the only aspect of language that lies within Factor I. To put it in terms of the passage cited in p. 76, Merge is the only mechanism that is "specific to FL, arising somehow in the course of evolution of language."[18]

Now, the passage we have just quoted continues by reducing the SMT to the following equation:

Interfaces + Merge = Language

which can be made more explicit by reading the reference in the passage above to language as "perfectly designed" as embracing the notion of "optimal computation." We might therefore expand the above equation to yield:

Interfaces + Optimal Computation + Merge = Language

In other words, language is the result of Merge operating under the conditions of interface legibility and computational efficiency. This seems be essentially what the SMT amounts to.

3.8 Two discourses, one thesis?

While it is not clear what was responsible for the emergence of the three factors framework, there is reason to suspect that Chomsky's collaborations with the biologists Marc Hauser and Tecumseh Fitch (see Hauser *et al*. 2002; Fitch *et al*. 2005) have played at least some role in this development. It may well be that the "recursion-only hypothesis" underlies Chomsky's later tendency to be more explicit about the tenets of minimalism, including the ontological status of Merge.

Of course, as some critics have noted, the hypothesis itself may be influenced by the minimalist program (MP). For instance, Pinker and Jackendoff (2005: 20) maintain that the "claim that the only aspect of language that is special is recursion lies in a presumption that the MP is ultimately going to be vindicated." Kinsella (2009: 129) goes even so far as to assert that "[t]he minimalist standpoint meshes with the claims of Hauser *et al*. quite obviously."

Now, there is no reason to leap to the other extreme and deny any connection between Chomsky's linguistic and interdisciplinary discourses, for, as we will see shortly, there are in fact similarities between the two that are too obvious to be denied. However, there are also uncertainties as to just how each discourse informs the other. For instance, when Hauser *et al*. (2002) refer to "recursion," do they mean by this what Chomsky in his linguistic discourse would mean? When they speak of the distinction between the faculty of language in the narrow sense (FLN) and the faculty of language in the broad sense (FLB), how does this distinction translate into Chomsky's linguistic discourse? Where does UG stand with respect to this distinction? These and similar questions will have to be confronted before equating (or distinguishing) Chomsky's linguistic and interdisciplinary discourses.

Thus, our primary task here is the bringing together of these two discourses so as to determine whether they are two sides of the same coin, or whether they

differ fundamentally. As the discussion will show, the correct answer lies between these two extremes. But before we begin, it must be pointed out that systematic exposition of relevant aspects of the evolution papers (Hauser *et al*. 2002 and Fitch *et al*. 2005), together with related work, will be given in the next chapter, where the focus will be on the empirical question of language specificity. Here we are merely concerned with those aspects that may bear on the three factors framework in particular and on Chomsky's linguistic discourse in general.

Let us begin with the distinction between FLB and FLN as defined by Hauser *et al*. (2002). We are told that FLN consists of the computational system of human language and is unique to our species. It comprises part of FLB, the remainder of which is shared with other species. As well as FLN, FLB includes (at least) the sensory-motor system and the conceptual-intentional system. In the light of this distinction let us ask: what are the possible counterparts of FLN and FLB in Chomsky's linguistic work? To examine this, consider the following passage from Hauser *et al*. (2002: 1574):

> Recent work on FLN ... suggests the possibility that at least the narrow-syntactic component satisfies conditions of highly efficient computation to an extent previously unsuspected. Thus, FLN may approximate a kind of "optimal solution" to the problem of linking the sensory-motor and conceptual-intentional systems. In other words, the generative processes of the language system may provide a near-optimal solution that satisfies the interface conditions to FLB.

Not surprisingly, Chomsky (1995a) appears in the reference list of Hauser *et al*. as an example of this "recent work on FLN." But more importantly, and leaving terminology aside, this passage makes an implicit but clear reference to the SMT in stating that "FLN ... may provide a near-optimal solution that satisfies the interface conditions to FLB." Now, consider the continuation of the passage, which we have already cited in a different connection (see p. 43):

> Many of the details of language that are the traditional focus of linguistic study [e.g. subjacency, Wh- movement, the existence of garden-path sentences ...] may represent by-products of [solving the problem of linking sound and meaning], generated automatically by neural/computational constraints and the structure of FLB – components that lie outside of FLN.

Clearly, this echoes the notion of shifting the burden of explanation from Factor I (i.e. the topic of UG) to Factor III (i.e. the conditions of interface

legibility and efficient computation). Thus, one may be inclined to equate Factor I and Factor III with FLN and FLB, respectively. However, a closer inspection will reveal that the relationships under consideration are not as straightforward as it may first seem.

Consider first the possible connection between FLB and Factor III. As we shall see in the next chapter, Fitch *et al*. (2005: 203) suggest that even if some aspects of language escape the force of the SMT, these aspects "would not automatically be part of FLN." This is because although there exists the possibility that some language properties might not be deducible *via* reference to the satisfaction of interface conditions in an optimal way (i.e. Factor III), there would still be the chance of locating these properties in FLB by arguing for their existence in the general domains of either animal or human cognition. An immediate implication of this is that the contents of Factor III constitute only a subpart of FLB.

As to the possible association of FLN with Factor I (UG), the parallelism in this case seems to be stronger. Thus, according to Hauser *et al*. (2002), FLN (i) refers to what is unique to humans and specific to language, (ii) represents a novel capacity that has emerged recently in the course of human evolution, (iii) includes only recursion, and (iv) approximates an optimal solution to the problem of linking the conceptual-intentional system and the sensory-motor system. Now, UG receives a similar characterization in Chomsky's (2005, 2007b, 2008a) linguistic discussions. Thus, UG (i) refers to the distinguishing aspects of human language, (ii) constitutes a recent human development in the course of evolution, (iii) contains only Merge, and (iv) approximates an optimal solution to the problem of linking the conceptual-intentional system and the articulatory-perceptual system.

In fact, the parallelism between FLN and UG goes one step further. Fitch *et al*. (2005) speculate that the content of FLN might also turn out to be empty, in which case FLN might be confined to the specific configuration that determines how the language mechanisms are integrated together in one way rather than in another. They write:

> The contents of FLN are to be empirically determined, and could possibly be empty, if empirical findings showed that none of the mechanisms involved are uniquely human or unique to language, and that only the way they are integrated is specific to human language. (Fitch *et al*. 2005: 181)

Chomsky (2007b) provides a similar speculation with respect to the content of UG, suggesting that if empirical evidence indicates that Merge is not

language-specific but rather recruited from other cognitive systems, then "there still must be a genetic instruction to use Merge to form structured linguistic expressions satisfying the interface conditions" (Chomsky 2007b: 5). It is perhaps worthy of notice, in passing, that this speculation, regardless of its plausibility, is consistent with Chomsky's assertion that "Factor I must be non-empty" (see Section 3.7).[19]

On the basis of the above remarks, one might be inclined to conclude that FLN is, *mutatis mutandis*, identical to UG (i.e. Factor I). However, the qualifications that are required by the *mutatis mutandis* clause have empirical implications that are too important to ignore. This is particularly so when one asks whether what Chomsky means by "Merge" is what Hauser *et al*. recognize as "recursion." It appears that the latter is much more general and inclusive, assimilating a range of technology beyond Merge into the language-specific recursive device. If this is true, as I will argue in the next chapter, it follows that the claim that FLN contains only recursion will have empirical content different from that of the claim that UG contains only Merge. As the next chapter will illustrate, it is through failure to appreciate this point that the recursion-only hypothesis has created considerable confusion, not only among critics, but also among supporters.

4 *The SMT in an evolutionary context*

4.1 Introduction

The rest of this book will be devoted to a thorough evaluation of the plausibility of the strong minimalist thesis (SMT) as formulated at the end of Section 3.7. That formulation has three components, Merge, interface conditions, and optimal computation, and it is the first of these that concerns us in this chapter. Since the role of Merge has also received considerable attention in recent discussion of language evolution, this latter will provide the context for what follows and provide us with the broader aim of evaluating the SMT from an evolutionary perspective.

The reader will recall that Hauser *et al*. (2002) advance the hypothesis that recursion is the only aspect of the language faculty that is unique to language and to humans and that Chomsky (2010: 52) takes the SMT to include the proposition "that language keeps to the simplest recursive operation, Merge." For convenience, we will refer throughout the present chapter to these two proposals as the "recursion-only hypothesis" and the "Merge-only hypothesis," respectively. Regarding these two hypotheses, the question that immediately arises is that of how they relate to each other.

Some scholars have attempted to assess the recursion-only hypothesis by assimilating it to the minimalist framework (see, among others, Scheer 2004; Kinsella 2009; Samuels 2009; and Progovac 2010). Although they differ significantly in the conclusions they reach about the content of the "narrow language faculty" (FLN), what all these discussions have in common is that they take for granted the assumption that what Hauser *et al*. mean by recursion is just Merge in the minimalist vocabulary.[1] However, a closer inspection of Hauser *et al*. (2002) and Fitch *et al*. (2005) makes it necessary to treat this assumption with caution. By focusing on the notion of recursion as employed in these two articles, this chapter shows how the recursion-only hypothesis differs from the Merge-only hypothesis, and assesses the implications of this difference for the evaluation of both hypotheses. This will prepare the ground for our assessment of one aspect of the SMT, namely the Merge-only hypothesis.

The chapter is organized as follows. The next section focuses on the recursion-only hypothesis, the context in which it emerged, and its reception in the literature. Section 4.3 discusses two extreme and opposed positions on the content of FLN: those of Kinsella (2009) and of Samuels (2009, 2011). In Section 4.4, the discussion turns to the notion of recursion to clarify what Hauser *et al*. consider to be the content of FLN, and in Section 4.5 the conclusions of previous sections are used to evaluate the Merge-only hypothesis.

4.2 The recursion-only hypothesis

As observed on several occasions already, the hypothesis that recursion is the only uniquely linguistic and uniquely human property of the language faculty was first advanced in Hauser, Chomsky, and Fitch (2002). Before we go into some of the details of this contribution, some general background on the evolution of language will be helpful to provide a context for our exposition.

Language is a property exclusive to humans, in the sense that its use constitutes one of the most striking qualities that differentiates *Homo sapiens* from other species. If it is indeed the case that the reason for the existence of this human prerogative lies in the strikingly small genetic differences between us and our closest relatives, then it is only sensible to ask *when* and *how* language emerged in humans. There is a huge literature debating these questions (e.g. Christiansen and Kirby 2003; Cangelosi *et al*. 2006; Larson *et al*. 2010; Scott-Phillips *et al*. 2012), but given the fact that spoken language leaves no evidence in the fossil record, debate on the evolution of language continues to be highly speculative. A feature of this debate is that it has been largely shaped by general views on the nature of evolution.

Consider, for instance, the question of whether evolution occurs gradually or in saltations. Discussion of this particular question has a long history and it continues to be a controversial topic to this day. Prior to Darwin, almost all evolutionists were saltationists, and "[a]mong those who accepted evolution after 1859 were not a few who were far more impressed by the occurrence of sudden mutations than was Darwin" (Mayr 1982: 544). In his book *On the Origin of Species by Means of Natural Selection* (2003 [1859]: 194), Darwin emphasizes the old aphorism "natura non facit saltum" to suggest that the history of life proceeds gradually under the force of natural selection; "she [i.e. Nature]," he writes, "can never take a leap, but must advance by the shortest and slowest steps." Although he was aware of the fact that the fossil record did not support this gradualist view, he believed the explanation for this apparent refutation of his theory lay "in the extreme imperfection of the geological

record" (2003 [1859]: 280). This Darwinian view became the norm among many biologists in the middle decades of the twentieth century, but the debate between gradualists and saltationists erupted again with the publication of Eldredge and Gould (1972). Their arguments echoed those of old saltationists by stressing the fact that the fossil evidence does not support the Darwinian gradualist view on evolution. For instance, they pointed to the fact that a significant number of fossil types appear suddenly and remain unaltered thereafter, which they argued indicated that breaks in paleontological data are real and not merely a matter of imperfection in the geological record.

Now, the question of whether the evolution of language is gradual and piecemeal or sudden is a special case of this broader question and, for this reason, it should not be a surprise to discover that the long-running debate in evolutionary biology between gradualists and saltationists resonates among scientists across various language-related disciplines. A case in point is Pinker and Bloom (1990), the purpose of which was to challenge the saltational views of Gould and Chomsky (cf. Chomsky's saltationist views in Section 2.6). A brief exposition of the major claims made in this article will suffice for our purposes.

According to the authors, the origin of language can be successfully explained by the theory of natural selection. This claim is based on two premises which Pinker and Bloom believe are basic to generative grammar and evolutionary theory, respectively: one, that language is a complex biological structure and, two, that "[t]he only successful account of the origin of complex biological structure is the theory of natural selection" (1990: 707). Thus, their intention is to show that evolutionary theory and generative grammar are perfectly compatible. This is evident from the abstract of their paper, in which they assert that "there is every reason to believe that a specialization for grammar evolved by a conventional neo-Darwinian process," a view that is reinforced by the following: "Since we are impressed both by the synthetic theory of evolution and by the theory of generative grammar, we hope that we will not have to choose between the two" (1990: 708).

Since invoking natural selection requires the specification of a certain function, Pinker and Bloom claim "that language shows signs of design for the communication of propositional structures" (1990: 712). If this claim is granted, then communication of propositional structures must have been beneficial to the species. The authors believe that this is indeed the case, arguing that "communication of knowledge and internal states is useful to creatures who have a lot to say and who are on speaking terms" (1990: 714). They further contend that language is an *adaptation*, in the sense that its various mechanisms have evolved to serve the purpose of communication.

Pinker and Bloom's paper is a response to the view that natural selection could not have been solely responsible for the emergence of language, and that the latter did not evolve for the purposes of communication – a view championed by Gould and Chomsky. As the authors put it, "when two such important scholars as Chomsky and Gould repeatedly urge us to consider a startling contrary position, their arguments can hardly be ignored" (1990: 708), and they go on to complain that, perhaps because of the enormous influence of these two scholars on cognitive science, "adaptation and natural selection have become dirty words," and those who invoke them are "open to easy ridicule as a Dr. Pangloss telling Just-so stories" (1990: 710–11).[2]

In their closing remarks, seemingly aiming to set an agenda for future research at the time they were writing, Pinker and Bloom (1990: 727) remark that "there is a wealth of respectable new scientific information relevant to the evolution of language that has never been properly synthesized," and optimistically assert "that there are insights to be gained, if only the problems are properly posed." However, the appearance of Hauser, Chomsky, and Fitch (2002) seems to have reset the research agenda, at least for some scholars.

The primary aim of Hauser *et al*. (2002) is, it is claimed, to promote interdisciplinary cooperation among scientists working in language-related fields in an effort to have a better understanding of the faculty of language (FL). The authors complain that many of the bitter debates on language evolution have derived from a failure to distinguish between two distinct but related aspects of language: *communication* and *computation*. They maintain that inquiries into the nature of language as a communicative system should be distinguished from inquiries into the abstract set of computations underlying this system. To help overcome this confusion and as a consequence to render the debate on language evolution more profitable, Hauser *et al*. set the stage by making a distinction between the faculty of language in the narrow sense (FLN) and in the broad sense (FLB). As already observed in the previous two chapters (Sections 2.3 and 3.8), FLN includes only the computational system and the mappings to the interfaces, while FLB comprises FLN, relevant aspects of the sensory-motor system (SM), language-related parts of the conceptual-intentional system (CI), and possibly other systems as well. Given this terminological distinction, the authors proceed to identify three key theoretical issues concerning the debate on language evolution.

The first issue revolves around the distinction between what is uniquely human in language and what is shared with other species. Most researchers maintain that there is a qualitative difference between animal communication and human language, in the sense that the former "lack[s] the rich expressive

and open-ended power of" the latter (Hauser *et al*. 2002: 1570). The evolutionary problem, therefore, lies in explaining this apparent discontinuity between humans and other living forms. The second issue relates to the "gradual *versus* saltational" distinction. As Hauser *et al*. observe, the difference between this second issue and the previous one lies in the fact that "a qualitative discontinuity between extant species could have evolved gradually, involving no discontinuities during human evolution" (2002). The third theoretical issue revolves around the "continuity *versus* exaptation" distinction. The crucial question here is whether language has evolved in a continuous fashion from pre-existing communication systems, or whether some aspects of it "have been exapted away from their previous adaptive function" (2002). Given this outline of the main theoretical problems concerning language evolution, the authors proceed to identify three hypotheses on how language could have evolved, the last of which is their own.

The first hypothesis maintains that the entirety of FLB, including FLN, is fundamentally similar to animal communication systems, the differences being a matter of degree rather than of essence. According to this hypothesis, then, there is a sense in which language is not a human prerogative, for it contains nothing that cannot be found in animal communication. The second hypothesis holds that language, as a whole, is a complex and genetically determined system that is unique to humans. This hypothesis requires that the whole of FLB is an adaptation for language, and even though it might be possible that FLB shares some of its mechanisms with other non-human communicative systems, these mechanisms must have been exapted away from their original function to the extent that it would be legitimate to consider them as uniquely human. According to this hypothesis, natural selection must have been a major evolutionary force in determining many aspects of FLB, since it is only by means of natural selection that the complexity of FLB becomes possible (i.e. the so-called "argument from design"). In contrast to these two hypotheses, Hauser *et al*. put forth their own, which consists of the following three claims: (1) FLN contains only recursion and the mappings to the interfaces; (2) it is recently evolved for reasons other than communication; and (3) it is the only component of FLB that is unique to the language faculty and unique to our species. Let us have a closer look at this last hypothesis.

First of all, by recursion Hauser *et al*. (2002: 1571) mean *syntactic recursion*, which "takes a finite set of elements and yields a potentially infinite array of discrete expressions."[3] This boundless expressive power of human language is captured by what the authors term the property of *discrete infinity*, itself the result of the computational mechanisms of *recursion*. It is a well-known fact

about human language that sentences are formed from discrete units and have no upper bound in terms of their length (but see Pullum and Scholz 2010). Hauser *et al*. (2002: 1576) believe that this capacity for discrete infinity is uniquely human:

> It seems relatively clear, after nearly a century of intensive research on animal communication, that no species other than humans has a comparable capacity to recombine meaningful units into an unlimited variety of larger structures, each differing systematically in meaning.

Given this observation, the authors "hypothesize that most, if not all, of FLB is based on mechanisms shared with nonhuman animals," and, in contrast, they suggest that "only FLN is uniquely human," and "compris[ing] only the core computational mechanisms of recursion as they appear in narrow syntax and the mappings to the interfaces" (2002: 1573). From this they formulate the implication that if it is true that FLN is limited to recursion and the mappings to the interfaces, then this fact would have "the interesting effect of nullifying the argument from design, and thus rendering the status of FLN as an adaptation open to question" (2002).

Moreover, Hauser *et al*. (2002: 1574) consider it implausible that human language has evolved gradually from animal communication systems, since "minor modifications to [these systems] alone seem inadequate to generate the fundamental difference – discrete infinity – between language and all known forms of animal communication." Instead, they suggest that the computational mechanism responsible for "discrete infinity," i.e. syntactic recursion, might represent a novel and recent development in the evolution of *Homo sapiens* and be unique to it. In addition, they claim that much of the apparent complexity in language might have been derived from FLB, "especially those [components] underlying the sensory-motor (speech or sign) and conceptual-intentional interfaces, combined with sociocultural and communicative contingencies" (2002: 1573).

While Hauser *et al*. concede that FLB may well be an adaptation, they argue that FLN was not adapted for communication. Indeed, they speculate that FLN has evolved for reasons other than communication (e.g. number quantification, navigation, social relationships, etc.), and it is only when the underlying computations proved to be useful for communication that they were later modified due to constraints at the interfaces. Accordingly, the authors do not discard the possibility that recursion might have precursors in non-human domains other than communication and, thus, they encourage researchers to seek out evidence for the existence of recursive mechanisms outside the communicative domain:

> Comparative work has generally focused on animal communication or the capacity to acquire a human-created language. If, however, one entertains the hypothesis that recursion evolved to solve other computational problems such as navigation, number quantification, or social relationships, then it is possible that other animals have such abilities, but our research efforts have been targeted at an overly narrow search space. (Hauser *et al*. 2002: 1578)

It should be clear that there is an apparent tension here. On the one hand, Hauser *et al*. consider recursion to be unique to language and to humans, and, on the other, they acknowledge that recursion may have precursors in animal non-communicative systems. I think that this tension can be resolved if we realize that there are two hypotheses on offer. There is the recursion-only hypothesis to which we have referred, and there is what might be called a "default hypothesis," one which the authors seem to be compelled to resort to when the former hypothesis is in danger of failing. Assuming the recursion-only hypothesis, Hauser *et al*. (2002: 1573) maintain that empirical data suggest that "uniquely human capacities [such as recursion] have evolved recently in the approximately 6 million years since our divergence from a chimpanzee-like common ancestor." However, when they entertain the possibility that recursion might have precursors in animal non-communicative systems, they fall back on a "just-in-case" hypothesis, according to which the property of recursion might have been exapted to the service of language. This is illustrated by the continuation of the passage we have just cited:

> If we find evidence for recursion in animals but in a noncommunicative domain, then we are more likely to pinpoint the mechanisms underlying this ability and the selective pressures that led to it. This discovery, in turn, would open the door to another suite of puzzles: Why did humans, but no other animal, take the power of recursion to create an open-ended and limitless system of communication? Why does our system of recursion operate over a broader range of elements or inputs (e.g. numbers, words) than other animals? One possibility, consistent with current thinking in the cognitive sciences, is that recursion in animals represents a modular system designed for a particular function (e.g. navigation) and impenetrable with respect to other systems. During evolution, the modular and highly domain-specific system of recursion may have become penetrable and domain-general. This opened the way for humans, perhaps uniquely, to apply the power of recursion to other problems. This change from domain-specific to domain-general

> may have been guided by particular selective pressures, unique to our evolutionary past, or as a consequence (by-product) of other kinds of neural reorganization.

Clearly, here we see an acknowledgement that evidence for recursive mechanisms in animals would falsify the hypothesis that recursion is unique to humans. Nevertheless, as the above passage shows, Hauser *et al*. are unwilling to relinquish their belief that something must be unique to human language; if not recursion, then perhaps the *way* in which this mechanism operates in humans. It will be necessary to return to assess the nature and consequences of this concession in Section 4.5. Here we close this section with a brief review of the impact Hauser *et al*.'s article has had since its publication, especially in connection with the notion of recursion. Our review will be selective rather than exhaustive.

One question that seems to have troubled some scholars is: what exactly do Hauser *et al*. mean by the term *recursion* (we shall be asking the same question in later sections, although not from a purely terminological point of view)? Kinsella (2009) takes the trouble of listing most of the definitions of "recursion" that have been proposed in the fields of linguistics and computer science, believing that this is a necessary first step in evaluating the recursion-only hypothesis. Tomalin (2007) tracks the roots of "recursion" in mathematical logic and linguistic theory, and proceeds to evaluate the hypothesis of Hauser *et al*. against the background of five formal definitions of "recursion" drawn from the works of Peano, Gödel, Church, and Turing. However, Chomsky seems to throw cold water on such terminological efforts by saying:

> [T]here's a lot of talk about recursion and it's not a mystical notion; all it means is discrete infinity. If you've got discrete infinity, you've got recursion. There are many different ways of characterizing that step, but they are all some sort of recursive operation. Recursion means a lot more than that, but that's the minimum it means. There are different kinds of recursion – partial recursive, general recursive – but we don't need to worry about them. (Chomsky in Piattelli-Palmarini *et al*. 2009: 387)

Thus, Chomsky simply equates recursion with *discrete infinity*. Discrete infinity in language is a property of linguistic expressions generated by the potentially *infinite* application of recursive Merge to a finite set of *discrete* lexical items or syntactic objects. It will be helpful to keep this definition in mind when we

discuss Chomsky's speculations about the relationship between language and arithmetic in Section 4.5.

The recursion-only hypothesis seems to be undermined by the fact that recursion is so omnipresent that it cannot be unique to human language. There is a large body of literature attesting to the presence of recursive patterns in nature (e.g. Mandelbrot 1982; Penrose 1989) and in music (e.g. Hofstadter 1999 [1979]; Giblin 2008). Outside syntax recursion has been argued for in phonology (Scheer 2004; Schreuder and Gilbers 2004; van der Hulst 2010), and in general cognition (Jackendoff and Pinker 2005; Jackendoff 2008; Kinsella 2009). From the opposite perspective, some authors have followed Everett (2005) in arguing that the Pirahã language lacks recursion (e.g. Kinsella 2009; Evans and Levinson 2009; Sakel and Stapert 2010), and others have gone further and argued that Everett's claims should not be surprising since the property of discrete infinity itself is not a universal property of human language (Pullum and Scholz 2010).[4] Some authors, while not disputing the essentiality of recursion to human language, have not considered it to be the property that best describes its unique character. Rather than recursion, a number of proposals have been offered, including the whole of syntax (Bickerton 2003), the linguistic sign (Bouchard 2006), and parametric variation (Smith and Law 2007).

From the point of view of empirical experimental work, the recursion-only hypothesis has sparked a new wave of research designed to explore the differences between human and non-human primates in terms of their learning capacities. A well-known example is an experiment on cotton-top tamarind monkeys by Fitch and Hauser (2004). The upshot of this experiment was that, unlike normal humans, cotton-top tamarins are unable to master a *phrase structure grammar*. Although Hauser *et al*. (2002: 1578) cite this experiment as providing supporting evidence for the recursion-only hypothesis, some critics have pointed out methodological flaws in the study and accused the authors of overstating their results (see, for example, Perruchet and Rey 2005).

Hauser *et al*.'s hypothesis has also sparked a debate between Chomsky and his collaborators on the one hand and Jackendoff and Pinker on the other (Pinker and Jackendoff 2005; Fitch *et al*. 2005; Jackendoff and Pinker 2005). Pinker and Jackendoff adopt an adaptationist perspective on language which is in many ways similar to that we have already seen from Pinker and Bloom (1990). For the purposes of this chapter, I shall presuppose familiarity with the broad terms of this debate, and raise any specific points relevant to the issues I am discussing at the appropriate point. Here I shall restrict myself to commenting on the outcome of the debate.

This outcome can best be described as inconclusive and the jury is still out on many of the issues that have been raised in this section and which structure much of the debate in question, including the fundamental question as to which aspects of language are uniquely human and uniquely linguistic. All contributors to the debate agree on the empirical character of this question, but they differ as to how the empirical evidence should be interpreted. To mention just one example, the evidence for the existence of a descended larynx in non-human animals has received different interpretations by both sides. While Hauser *et al*. (2002: 1574) interpret this evidence as supportive of their claim that speech is not special, Pinker and Jackendoff (2005: 9) maintain that it is more plausible that a descended larynx in non-human animals was subsequently modified by natural selection to facilitate speech in humans. Thus the same piece of evidence has served to both exclude speech from, and include it in, FLN.

Even worse, in at least one case, the two parties to the debate seem to confuse what the relevant evidence is. This is the case where the two sides disagree on whether word learning is a property that is uniquely human and uniquely linguistic. Thus, when Hauser *et al*. (2002: 1576) refer to Bloom and Markson (1998) to support their claim that "human children may use domain-general mechanisms to acquire and recall words," Pinker and Jackendoff (2005: 12–13) respond by saying that the experiment by Bloom and Markson "did not conclude that words are acquired by a *domain-general* mechanism," but "showed only that children display similar levels of recognition memory for a newly learned word and a newly learned fact." Now, the experiment to which Pinker and Jackendoff refer was conducted by Markson and Bloom (1997) and showed, as the authors correctly observe, that word learning and fact learning in children share the same underlying cognitive abilities. But it should be noted that Hauser *et al*. do *not* refer to this experiment; rather, they refer to Bloom and Markson (1998), which is simply a review of various studies which "support the view that young children's remarkable ability to learn words emerges from more general cognitive capacities: intentional, conceptual, and syntactic," and that "some of [these capacities] are shared by other species" (Bloom and Markson 1998: 72). Interestingly (or rather confusingly), rather than drawing their opponents' attention to the relevant evidence, Fitch *et al*. (2005: 201) respond by saying that their opponents "are correct that we misrepresented the results of (Markson and Bloom 1997) in saying that children 'may use domain-general mechanisms for learning both words and facts.'" But the truth is that Hauser *et al*. could not possibly have misrepresented the experiment by Markson and Bloom (1997), for they did not even refer to it in their paper!

Finally, and more importantly, the charge of lack of falsifiability has been levelled by the debate participants against each other. On the one hand, Fitch *et al.* (2005: 193) complain that the "speech-is-special" hypothesis is not strong enough to be readily falsifiable. The reason they give for this is that since speech involves a host of varied mechanisms, evidence for the existence of a single speech-related mechanism in non-human animals would not undermine the hypothesis that speech is unique to humans. On the other hand, Pinker and Jackendoff (2005: 18) assert that "any theory can be rescued from falsification if one chooses to ignore enough inconvenient phenomena." Moreover, Jackendoff and Pinker (2005: 214) express no surprise that all the data they cite in support of a rich FLN have been assigned by their opponents to FLB; they believe that the reason for this is that the FLN/FLB distinction is applied by their opponents in the absolute sense, using *any* similarity between a language trait and anything else to justify excluding the trait from FLN.[5] What this charge implies is that the recursion-only hypothesis is not strong enough to be readily falsifiable, for the boundary line between FLN and FLB is not sharp enough to assess the plausibility of the hypothesis. For instance, Jackendoff and Pinker (2005: 216–17) point to an ambiguity in the formulation of the hypothesis, drawing attention to two possible readings for Hauser *et al.*'s statement that FLN includes "only the core computational mechanisms of recursion as they appear in narrow syntax and the mappings to the interfaces." This statement can be interpreted in two different ways, depending on where the brackets are placed:

(i) "mechanisms of recursion as they appear in [syntax and the mappings to the interfaces]";
(ii) "[mechanisms of recursion as they appear in syntax] and [the mappings to the interfaces]."

The authors argue that, under the first reading, the evidence they have provided suffices to falsify the recursion-only hypothesis. Under the second reading, however, they maintain that the hypothesis is rather uninteresting, because the nature of the "mappings to the interfaces" is not sufficiently specified by their opponents.

4.3 The content of FLN: two extreme views

Despite the fact that the debate briefly discussed in the previous section seems to be inconclusive, largely because of the difficulty in evaluating the claims of the two sets of protagonists, Kinsella (2009) and Samuels (2009, 2011) have found it possible to adopt two extreme, and opposing, positions. Kinsella, endorsing

the adaptationist position of Pinker and Jackendoff, argues that the minimalist conception of the language faculty and evolutionary theory "are incompatible" (Kinsella 2009: 186). She asserts "that FLN is complex and intricate, and that an evolutionary account which denies this is mistaken" (2009: 159). By contrast, Samuels (2011: 10) – who sides unreservedly with Hauser *et al*. (2002) – attempts to lend support to the "view ... that FLN is very small, perhaps consisting only of some type of recursion (i.e. Merge) ... and the mappings from narrow syntax to the interfaces." Specifically, she argues that phonology is neither uniquely linguistic nor uniquely human; in short, nothing in "phonology ... is part of FLN" (Samuels 2011: 36). We shall return shortly to discuss these opposing views, but for now, it is important to note that these two extreme positions share a common assumption, namely that Chomsky's linguistic discourse is identical in the relevant respects to his interdisciplinary discourse.

For instance, Kinsella (2009: 159) asserts that "it is obvious that the simplicity and atomicity which underpin the recursion-only hypothesis have been directly inspired by the development of minimalism within generative grammar." Elsewhere, as mentioned in the previous chapter (p. 78), she goes so far as to claim that "[t]he minimalist standpoint meshes with the claims of Hauser *et al*. quite obviously" (2009: 129). Similarly, Samuels (2009: 16, n. 2), despite acknowledging Fitch *et al*.'s claim that their views and those developed in the minimalist program are independent, believes that "they are two sides of the same coin."[6]

It should also be noted that Kinsella and Samuels are not alone in their adherence to this assumption. To mention just one example, Progovac (2010: 194) asserts that "the recursive power of language cannot be attributed to Merge alone, contra the hypothesis put forth in Chomsky (2005[a]); Hauser *et al*. (2002); and Fitch *et al*. (2005)." Clearly, Progovac takes it for granted that the Merge-only hypothesis is characteristic of both discourses.

Now, the reader will recall from the previous chapter that, although there are undeniable similarities between Chomsky's two discourses, there are also uncertainties as to how each discourse is supposed to inform the other. Given these uncertainties we have suggested that the idea of equating the two discourses should be treated with caution. It is the purpose of this and the next section to develop this line of thought, and to clarify its implications for any assessment of both the SMT and Hauser *et al*.'s hypothesis.

We have observed in the previous chapter (Section 3.8) that Hauser *et al*. (2002: 1574) make an implicit but clear reference to one formulation of the SMT in stating that "FLN ... may provide a near-optimal solution that satisfies the interface conditions to FLB." But how does the SMT relate to

the recursion-only hypothesis? The answer to this requires us to consider again the relationship between the two discourses. Suppose we start from the interdisciplinary perspective. In this case, an answer to our question is suggested by the following passage from Hauser *et al*.:

> Even novel capacities such as recursion are implemented in the same type of neural tissue as the rest of the brain and are thus constrained by biophysical, developmental, and computational factors shared with other vertebrates. Hypothesis 3 [i.e. the recursion-only hypothesis – FA] raises the possibility that structural details of FLN may result from such preexisting constraints, rather than from direct shaping by natural selection targeted specifically at communication. (Hauser *et al*. 2002: 1574)

What this passage seems to suggest is that the correctness of the recursion-only hypothesis is at least consistent with the SMT. Indeed, if the content of FLN is very small, a thesis (i.e. the SMT) that says that much of the apparent complexity of FLN is a by-product of non-language specific constraints becomes a reasonable possibility.

But now suppose that we start with the SMT itself. In this case, a different answer to our question arises. To see this, we recall from the previous chapter (Section 3.7) that we have identified the SMT with the equation *language* = *Merge* + *interfaces* + *optimal computation*. Now, it should be clear that the correctness of the SMT would entail that of the recursion-only hypothesis, for the latter is contained in the former.

Thus, although starting from either discourse, we can discern some sort of relationship between the SMT and the recursion-only hypothesis, neither starting point suggests that this relationship amounts to one of identity. Yet, so long as we regard Merge and recursion as "two sides of the same coin," there would appear to be a case for treating the linguistic and interdisciplinary discourses as relevantly identical and the fate of the SMT and the recursion-only hypothesis as being inextricably linked. We will now submit this identity assumption to critical scrutiny.

As noted in Chapter 2 (Section 2.3), the general notion of recursion is *instantiated* by the syntactic operation Merge. Since, with the advent of minimalism, the various recursive techniques of earlier frameworks (e.g. rewriting rules, X-bar theory, etc.) have been superseded by Merge, one might argue that this computational operation constitutes the *only* mechanism responsible for implementing recursion in language in the minimalist framework. Put like this, then, it makes some sense to say that recursion and Merge are "two sides of the same coin." Kinsella (2009: 129, n. 20) seems to have this in mind when she

writes: "Recursion is useless without Merge. In other words, although Merge is not recursion, Merge is necessary for recursion to be implemented in language." Notice that this is true only if we assume, as we have, that the only recursive mechanism available for language is Merge – an assumption which, as we shall see in the next section, does not necessarily reflect what Hauser *et al*. (2002), Chomsky *et al*. (2004: appendix), and Fitch *et al*. (2005) are proposing. Notice further that an implication of this is that Samuels's core claim that phonology lies outside FLN entails the absence of Merge inside phonology. Indeed, it is precisely this latter claim that Samuels (2009, 2011) and Samuels and Boeckx (2009) advocate.

Having clarified what the identity assumption amounts to, I will now demonstrate that, by adopting such an assumption, both Kinsella and Samuels have gone astray in their evaluation of Hauser *et al*.'s hypothesis. To see this, we must first be on guard against a confusion that is apparent in Hauser *et al*.'s own characterization of FLN. Consider, for instance, the following two statements in which the authors specify the content of FLN:

> FLN includes the core grammatical computations that we suggest are limited to recursion. (Hauser *et al*. 2002: 1570)

> We hypothesize that FLN only includes recursion and is the only uniquely human component of the faculty of language. (Hauser *et al*. 2002: 1569)

These two statements do not mean the same thing. Unlike the second, which clearly limits FLN to recursion, the first statement allows for the possibility that language mechanisms other than recursion may be part of FLN.[7] Now, it may be suggested that what the authors intend by the first statement is that FLN includes *only* the core computational operations, and that these operations are limited to recursion. If this were the case, it would follow that Hauser *et al*. restrict FLN to narrow syntax. But this cannot be what they have in mind, for they say:

> We assume, putting aside the precise mechanisms, that a key component of FLN is a computational system (narrow syntax) that generates internal representations and maps them into the sensory-motor interface by the phonological system, and into the conceptual-intentional interface by the (formal) semantic system. (Hauser *et al*. 2002: 1571)

With its reference to "a key component" and its explicit commitment to (at least part of) phonology and formal semantics being part of FLN, this passage clearly entails that FLN cannot be identified with narrow syntax. If this is true, it

follows that both Kinsella (2009) and Samuels (2009, 2011) are misguided in their understanding of what Hauser *et al*. consider to be the content of FLN, and there are further reasons for this conclusion.

Kinsella, to begin with, refers to various studies supporting the view that aspects of the lexicon, phonology, and morphology are specific to language. From this she concludes "that the proposal of Hauser *et al*. – that recursion is the sole defining property of language that makes up the faculty of language in the narrow sense – is flawed" (Kinsella 2009: 133). But this conclusion seems to be based on an erroneous interpretation of where Hauser *et al*. consider the boundary between FLN and FLB should be drawn. Indeed, Kinsella (2009) seems to be arguing against a straw man when she suggests that, since language properties other than recursion are uniquely human and uniquely linguistic, Hauser *et al*.'s dividing line between FLN and FLB "is not in the right place." For, from what we have just observed in the passage quoted above, these authors seem to be committed to the view that at least some aspects of phonology and semantics form part of FLN. In their second article, they confirm this commitment in saying:

> [W]e suggest that a significant piece of the linguistic machinery entails recursive operations, and that these recursive operations must interface with SM and CI (and thus include aspects of phonology, formal semantics and the lexicon insofar as they satisfy the uniqueness condition of FLN, as defined). (Fitch *et al*. 2005: 182)

Turning now to Samuels, her thesis that nothing in phonology is part of FLN is defended by reference to studies on animal cognition and behaviour, which she regards as "provid[ing] ample evidence that Pinker and Jackendoff's (2005) criticism of Hauser *et al*. (2002) concerning phonology is unfounded" (Samuels 2011: 58). On these grounds, she concludes that "phonology thus provides no challenge to the idea that FLN is very small" (Samuels 2011: 59). Thus, although Samuels reaches the opposite conclusion to Kinsella, she begins from the same erroneous premise, namely that Hauser *et al*. locate the whole of phonology outside FLN.

It is not hard to see why Kinsella thinks of her criticism as refuting the recursion-only hypothesis, and why Samuels thinks of her perspective as corroborating it. There appears to be an underlying argument to these two positions, one which neither Kinsella nor Samuels articulates, but which appears manifest when one reflects on their opposing views. We may express this underlying argument as follows:

Premise I:	Hauser *et al*. argue that FLN is limited to recursion.
Premise II:	Recursion and Merge are one and the same thing.
Conclusion:	Hauser *et al*. argue that FLN is limited to Merge.

Now, from a minimalist perspective, Merge is the narrow syntactic operation *par excellence*, and the computational system (narrow syntax) is said to be exhausted by this operation (see, for instance, Berwick (1998) and the passage cited from Chomsky (2010) in the next section, p. 98). Thus, it should not be surprising, given the argument sketched above, that both Kinsella and Samuels read Hauser *et al*. (2002) as claiming that (i) FLN can be identified with narrow syntax, and (ii) that the presence or absence of recursion in phonology can be considered as a valid criterion for the evaluation of the recursion-only hypothesis. We have already shown that the first claim does not seem to be supported by what the proposers of the hypothesis say in their two articles. As to the second claim, it is difficult to tell what Chomsky and his co-authors have in mind. But let us suspend judgement on this issue until we have given further consideration to the notion of recursion, a topic to which we now turn.

4.4 The where and how of recursion

If FLN cannot be identified with narrow syntax as Hauser *et al*. (2002: 1571) suggest, and if we are to follow these authors in their view that a defining feature of FLN is recursion, does it follow that those components of FLN that are not part of narrow syntax also exhibit recursion? According to Atkinson (p.c.), a positive answer to this question is consistent with his suggestion that recursion, as used in Hauser *et al*., must be understood as a property of the whole mapping – it is a *recursive mapping* between SM and CI with Merge at its core. From this he suggests that it follows that all the technology that goes into this mapping (Matching, Agree, Deletion, etc., alongside bits of phonology and (formal) semantics), becomes part of the recursive device.

Atkinson's suggestion involves two related claims: one about the *range* of recursion, and one about its *implementation*. As to the former, it suggests that recursion may not be limited to narrow syntax but may be present in the mappings to the interfaces *via* the phonological system and the semantic system.[8] This seems to be supported by what Hauser *et al*. (2002: 1573) say regarding their hypothesis:

> FLN comprises only the core computational mechanisms of recursion as they appear in narrow syntax and the mappings to the interfaces.

Thus, one might argue that this statement should be interpreted according to the first of the two readings as outlined at the end of Section 4.2, suggesting that in addition to syntax the parts of phonology and (formal) semantics that may be included in the content of FLN exhibit recursion. If this is true, the question of which parts of phonology and semantics exhibit recursion (and, therefore, should be included inside FLN) becomes essential for the empirical evaluation of the recursion-only hypothesis. For our purposes here, if Hauser *et al*. understand recursion as encompassing both the computational system and the mappings to the interfaces, the question arises of whether they consider Merge to be the *sole* mechanism responsible for implementing such a property. This question brings us to the second claim which Atkinson's suggestion involves, namely that Merge occupies a core place in the recursive device mapping between SM and CI. Note that this claim in itself does not rule out the possibility of, say, some phonological rules exhibiting recursion. Note further that, as far as the question here is concerned, in neither of their two articles do Chomsky, Fitch, and Hauser make any explicit mention of Merge. Indeed, as is clear from the passages we have thus far quoted, the authors themselves choose to be silent on this question.

It is of interest that Chomsky *et al*. (2004: 2) do refer to Merge in an unpublished discussion that was originally intended to be an appendix to Fitch *et al*. (2005). Here they write that "[t]he core computational mechanisms of recursion include the indispensable operation *Merge* and the principles it satisfies." Leaving aside this reference to the principles which Merge satisfies (and which the authors conjecture may be derived from general principles that are not specific to language – a conjecture that affords an additional basis for recognizing a close affinity between Chomsky's two discourses), it seems that we are facing here the same ambiguity as that identified in the first of the two statements cited in the previous section (p. 169), namely that "FLN includes the core grammatical computations that we suggest are limited to recursion" (Hauser *et al*. 2002: 1570). For to say that the core recursive mechanisms *include* Merge does not rule out the possibility that other mechanisms may also be responsible for recursion. By contrast, this ambiguity is not present in Chomsky's linguistic discourse. For instance, referring to a collection of essays published under the title *Interfaces + Recursion = Language?*, Chomsky (2010: 52) remarks:

> Proceeding beyond [Interfaces + Recursion = Language?], we therefore can inquire into the validity of SMT:
>
> (SMT) Interfaces + Merge = Language

> The associated question mark is even more forceful than before, since SMT reduces the options for recursion. In fact it bars almost everything that has been proposed in the course of work on generative grammar. Any stipulated device beyond Merge carries a burden of proof: the complication of UG must be based on empirical evidence.

What this seems to suggest is that a hypothesis that reduces the options for recursion is empirically stronger than one that leaves them open. The SMT, as formulated above, embraces this reduction, and, given the fact that Hauser *et al*. (2002) are silent on such a reduction, we may be justified in supposing that what Hauser *et al*. mean by recursion must be understood as being broader than the minimalist's Merge operation and cannot be regarded as equivalent to it. If this is correct, it confirms that the notion of recursion involves more than Kinsella (2009) and Samuels (2009, 2011) are disposed to admit. A further consequence is that the recursion-only hypothesis will have an empirical content different from that of a hypothesis which limits the uniquely linguistic and uniquely human component of the language faculty to the computational operation Merge. Moreover, the content of FLN will become more difficult to investigate empirically than the content of UG; this will be (at least theoretically) true, especially if one subscribes to Fitch's (2010: 23) suggestion that, given "the empirical difficulties of studying mechanisms unique to humans, biologists should be happy if most mechanisms involved in language do not fall into the FLN – the fewer its contents, the better." These consequences should be kept in mind for any serious assessment of the SMT and the Merge-only hypothesis which it involves (see Section 4.5).

However, we are not entirely out of the woods yet, as there are two matters that need to be addressed before we can proceed to the next section. The first relates to a question raised but left unanswered at the beginning of this section – do the components of FLN that are not part of narrow syntax also exhibit recursion? To be sure, we have observed that Fitch *et al*. (2005) are unequivocal in their commitment to (at least part of) phonology and formal semantics being part of FLN; we may, therefore, feel justified in suggesting that such commitment gives a positive, albeit partial, answer to this question. However, the problem with this suggestion is that it sits awkwardly with the authors' view on whether recursion can be identified in phonology. To see this, consider what they say in response to Pinker and Jackendoff in the debate discussed above:

> The discovery of a recursive mechanism in phonology would first raise the empirical questions "is it the same as or different from that

> in phrasal syntax?" and "is it a reflex of phrasal syntax perhaps modified by conditions imposed at the interface?" Second, given that the phrasal structure of music shows no obvious limit on embedding, we might ask "is phonological recursion the same as or different from that in musical phrases?" or in the phrases of birdsong. If the answer to all of these questions were "same," we would reject our hypothesis. (Fitch *et al*. 2005: 201)

Clearly, the authors believe that if phonology exhibits unambiguous evidence of recursion, this evidence should be considered as a falsification of their hypothesis. Thus, there appears to be an incompatibility between how they define the content of FLN and how they conceptualize the place of recursion within such content. More explicitly, their implicit suggestion that bits of phonology form part of FLN seems inconsistent with their proposal that phonology (probably) lacks recursion. Note that this inconsistency is only made possible by the fact that the authors define recursion as the sole defining property of FLN. In other words, given the assumption that if a property (or a component) of language forms part of FLN, then it must exhibit recursion, it is inevitable that inconsistency exists between the two suggestions just noted. In order to resolve this inconsistency, we need only turn the assumption the other way round. Thus, we assume that if a property (or component) of language exhibits recursion, then it must be part of FLN. Phrased in this way, this assumption allows us to locate some aspects of phonology inside FLN without requiring them to be recursive. It should be noted, however, that resolving the inconsistency in this way comes at the expense of rejecting Hauser *et al*.'s claim that recursion is the sole defining property of FLN. Indeed, if something can be an FLN property without itself being recursive, then recursion alone does not exhaust the content of FLN.

At this point, a further question arises: on what grounds should a non-recursive property of language be included inside FLN? There is an unsatisfactory answer to this question in a passage quoted in the previous section (p. 96), where Fitch *et al*. suggest that the recursive operations "include aspects of phonology, formal semantics and the lexicon insofar as they satisfy the uniqueness condition of FLN." Jackendoff and Pinker (2005: 217) object to this "insofar" clause on the grounds that it "turns this part of the hypothesis into a tautology," which says that "other than recursion, the uniquely human/uniquely linguistic subset of language consists of whatever aspects of phonology, semantics, and the lexicon prove to be uniquely human and uniquely

linguistic." This seems to me to be a valid objection, and I cannot see how Fitch *et al*. could meet it. For in addition to what Jackendoff and Pinker say, Fitch *et al*. (2005) strongly support the idea that when a certain trait is found in humans, the default assumption should be that that same trait is also found in other species unless empirical evidence points to the opposite conclusion. As they put it: "Human uniqueness is something to be demonstrated (as we do with recursion . . .), not assumed" (Fitch *et al*. 2005: 193). Clearly, the "insofar" clause above indicates that the authors fail to live up to their commitment, for it assumes without demonstration that other aspects of language may be part of FLN.

The final matter that we need to address concerns the relationship between FLN and UG. Although in the previous chapter we identified several parallelisms between these two constructs, we stopped short of drawing the conclusion that FLN was, *mutatis mutandis*, identical to UG. The reason we gave was that the qualifications that were required by the *mutatis mutandis* clause had empirical implications that were too important to ignore. One such qualification, already the subject of discussion in the previous section, concerns the suggestion that what Chomsky means by Merge differs from what Hauser *et al*. recognize as recursion. A second relates to a potential difference between FLN and UG in terms of their definition. While FLN is by definition the component of the faculty of language (FL) that is *genetically unique* to language and to humans, it is not clear how far this definition is applicable to UG.

Chomsky (2008a: 134) defines UG as the "theory of the genetic endowment" of the language faculty. What this definition seems to suggest is that UG concerns those properties of language that are *genetically determined*, and not necessarily *genetically unique*. This interpretation seems to be supported by what Chomsky says regarding "unbounded Merge" (to which we return in the next section), which he describes as "not only a genetically determined property of language, but also unique to it" (Chomsky 2007b: 5). Thus, it seems that there is an asymmetry here between FLN and UG; for if a language property is genetically unique, then it is also genetically determined, although the converse need not obtain. To put it differently, every FLN property is a UG property, but *not* every UG property is an FLN property. If true, then this is something that will have to be taken into account when attempting to evaluate the SMT from an evolutionary standpoint. On the other hand, however, Chomsky (2008a: 133) defines UG as "the theory of the distinguishing features of human language," and elsewhere he says that "UG consists of the mechanisms specific to FL, arising somehow in the

course of evolution of language" (Chomsky 2007b: 3). There appears to be room here for suggesting that FLN and UG are definitionally identical and it is clear that there is considerable scope for confusion regarding this matter. As we shall see in the next section, even Chomsky himself fails to achieve clarity in this context.

4.5 The Merge-only hypothesis

Our main task in this section is to examine Chomsky's position on the uniqueness of language. What we have seen in this and the previous chapter suggests that he entertains three different hypotheses. First, in his linguistic discourse, we find the hypothesis that UG is restricted to Merge. Second, in his interdisciplinary discourse, we encounter the hypothesis that FLN is limited to recursion. Finally, the third hypothesis, which is common to both discourses, may be regarded as a "default hypothesis"; it is one that, as its name implies, is entertained only when the other two hypotheses encounter difficulties or are seen as implausible for other reasons. It involves a variety of "just-in-case" hypotheses, including that which maintains that the set of properties that are unique to human language may be empty – the reader will recall the parallelism between FLN and UG described at the end of the previous chapter. The three hypotheses will not receive equal consideration in this section. Since the Merge-only hypothesis relates closely to the SMT, it receives the greatest emphasis. Less emphasis is placed on the default hypothesis, and as to the recursion-only hypothesis, we will have little to say since it has been discussed extensively in previous sections.

Let us begin by asking a very simple question: Is Merge specific to language? Chomsky (2007b: 5) admits that a negative answer to this question is suggested by the fact that this computational operation seems to have antecedents in other domains, notably the system of natural numbers. However, he speculates that the core component of the mathematical capacity, arithmetic, may somehow be parasitic on language, in the sense that it is derivative from it.[9] "If the lexicon," Chomsky (2007b: 5) writes, "is reduced to a single element, then Merge can yield arithmetic in various ways." In other words, arithmetic is possible only because of the existence of language.[10] A full assessment of this claim and its intended implication for the specificity of Merge requires, *inter alia*, a discussion of the foundations of mathematics, a topic that is beyond the scope of this work. Suffice it to make here three observations that cast doubts on this claim and, consequently, on what I take to be its intended implication, *viz*. that Merge is specific to language despite the fact that it also shows up in mathematics.

First, while the reduction of the lexicon to a single element, as Chomsky suggests, may give rise to a form of arithmetic, it would be rash to take this as an indication that arithmetic is derivative from language. There are two reasons for caution here. First, all Merge does in this context is generate an infinite series by computing upon n discrete elements ($n \geq 1$), which yields the property of *discrete infinity*. Note that while the *infinity* arises from the unbounded application of recursive Merge, the *discreteness* is merely a property of those elements to which Merge can *only* apply. Chomsky rightly suggests that Merge can apply recursively to a one-membered lexicon to generate each immediate successor of its own output *ad infinitum*, but this suggestion does not apply to the real number system; for there is simply no such thing as an immediate successor of a real number.[11] Therefore, the fact that Merge generates an infinite series of elements (e.g. integers, linguistic expressions, etc.) says nothing about the evolutionary precedence of language over arithmetic, but only shows that Merge functions on a discrete basis, that is, it can only apply in domains where discrete elements are available (e.g. lexical items in language, integers in arithmetic, etc.). Now, since arithmetic is not exhausted by discrete numbers, there is no warrant for claiming that arithmetic is derivative from language. In fact – and here comes the second reason for caution – there is no warrant even for holding that the natural number system is derivative from language. For if we grant that the number system originated with primitive humans who counted on their fingers, and even if we assume that this historical event took place after the emergence of language, we should still guard against the identification of the social-historical order of the development of mathematics in human culture with the natural-biological order of its development as a cognitive capacity in the species. There is no reason to discard the possibility that a number system (*qua* cognitive system) evolved before language evolved, and that the fact that the cultural history of mathematics began with the integers is merely due to these numbers being discrete countable magnitudes, which makes them a ready target for representation by the discrete language system that evolved at a later stage. This possibility is defended by Gallistel *et al.* (2006), who argue that it is the real numbers (and not the natural numbers) that are psychologically primitive. They provide ample empirical evidence showing that the real number system is shared with other species, and based on this evidence they suggest that "when language evolved, it picked out from the real numbers only the integers, thereby making the integers the foundation of the cultural history of the number" (Gallistel *et al.* 2006: 247). If this is true, it follows that the property of discrete infinity (and, therefore, Merge) cannot be said to be special to language, nor seen as unique to humans.

Second, Hinzen (2009: 133), who is very sympathetic to minimalism, reads Chomsky's speculation on language and arithmetic as saying that "Merge in language ... is simply an instance of a more general operation that generates the natural numbers in arithmetic, too, yielding a discretely infinite space in both cases." He points out that Merge creates a single-dimensional space in both arithmetic and language, in the sense that it yields only one category of objects (one ontological kind). As such, Hinzen argues that Merge will not be ontologically productive. He sees this outcome as unsatisfactory, for "just as Merge is a far too poor basis to yield the rest of arithmetic (all that goes beyond the integers), a naïve adaptation of Merge ... to the domain of language does not give us its full richness either" (Hinzen 2009: 133). He suggests that to go further one needs more operations to create a multi-dimensional space with new kinds of objects (i.e. new ontological kinds), in the same way that the two basic arithmetic operations of subtraction and division create new mathematical spaces with new kinds of objects (e.g. the negative and rational numbers). In short, Hinzen (2009) contends that "if arithmetic is to be an evolutionary offshoot of language, as Chomsky (2005[d]) plausibly suggests, basic structure-building operations in language might therefore be richer as well."[12] Whatever the (de)merits of Hinzen's view might be, it appears to entail that Merge is neither specific to the faculty of language nor is it sufficiently differential as a defining property of this faculty.[13]

The third observation to be made in connection with Chomsky's speculation regarding the relationship between language and mathematics is that such speculation might be seen to predict an empirical connection between severe agrammatic aphasia and syntactic mathematical impairments. Indeed, if arithmetic is parasitic on language, it might be thought that an aphasic patient with no sensitivity to the structural dependency of linguistic expressions, or to the recursive application of linguistic rules, should also be insensitive to these same properties in mathematical calculations. Unfortunately, experimental studies on this issue show inconsistent results. For example, a study conducted by Varley *et al*. (2005) indicates that aphasias resulting in insensitivity to structural dependency and recursiveness may leave mathematical computational ability intact, while another study by Semenza *et al*. (2006) concludes, not only that aphasia is correlated with acalculia, but that the type of the latter depends on the type of the former. Moreover, Chomsky himself appeals to the distinction between competence and performance to discredit empirical evidence which seems to undermine his speculation concerning the relationship between language and mathematics. Thus, although he acknowledges the existence of some empirical phenomena that seem to undermine

his speculation, like the "apparent dissociation with lesions and diversity of localization," he adds: "The significance of such phenomena, however, is far from clear. They relate to the use of the capacity, not its possession; to performance, not competence" Chomsky (2007b: 5). It is not clear to me, *pace* Chomsky's skepticism, how else *empirical evidence* of any competence will be obtained if not by investigation of some kind of performance.

In light of the above observations, especially the first two, it may be tempting to conclude that Merge may not be special to language. Chomsky (2007b: 5) does not deny that this may indeed be the case, suggesting that "[t]he conclusion that Merge falls within UG holds whether such recursive generation is unique to FL or is appropriated from other systems." Now this suggestion is confusing when compared with Chomsky's own definition of UG from the previous section, namely that "UG consists of the mechanisms specific to FL" (Chomsky 2007b: 3). There is a clear inconsistency between these two statements. If Merge falls within UG whether or not it is unique to FL, then how could UG possibly be defined as the mechanisms specific to FL? Confronted with this question Chomsky (p.c.) has said:

> UG is by definition the theory of the genetic component of the language faculty, which means of course the part that is specific to FL (the process of cell division is involved in the language faculty, but UG is not concerned with that). We might similarly define UMV as the theory of the genetic component of the faculty of mammalian vision; that is, what's specific to this faculty. That's entirely consistent with the assumption (false as far as we know) that Merge is recruited from other systems and is specific to FL in that it yields structured expressions mapping to the interface.

As it stands, this reply is not satisfactory. If I am interpreting Chomsky correctly, he seems to be suggesting that the language-specificity of Merge can be saved by the assumption that such a recursive operation may be special to FL, not in the sense of belonging exclusively to language, but in the sense of "yield[ing] structured expressions mapping to the interface." But this trivialises the notion of language-specificity. For, by the same token, it could also be suggested that Merge is special to the thought system in that it yields structured thoughts, or that it is special to the vision system in that it yields structured images, and so on, which would indicate that the language-specificity of Merge amounts to no more than the nature of inputs and outputs of the computational system in which it happens to operate. If this is true, then Chomsky's reply amounts to saying that Merge is language-specific because it operates in the language

faculty – a faculty whose main function is to "yield structured expressions mapping to the interface." If this is indeed what Chomsky suggests, it would seem that we are left with an unacceptable circularity.

It should be noted that, as a consequence of this circularity, the empirical question of whether Merge is unique to language cannot even be stated without implying an affirmative answer, which implies that (at least one aspect of) the SMT is irrefutable in principle (cf. the conclusion we have reached in the previous chapter, Section 3.6).

Setting aside the issue of circularity, empirical difficulties in testing assertions or hypotheses about the biological uniqueness of language remain. Let us now turn to some of these difficulties.

To begin with, let us suppose, for the sake of argument, that we want to test the soundness of the Merge-only hypothesis by focusing on language mechanisms other than Merge, say the technology of AGREE(ment), probe-goal relations, feature valuation, etc. If we find no evidence for the presence of this technology in non-linguistic cognitive domains or in non-human communication systems, we may conclude from this that the hypothesis is probably incorrect. We feel justified in using this criterion because we believe that the Merge-only hypothesis, as understood in the previous section, entails the presence of such technology outside the domain of human language. We may also want to consider the empirical plausibility of the recursion-only hypothesis of Hauser *et al*. (2002), in which case the same criterion applies but in a reverse manner: i.e. the presence of that technology outside FLN should be taken as a undermining this latter hypothesis. Theoretically speaking, the task appears simple and well-defined. But recall that we are dealing with an *empirical* task, and unless we are provided with a criterion that enables us to identify instances of, say, Agree in domains external to language, the task itself cannot be performed, and, therefore, the hypothesis would not be falsifiable. Indeed, we are given the almost impossible task of searching for abstract operations in the realm of nature, for it is not clear how we might go about searching for, say, probe-goal relations or case assignment under such a relation in animal communication systems or non-linguistic cognitive systems in general.

This difficulty in testing the empirical validity of the Merge-only hypothesis is aggravated by a distinction which Chomsky introduces in his more recent work between Merge and *unbounded* Merge. Consider, for instance, how he understands the specificity of language:

> [T]he crucial thing about language is not Merge; it is unbounded Merge. So just the fact that things are hierarchic elsewhere doesn't really tell

> you anything. They have to be *unboundedly* hierarchic. (Chomsky in Piattelli-Palmarini *et al*. 2009: 52, emphasis in original)

This seems to suggest that what makes Merge special is not so much the embedding hierarchies it gives rise to, but the *self*-embedding that yields an unbounded way in which these recursive structures "manifest" themselves. Taking this suggestion seriously makes the Merge-only hypothesis even harder to evaluate. Indeed, the difficulty with this suggestion arises from its implication that access to competence whatever that might mean is the only means of inquiring into the unboundedness of Merge. But as far as animal cognition is concerned, this sort of inquiry is uncertain at best and impossible at worst.[14]

Our discussion so far has focused on some of the conceptual and empirical difficulties surrounding the Merge-only hypothesis. Some of these difficulties suggest a negative answer to the question we posed at the beginning of this section, whereas others indicate no definite answer. Now, if recursive Merge is not the key evolutionary step that gave rise to language, what other option is available to Chomsky to sustain his thesis that either the genetic component of language is non-empty or language acquisition is a miracle (see Section 3.7)? In other words, if in any case *something* must be specific to language, what could this something be?

Chomsky's (2007b: 5) answer to this question is that if Merge is not specific to language, then "there still must be a genetic instruction to use Merge to form structured linguistic expressions satisfying the interface conditions."[15] Clearly, the "something" refers to the "genetic instruction to use Merge." Chomsky seems to be trying very hard to resist having to take this option seriously. Examples of this can be seen in the preceding discussion. His attempt to derive arithmetic from language, his appeal to the distinction between Merge and unbounded Merge, his recourse to the competence/performance distinction, all reflect the same purpose – to preserve the integrity of the Merge-only hypothesis.

The reason why Chomsky might feel reluctant to resort to the (so far undiscovered) "genetic instruction to use Merge" should be obvious; taking this option would mean that we are left with the uncomfortable circularity we noted above in connection with the language-specificity of Merge. To see this, let us accept for the sake of argument that what is specific to language may be limited to a genetic instruction allowing Merge to satisfy interface conditions by forming structured linguistic expressions. Intuitively, this implies that there must be other genetic instructions allowing Merge to apply in domains other than language, say, in vision or cognition at large (note that there should

be nothing surprising about this implication since the very idea of a genetic instruction has been introduced on the grounds that Merge may not be specific to language). The question that immediately arises is, what makes the genetic instruction to use Merge in the language domain different from another instruction which allows its instantiation in another domain? The nature of the domain in which Merge applies seems to be the only answer that can be given to this question, which brings us back to the circularity problem.

I suspect that the body of literature noted earlier (Section 4.2), arguing for the presence of recursion outside the syntax of human language and in nature at large, together with Hauser *et al*.'s own concession that recursion might have precursors in animal navigation systems, leaves Chomsky with no alternative but to adopt this very unsatisfactory position (*viz*. the genetic instruction to use Merge in the language system).

To summarize, the aim of this chapter was to evaluate the SMT from an evolutionary perspective. We distinguished between the recursion-only hypothesis and the Merge-only hypothesis – two hypotheses that are central in Chomsky's linguistic and interdisciplinary discourses, respectively. Contrary to a widespread opinion, we have argued that these two hypotheses are not equivalent. In particular, it has been suggested here that recursion is much more general and inclusive than Merge, assimilating a range of technology beyond the latter into the language-specific recursive device. In consequence, this chapter has argued that the recursion-only hypothesis has an empirical content different from that of the Merge-only hypothesis, and that the latter is beset with conceptual and empirical difficulties. The next chapter continues the evaluation of the SMT by focusing on its explanatory status.

5 *The SMT as an explanatory thesis*

5.1 Introduction

The previous chapter has focused on one aspect of the strong minimalist thesis (SMT) – the recursive operation Merge. We now turn to the remaining two aspects of this thesis: interface conditions and optimal computation. These form the *explanantia* of a minimalist explanation and will be the focus of this chapter.

The organization of the chapter is as follows. The first three sections deal with interface conditions. I begin by a discussion of the minimalist appeal to an interface-based explanation (Section 5.2), and proceed to identify two major problems with such an explanation: tautology (Section 5.3) and teleology (Section 5.4). The remaining part of the chapter is devoted to optimal computation. In Section 5.5, I first consider the role of optimal computation and argue that it lacks any independent explanatory status. Next, in Section 5.6, I examine some of the attempts to ground optimal computation in physical principles, and I argue that they fail to offer genuine correlates between the principles of language and the laws of physics. I then proceed in Section 5.7 to consider the explanatory status of the kind of physics which some minimalists believe is relevant to the minimalist program (MP), and I argue that it is of a kind that has been described as teleological and is no longer acceptable in modern physics.

5.2 Minimalist explanation: interface conditions

The SMT embraces what Chomsky (2004b: 158) describes as "two sources for principled explanation": interface conditions and optimal computation. Leaving aside the latter for the time being, consider what Chomsky (2004b: 158–9) has to say regarding the former:

> The language organ is going to be interacting with [the performance] systems: They impose their own requirements – that's the interface conditions. And if you can show that some property of the internal

> state of the language faculty satisfies those conditions, you also have a principled explanation.

With regard to this passage two notable questions arise. What justifies the minimalist appeal to an interface-based explanation? How plausible is such an explanation? I shall deal with the first question here and tackle the second in the next two sections.

Let us begin by reminding ourselves of the evolutionary considerations that we have suggested may, at least partly, underlie the shift to minimalism. As we saw in Chapter 2 (Section 2.6), Chomsky (2008b, and other places) puts forth an argument to the effect that if language is part of our genetic endowment, then it ought to have evolved in some way or another, but since its emergence in the course of evolution appears to be quite recent in evolutionary terms, it follows that not much of it could have evolved.

Given this argument, one can understand why minimalists might feel attracted to the notion of "interface" *qua explanans*. This is because, unlike adaptationist explanations based on natural selection, an interface-based explanation refrains from invoking any path-dependent evolutionary history; rather, it posits that certain language properties are not rooted in genetics but arise merely as by-products of the interaction between the language faculty and its neighbouring systems (cf. e.g. Chomsky 2005 and Hinzen 2006b). This may provide a plausible answer to the first of the above questions, as it suggests that invoking the notion of interface as a category of explanation is justified by the limited time frame of language evolution.

However, while this argument from the evolutionary time frame of language has obvious attractions, it needs to be treated with some caution. The reason for this has to do with a crucial premise in Chomsky's argument, according to which the emergence of language is quite recent; according to one estimate favored by Chomsky (2002: 148–9), language emerged in the species only about 100,000 years ago. This, however, need not be the case. Various hypotheses have been offered which differ significantly in this respect, with estimates ranging from millions of years (for instance, Pinker and Bloom 1990 appear to favor an estimate of between 3.5 and 5 million years) to a mere 40,000 years. Johansson (2006), in his review of the literature on this topic, notes that there is no consensus, neither on when language evolved, nor on whether its evolution was sudden or gradual. Accordingly, this justification for interface-based explanation rests on an unverified (and controversial) assumption about the timing of language emergence, and it is for this reason that caution is required.

Regardless of the empirical question of when language evolved, there are factors other than evolutionary considerations that must be taken into account if a proper explanation of language and its properties is to be achieved. One such factor has to do with boundary conditions arising from the requirement that the language system has to interface with other systems in order to be usable by them. As Chomsky (2000c: 25) puts it:

> The extralinguistic systems include sensorimotor and conceptual systems, which have their own properties independent of the language faculty. These systems establish what we might call "minimal design specifications" for the language faculty. To be usable at all, a language must be "legible" at the interface: the expressions it generates must consist of properties that can be interpreted by these external systems.

If this is true, then an interface-based explanation is called for. More specifically, the explanatory role attributed to the interfaces derives its legitimacy from the claim that the legibility conditions imposed on language by external systems are a *sine qua non* of its usability by these systems. But here again caution is necessary, for there is no reason why we should conceive of the interaction between language and other systems in terms of certain requirements that are imposed on the former by the latter, rather than the other way round. Let us see what this means in some detail.

Chomsky (2002: 108) starts by claiming that language "has to interact with [the external] systems, otherwise it's not usable at all." But how do we move from this claim to the assertion that the outside systems impose legibility conditions on language? The passage just cited from Chomsky continues:

> So, we may ask: "Is [language] well designed for the interaction with those systems?" Then you get a different set of conditions. And in fact the only condition that emerges clearly is that, given that the language is essentially an information system, the information it stores must be accessible to those systems, that's the only condition. We can ask whether language is well designed to meet the condition of accessibility to the systems in which it is embedded. Is the information it provides "legible" to those systems?

It seems to me that there is a large step from the assumption that there must be *some kind* of interaction between language and its neighbouring systems, to the conclusion, implicit in the passage just quoted, that language is designed for meeting interface requirements. Why could not this conclusion be reversed? That is to say, there seems no reason why the interaction should not be viewed in

terms of usability requirements that are imposed on the outside systems. Alternatively, it is conceivable that the external systems impose no conditions on language and that the latter comes equipped with instructions as to its use which the external systems happen to decode and execute for their benefit. Of course, against this, one might argue that, if the instructions provided by language are not based on what the external systems demand (i.e. on legibility conditions), the fact that language is used by these systems would be a mystery. But it is no less mysterious to assume (as minimalists do) that the language faculty contains only those properties which satisfy legibility conditions.[1]

I should make it clear that I am not advocating one proposal over another. In fact, my intention is quite the opposite. The fact that there are (at least) two conceivable ways of viewing the interaction between language and the external systems suggests caution in advocating either direction for this interaction without additional evidence. However, a position opposed to that advocated by Chomsky has been proposed by Hinzen (2009). Let us briefly examine his views on the relationship between syntax and the thought system.

Hinzen argues that language could "be used, even if such independently constituted systems [i.e. the 'outside systems' of thought] did not exist" (Hinzen 2009: 127). He proposes what he calls a "radical approach" to syntax, in which there is "no semantic component, no independent generative system of 'thought,' no 'mapping' from the syntax to such a system, no semantic 'interface'" (2009: 128). Under this approach, syntax is conceived of "as the skeleton of thought," that is, it "literally constructs a thought and gives it its essential shape, much as our bones give shape and structure to our body" (2009: 129).

Clearly, Hinzen is adopting a point of view which is the complete opposite to that of many minimalists (cf., however, Uriagereka's "co-linearity thesis" mentioned in the next section). For he seems to be suggesting that the thought system (i.e. the conceptual-intentional system) is the way it is because the language faculty is the way it is, not the other way round as implied by the SMT. Although this view may be correct, it is, of course, speculative and no less so than that advocated by Chomsky. Indeed, it might be felt that the cost involved in adopting Hinzen's radical approach is too high. For one thing, by depriving the SMT of one of its *explanantia*, this radical view places the burden of minimalist explanation entirely on optimal computation, which would call into question the grounds on which much of the reduction of the descriptive apparatus of pre-minimalist approaches has been based.[2] For another thing, if syntax does indeed "construct" semantics, and if we (as humans) share with other species part of that semantics, it follows that other species should also have syntax, though one which is less developed than ours. Hinzen (2009: 130)

seems to be aware of this latter consequence, for he says "that to whatever extent non-human animals partake in the systematicity and propositionality of human thought, they partake in whatever fragments of the computational system are needed to think them." But we have observed that there is so little consensus over the evolution of language that it is difficult to draw any conclusion on this topic. Moreover, given our ignorance of what the "thought system" is, reference to "the skeleton of thought" involves a considerable amount of handwaving. Since adopting Hinzen's radical view on syntax requires one to decide on these two poorly understood topics, the caution I have been urging in this discussion becomes imperative.

The discussion so far has suggested that the appeal to interface-based explanation may be justified but that this justification must remain tentative. Nothing suggested so far, however, allows us to assert the plausibility of such explanations *qua* explanations. We turn immediately to this matter.

5.3 Tautology

To help frame the discussion, let us recall what Chomsky says in the passage quoted in p. 109. There, he suggests that "if you can show that some property of the internal state of the language faculty satisfies [legibility] conditions, you also have a principled explanation." This suggestion implies that it is possible that one may *fail* to show that some property of language satisfies legibility conditions, and in such an event, one may fail to arrive at a principled explanation. That such a conclusion can be reached, however, I maintain is *impossible* in view of what I regard as the tautological nature of interface-based explanations. If this characterization is correct, it implies that the minimalist goal of achieving a principled explanation in terms of the interfaces is not an empirical one; whatever property we attribute to language, it will always be possible to explain it by postulating some legibility condition.

Before we go further, let me clarify what I have just said. I am certainly *not* suggesting that the issues we are about to discuss are not empirical; indeed, legibility conditions and their relation to language are an empirical, though very complex, issue. Rather, I am suggesting that so long as minimalists resort to interface-based explanations *as currently conceived and practiced*, their aim of achieving a principled explanation is devoid of any empirical content: *this is a criticism of the form of explanation they provide, not of their subject-matter*. I suspect that when the enormous gaps in our knowledge of interfaces are filled in, if, indeed, this proves to be the path that understanding follows, the kind of tautological explanations that are currently on offer in minimalism will only

serve to remind us how little we knew of what we were trying to explain. With this important clarification in mind, let us now return to our discussion of interface-based explanation.

It is certainly not uncommon to come across an argument which invokes some legibility condition to explain why the computational system operates as it does. As one minimalist has complained, "it is still rather customary ... to postulate some constraints at the ... interfaces to performance systems and let them explain particular attested syntactic patternings" (Narita 2009: 1771). But if we ask why this should be so, a common answer would be something like: if the system fails to behave in such-and-such a way, its output crashes at the relevant interface. Yet observe that what this amounts to is really only that the derivational system yields a representation which *violates* the condition of convergence at the relevant interface. Thus, what we have here is a clear example of tautological explanation in which the *explanandum* and the *explanans* are logically equivalent; $p \rightarrow q$ *because* $\neg q \rightarrow \neg p$. Specious explanations of this form abound in the minimalist literature, but for reasons of space only two examples are mentioned here.

Freidin (2007: 55) provides one instance:

> The derivation to LF does require the covert movement of wh-phrase subject to [Spec, CP] in order to create a proper quantifier/variable structure. Otherwise, the derivation crashes at LF because the result violates Full Interpretation (FI) – a quantifier that does not bind a variable cannot be interpreted.

In other words, the need to satisfy Full Interpretation (FI) by "creat[ing] a proper quantifier/variable structure" at LF explains why the element involved undergoes a covert movement; because if this element fails to move, FI will be violated and thus the derivation will crash at LF. This sort of non-explanation can also be found in Freidin and Vergnaud (2001: 642), who write:

> Case features are extraneous to interpretation at the LF interface at least and therefore should be eliminated from the derivation before LF. Otherwise, these uninterpretable features at the LF interface will violate FI, causing the derivation to crash. Thus legibility conditions imposed by the cognitive system that interfaces with [the computational system] determines (*sic*) how derivations proceed with respect to Case features.

In this and the previous example, the only evidence for the *explanans* on offer is its corresponding *explanandum*. This kind of explanation is clearly circular,

and as such it is no better than an explanation in which a house fire is explained by reference to an electrical malfunction, the only evidence for which is the house fire itself. Circular explanations like these can never fail to be true, and, therefore, they can never be empirically falsified. Given the tautological form in which they are stated, they can explain everything, and it is for this reason that they explain nothing.

Freidin and Vergnaud (2001: 643) affirm that an interface-based explanation is "a more promising explanatory account than the postulation of various constraints internal to [the computational system] that basically reflect the complexity of the phenomena in an essentially descriptive fashion." But observe that the same can be said about an explanation that appeals to interface conditions. If every syntactic property is explained by postulating an interface condition with which it is said to correlate, there will be as many postulates as syntactic properties. As Hinzen (2006a: 6) puts it, if "syntax is motivated by interface conditions imposed by outside systems . . . then the syntax resulting from and explained by this can only be as rich as these very outside systems." He correctly observes (2006a: 7) that resort to such a mode of explanation makes it "unclear . . . whether we have explained language, or explained it away."

I suspect that the root of this weakness comes from the way Chomsky seeks to justify the appeal to the notion of legibility. As a passage quoted earlier (p. 111) indicates, this justification relies on an appeal to the definition of language itself. In fact, every legibility condition we can think of is suggested in great measure by what we know about language itself, something Chomsky (2004b: 165) himself acknowledges: "We don't know very much about the language-external conceptual-intentional systems," because "it's almost impossible to study them except through language," and therefore we "don't get independent information about them." Yet it is these very systems that are supposed to account for language and its properties by imposing on the latter their own legibility conditions. It is simply tautological to attempt to infer the existence and nature of legibility conditions from properties of language and then proceed to use these same conditions to explain those properties.

As further support for this position, I will now argue that the charge of panglossianism that has been levelled against adaptationists is as applicable to minimalist explanation as to Darwinian explanation.

In the previous chapter, we observed that Pinker and Bloom (1990) defended neo-Darwinian explanation against the charge of panglossianism. "Adaptation and natural selection," they complained, "have become dirty words," and those who invoke them are "open to easy ridicule as a Dr. Pangloss telling Just-so

stories" (1990: 710–11). Indeed, adaptationist explanations have been vehemently ridiculed by many scholars (see, for example, Gould and Lewontin 1979; Fodor 2007). Chomsky, understandably enough, joined in this ridicule. He writes that the stories about how language evolved "are free and, interestingly, they are for the most part independent of what the language is" (Chomsky 2002: 146). Such stories, according to him, resemble the Lamarckian story about giraffes' necks, where

> giraffes get a little bit longer neck to reach the higher fruits, and they have offspring and so giraffes have long necks. It was recently discovered that this is apparently false. Giraffes don't use their necks for high feeding, end of that story. You have to figure out some other story: maybe sexual display like a peacock tail or some other story, but the point is that the story doesn't matter. You can tell very plausible stories in all sorts of cases but the truth is what it is. You can tell stories about the planets, as the Greeks did, in fact: nice stories, but things don't work that way. (Chomsky 2002: 149)

According to this passage, adaptationist explanations are nothing more than "Just-so stories" that do not explain anything. This may well be true, but the question remains as to why interface-based explanations should be different in this respect. Consider the following passage from Chomsky (2000b: 98):

> 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.

Chomsky seems to suggest here that the way to proceed in order to explain the language faculty is to make tentative proposals about the two sides of the interfaces (i.e. about the internal properties of the language faculty and the external properties of the performance systems); owing to the complexity of language, that is all one can hope for. If this is true, there is no reason why this aspect of the minimalist approach should not be subject to the same criticism as the adaptationist approach. As noted above, Chomsky has it that adaptationist stories about language evolution are "free" and largely independent of what language actually is. But it should be clear from the discussion above that the

same can be said about an interface-based explanation; virtually every syntactic property can be explained by an "interface story" referring to legibility constraints. Accordingly, one may rightly argue that the charge of telling "just-so-stories" is as applicable to this feature of minimalist explanation as to Darwinian explanation.

In fact, from an empirical point of view, minimalist explanation seems to be in a worse position than adaptationist explanation. The latter, as Chomsky himself implies above, is open to falsification, while the former, as noted in Chapter 3 (Section 3.6), advocates a strong minimalist thesis without specifying independent grounds on which such a thesis can be falsified. Moreover, in the case of the adaptationist explanation of the giraffe's long neck, both the *explanans* and its *explanandum* are at least in principle amenable to experimental intervention; it is possible to alter the giraffe habitat or modify the genetic code of a giraffe embryo. By contrast, this is not the case for interface-based explanations; as the last passage cited above makes clear, both the language properties and the external conditions remain as mere postulations. Furthermore, it is possible to study the giraffe's habitat and obtain information about its behavior independently from its internal anatomy or genetic make-up. In contrast to this, it seems likely that the study of conceptual-intentional systems cannot be carried out independently from the study of language (cf. the fragments cited from Chomsky 2004b: 165 in p. 115).

In light of the above, it is perhaps not surprising that some authors, including prominent minimalists, have argued for the elimination of legibility conditions. Hinzen (2009) is one example already discussed above. Another is Uriagereka (2008: 224), who has argued for a view similar to that of Hinzen. He proposes a so-called *co-linearity thesis*, the radical version of which suggests that semantics is dynamically constructed by the syntactic system. Narita (2009: 1772) provides a further case. Dissatisfied with the explanatory status of legibility conditions, he argues that the SMT should be stated in a more simplified way, namely: "Language is optimal in terms of the third factor." It should be noted that Narita (2009: 1771) does not consider legibility conditions to be part of the third factor and defines the latter in terms of considerations that are related only to optimal computation. However, we have already referred to the enormous cost involved in proposals like these and the uncertainty they lead to (cf. the discussion in the previous section of Hinzen's radical approach to syntax).

Our discussion so far has concentrated on the tautological character of interface-based explanation. We have seen that the tautology results from an interplay between the two notions of "convergence" and "crash." This interplay, as the following discussion will argue, reveals a further difficulty for this type of

explanation, namely its *teleological* character. Eight years ago, I expressed my concern to Chomsky about the seemingly teleological character of the strong minimalist thesis (SMT). It seemed to me at the time that this thesis, at least in the formulation "Language is an optimal (computational) solution to legibility conditions," ascribed to language some sort of *finality*, in the sense that it does not appear to differ much from a teleological, adaptationist thesis, which may be viewed as asserting that a biological trait is a successful (adaptationist) solution to environmental conditions. Interestingly, Chomsky did not explicitly deny that there was any teleological aspect involved in the minimalist thesis, although he seemed to believe that the teleology was only apparent. In a personal communication, he suggests that "[t]he feeling of teleology comes from the fact that the 'language organ,' like others, has to interact with others, which set some conditions on what the system may be: in the case of language, with other cognitive systems." However, it is my contention that, *pace* Chomsky, the "feeling of teleology" is for real and not just an impression, as the discussion that follows will seek to demonstrate.

5.4 Teleology

Let us first be clear on the sense of teleology before we ask whether or not an interface-based explanation amounts to a teleological explanation. Teleology can be defined in various ways, but for our purposes we will define a teleological explanation as an "explanation-by-function," in which we explain something by reference to the function it fulfils or by the purpose it serves. Thus, the states of affairs in which a student reads a book, a bird has wings, or a solar eclipse takes place receive functional/teleological explanations by an appeal to certain functions or purposes: to pass an exam, to fly, or to remind people of their sins. Functional/teleological explanations abound in many fields, scientific and otherwise, and, as we shall see in Section 5.7, they were frequently adhered to in the physics of the pre-nineteenth century.

We now turn to consider whether the appeal to interface conditions in motivating the design of language constitutes an instance of a teleological explanation. The reader will recall from Chapter 2 (Section 2.4) that minimalism defines the interaction between the faculty of language (FL) and the performance systems in terms of the *function* and *quality* of that interaction; that is to say, the interaction is directed towards satisfying the *interface conditions* imposed on the FL by the performance systems, while the way in which it is achieved is assumed to be *optimal*. Continuing to set optimality aside, the first passage cited in Section 5.2 (p. 109) can be read as suggesting

that a property P of language receives a principled explanation insofar as it can be shown that language has P in order to satisfy some interface condition C; that is, P exists in order to satisfy C.

We have already seen examples in the previous discussion, where the principle of full interpretation (FI) has been invoked to account for certain language properties. Observe, however, that teleological considerations are not restricted to interface-based explanations that appeal to the notion of "convergence," but they also extend to those invoking the notion of "crash." Indeed, this latter notion seems at odds with a framework that views the syntactic system as a "blind watchmaker," that is to say, one which must satisfy legibility conditions, but in doing so it must also not "look ahead" and anticipate what may go wrong at the interface. As we shall see later (Section 5.6), the teleological notion of "look-ahead" is also induced by explanations that refer to optimal computation.[3]

It may be objected that all I have shown is that minimalist explanation adheres to a functionalist/teleological terminology, and that this in itself does not make minimalist explanation functional/teleological in character. To be sure, Chomsky (2000a: 9, 2000b: 94) invites us to think of the task of satisfying design specifications as a problem for a "super-engineer," but we are not to suppose from this that he is committed to the view that the emergence of language is due to the conscious design of an agent; rather, the evolutionary fable of a super-engineer is intended for expository purposes only and does not reflect some deep commitment to teleology. However, there *is* more than just terminology at stake here. This is a functionalist teleology according to which the design of language is not arbitrary but can be explained in terms of the functions it serves; targeted properties gain their legitimacy as being part of language design insofar as they serve an interface condition. Indeed, this facet of minimalist explanation promotes an "ends justify the means" approach to syntax. Moreover, it has been well known since Darwin that natural as opposed to artificial teleology does not require the postulation of a self-conscious design agent. Indeed, as Ayala (2007: 27–48) points out, Darwin's greatest discovery was the idea of "design without designer." If this is true, and since Darwinian explanation in terms of natural selection is more or less commonly regarded as being teleological, it follows that teleological explanation is not confined to the argument from intelligent design. Accordingly, although Chomsky's evolutionary fable of a super-engineer is intended for expository purposes only, recognition of this fact does not entail that interface-based explanations of language design provide a teleology-free form of explanation.

It should be noted that what I am maintaining here is that, as far as the functional/teleological status of explanations are concerned, interface-based minimalist explanation is no different from that offered by the neo-Darwinians who propose that language evolved because of its role in communication (e.g. Pinker and Bloom 1990; Pinker and Jackendoff 2005). Of course, Chomsky is not a *functionalist* in the standard sense, that is, someone who is committed to the view that the structure of language is better understood in terms of its use in communication (cf. Carnie and Denton-Mendoza 2003; Golumbia 2010). As noted already (see, in particular, Section 2.6), Chomsky's position has always been unambiguous in its denial that communicative use was primary to language or that language evolved for communication – a position that he has articulated since at least the publication of his *Cartesian Linguistics* (1966). However, there is more to being a functionalist than just being committed to viewing language as an instrument of communication, just as there is more to being a teleologist than just being committed to an argument from intelligent design. For let us not forget that in relegating the communicative function of language to secondary status, Chomsky was also proposing a different function as primary to language. For example, in Chomsky (2002: 106–7), we find:

> If you take a standard functionalist point of view, you would ask: "Is the system designed for its use? So, is it going to be well designed for the uses to which people put it?" And the answer there is "apparently not" . . . but it has to be designed well enough to get by. That's all that we discover: it's designed well enough to get by. That raises the question: can we find other conditions such that language is well designed, optimal for those conditions? I think we can, from a different perspective. So instead of asking the standard functionalist question, "is it well designed for use?," we ask another question: is it well designed for interaction with the systems that are internal to the mind? It's quite a different question, because maybe the whole architecture of the mind is not well designed for use.

What Chomsky appears to be doing here is simply substituting one function for another; i.e. interaction with the thought system and other internal systems for communicative use. Of course, he speaks here of two different questions (*viz*. "Is language (well) designed for use?" and "Is language (well) designed for interaction with other cognitive systems?"), but it should be clear that the only difference between the two questions is the assumed function of language. In fact, from an evolutionary perspective, the difference is even narrower. For just as adaptationists argue that language evolved because of the need to

communicate one's mental states to others (recall the claim of Pinker and Bloom (1990) in the previous chapter, Section 4.2), Chomsky and his followers argue that it evolved to communicate one's mental states to oneself (recall Chomsky's (1966: 13) long standing subscription to the Cartesian assumption that the primary function of language is the expression of thought; see also Chomsky 2002: 148).[4] Thus, the difference here lies not in the communicative function *per se*, but rather to whom the communication is directed. When the indirect object of the verb of communication is suppressed, the difference disappears, as in the following remark from Hauser *et al*. (2002: 1574): "The question is not whether FLN *in toto* is adaptive. By allowing us to communicate an endless variety of thoughts, recursion is clearly an adaptive computation."

The aim of the preceding discussion was to argue for the functional/teleological character of minimalist explanations which appeal to interface conditions. We now turn to the second and final task of this chapter – a critical evaluation of the explanatory role of optimal computation in minimalist arguments.

5.5 Minimalist explanation: optimal computation

A view widely held by minimalists is that the laws of physics are sufficient to ensure certain aspects of "good design" in organisms without the need for special mechanisms that are organism-specific. This view underlies the optimism regarding the prospects of unification between linguistics and physics, to which many minimalists aspire. Such aspiration manifests itself in the various attempts to substantiate claims of the existence of genuine connections between the principles of language and the laws of physics. Before we assess these claims and the status of the physics that is providing the explanatory basis, we will first consider the extent to which optimal computation is supposed to provide a kind of explanation different from interface-based explanation.

If we look carefully at how the appeal to optimal computation is justified in the minimalist literature on the one hand, and how it is supposed to lead to a principled explanation of the language faculty on the other, we are forced to see a tension between the two in terms of the explanatory status of optimal computation. Let us consider how this tension arises.

If language is regarded as epitomizing an "optimal design" on the grounds that its principles of computational efficiency follow from elegant, economical, and simple laws of physics, then a minimalist explanation referring to optimal computation must exhibit explanatory power on its own, independent of that of an explanation based on legibility conditions. Observe that this is what is required if talk about *two* sources for principled explanation is to be meaningful

(recall the quotation from Chomsky (2004b: 158), cited on p. 109). However, when we look at actual practice, not only do we see that the two modes of explanation are closely related, but we also see that explanations referring to optimal computation are often subsumed under interface-based explanations. Surely, in the event of interaction of the two modes, we would have expected the opposite, given the widely accepted view that physics provides the most fundamental level of scientific explanation.

Before taking this line of discussion further, let us consider a few examples, beginning with the following passage from Chomsky and Lasnik (1993, reprinted in Chomsky 1995a: 28):

> The principle of economy of derivation requires that computational operations must be driven by some condition on representations, as a "last resort" to overcome a failure to meet such a condition. Interacting with other principles of UG, such economy principles have wide-ranging effects and may, when matters are properly understood, subsume much of what appears to be the specific character of particular principles.

It is interesting to observe that while the authors allude to the possible subsuming of UG principles under more general principles of economy, they seem to overlook the fact that their description of the principle of economy of derivation suggests the subordination of this economy principle to legibility conditions. Now the question is: Why should the principle of economy of derivation, which is supposed to be somehow a consequence of some physical law, be operative in such a way as to counter the danger of some condition on representations being trampled on? To put it in provocative terms: Why should the laws of nature care whether a derivation in a language satisfies legibility conditions? It would certainly be absurd to say, for instance, that Boyle's law, relating pressure and volume, requires that the dynamics of the volume of the heart's chambers must be driven by some condition on adequate blood pressure and circulation, as a last resort to overcome a failure to meet such a condition. If physical laws are indifferent to whether the heart beats or ceases to beat, why should they care about whether language is usable or not?

Now, Chomsky (1995a: 220) assumes that full interpretation (FI) determines the subset of *convergent* derivations out of the set of all derivations, and he further assumes that economy principles apply only to convergent derivations to determine the subset of *admissible* derivations. Thus, it might be argued that optimal computation does after all seem to be indifferent to convergence requirements, and it also seems to have the last word in deciding on the design

of language. However, we can see that this is not the case from inspection of the line of reasoning which led Chomsky to adopt the above assumptions.

Very briefly, Chomsky (1995a: 220) suggests that "[l]ess economical computations are blocked even if they converge." The reason for this, as he sees it, is that a linguistic expression cannot be fully defined as a convergent derivation; it must also "be *optimal*, satisfying certain natural economy conditions" (1995a). So far, this indicates that optimal computation has an independent explanatory status. But Chomsky (1995a: 220–1) goes on to argue that the most economical derivation is sure to crash since it applies no operations at all, and if crashed derivations can block others, it follows that the most economical derivation will block all other derivations, a conclusion which he describes as "an unwelcome result." To overcome this problem Chomsky (1995a) adopts the proposal of the previous paragraph that crashed derivations do not block others, and that economy considerations hold only among convergent derivations.

Now, Chomsky's proposal may solve the problem he raises, but, crucially, it indicates that recourse to optimal computation is not fundamental, but is tempered by a requirement to take account of legibility conditions. If this is so, then considerations of optimal computation can in principle be dispensed with in favor of considerations of legibility conditions. It is perhaps worth mentioning that this is precisely what Chomsky himself has done in at least one case. Thus, in discussing the status of the minimal link condition (MLC), which requires movement to be the shortest possible, Chomsky (1995a: 267–8) proposes – as a way to address the problem of excessive computational complexity which such a condition can lead to – to deprive this condition of its status as an economy condition and conceive of it instead as "part of the definition of Move." This proposal is characterized by Chomsky as a preferred one, since in this case the difficult question of how to compare derivations "does not arise," as "violation of the MLC is not a legitimate move in the first place" (1995a: 268). I take it that this entails that the MLC is a *legibility*, rather than an economy, condition, in the sense that its violation results in a crashed derivation. Now, since Chomsky's suggestion to reduce the MLC to the status of a legibility condition is motivated by an attempt to constrain the "globality" of economy of derivation, it may be argued that the conclusion which I drew above (*viz*. that considerations of optimal computation can in principle be dispensed with in favor of considerations of legibility conditions) misses the point. To this my answer is simple: globality of economy of derivation is a *problem* because it suggests that language is computationally intractable and, therefore, is *not usable*. I take it that "usability" is a notion more closely connected with the interfaces and their legibility conditions.

The important point I am trying to make here is perhaps best clarified by asking ourselves the following question: why should computations be minimized? One possible answer is: in order to enhance the ability of the external systems to access the information provided by the computational system – and here I have in mind in particular the notion of "active memory" in phase theory (see Chomsky 2001). A different answer is that computations are minimized as a consequence of the contribution of simple (optimal, economical) physical laws to the design of language. Clearly, the two answers are distinct; the former highlights the need to satisfy legibility conditions in an "efficient" way, and the latter takes a "naturalist" perspective. Now, to say that economy considerations are driven by some condition on representations is simply not consistent with the proposition that optimal computation constitutes a mode of explanation *sui generis*.

Of course, it might be argued that there is no reason for concern here; language (*qua* cognitive system) has a fundamental task to perform, namely satisfying legibility conditions, and (*qua* biological system) it exploits the laws of physics to accomplish its task. Let us grant this. But then it must also be admitted that, from a minimalist perspective, language (*qua* natural object) is *subject to these same laws*, and it is precisely because of this that we should not adopt an *à la carte* attitude towards the notion of optimal computation with economy considerations being welcome *insofar as they contribute to the usability of language*. This attitude is clearly manifested in Hornstein *et al*. (2005: 324), who, upon realizing that the most economical option available at a certain point of a derivation would lead to a crashed structure, are led to assert "that less economical operations are permitted *if* the more economical options don't lead to a convergent result" (emphasis in original). But this assumption seems absurd. For if we adopt the view that economy principles are inherent to language in the *substantive* sense, that is, in the sense of being a consequence of the *necessity* of physical laws, then it is incoherent to imagine a computational system guided by an economy principle of derivation in such a way that sometimes, but not always, the derivation is most economical.

In short, such a "pick-and-choose" attitude is not only at odds with the minimalist conception that optimal computation derives its explanatory power from the necessity of physical laws, but it also undermines it by suggesting that economy considerations are contingent on demands coming from the interfaces. Consider, as yet another example, Uriagereka's (2000, 2001, 2002) so-called "entropy condition," which suggests to him a correlation between the notion of "entropy" in thermodynamics and the notion of computational economy in syntactic theory. As he defines it, this is a condition which states that, at a certain

derivational juncture, "derivations take those steps which maximize further convergent derivational options." We shall have occasion later to comment on this condition, but here suffice it to say that, as stated, Uriagereka's condition clearly puts computational optimality at the service of the notion of convergence. In this connection, it is interesting that Uriagereka (2001: 897) offers the following perspective on the SMT: "once you give up a functionalist approach to optimality in grammars, what else could it be other than the familiarly optimal universe (in physics and chemistry) manifesting itself in biology?" This may reflect his dissatisfaction with the standard minimalist conception of the SMT – a dissatisfaction which, as mentioned in p. 117, has led him to propose his co-linearity thesis. In this sense, Uriagereka's conception of the SMT could be interpreted as implicitly recognizing the tension that economy principles like his own are liable to create between viewing optimal computation as needing to take account of legibility conditions and as a consequence of physical laws.

Fukui (1996), whose views we will consider further in the next section, makes an attempt to keep these two views of optimal computation separate, but (I think) with little success. For instance, he maintains that, unlike the condition on economy of representation, the condition on economy of derivation is not related to legibility considerations, adding that the latter condition, which he considers to be akin to a physical law, is "computationally intractable," and that this indicates that language is "fundamentally unusable" (Fukui 1996: 64–6). However, in order to explain the fact that language *is* used, he falls back on legibility considerations under the label "computational tricks," tricks which are "embedded in economy of derivation" and which "have the function of facilitating usability of language."

But how is optimal computation supposed to be an instance or a consequence of a physical law? What this question is really asking for is independent evidence for the role of optimal computation as an *explanans* in minimalist explanation. If optimal computation, as linked to physical laws, is to be genuinely explanatory, then it is necessary to explore this link and have the exploration yield positive outcomes. Otherwise, one is left with the notion of optimal computation as a *primitive* that is not grounded in physical principles (see Chapter 6). Now, some minimalists have sought to justify the view that optimal computation can be properly grounded in general physical principles. Their efforts have resulted in various strategies aimed at establishing a connection between the principles of language and the laws of physics. These strategies in turn, however, have resulted in nothing more than loose correlations, as I now seek to demonstrate.

5.6 Loose correlations

The strategies to which I have just referred have certain common features: for example, they contain various citations from the works of prominent scientists, especially physicists of great reputation; they seek to establish a comparison between the basic tenets of the minimalist program and some empirical findings or general practices in the core sciences; and often end with the suggestion that there are genuine analogies between the two sides of the comparison (see, among others, Fukui 1996; Uriagereka 1998; Freidin and Vergnaud 2001; Epstein and Seely 2002; Boeckx and Piattelli-Palmarini 2005; Boeckx 2006; Ott 2007; Boeckx and Hornstein 2010). These strategies may be classified into two types, according to whether they are oriented towards examining the *methodological* or *substantive* aspects of minimalism. When methodology is the target of comparison, the conclusions reached are sometimes made in such a way as to imply – at least implicitly – that linguistics is almost theoretical physics in disguise. When substantive links between minimalism and the physical sciences are sought, the conclusions arrived at are usually made with optimism, to the effect that only a future physics of the brain will determine whether or not the connection between the principles of language and the laws of physics are merely metaphorical. It is these strategies of this latter type that are the primary focus here.

For reasons of space I shall limit myself to considering two examples by means of which I look to demonstrate that the strategies in question have resulted in nothing more than the postulation of erroneous and vague analogies between the principles of language and the laws of physics. The two examples are Fukui (1996) and Uriagereka (2000, 2002), and their selection is not arbitrary. The former is one of the earliest and most influential (at least in the minimalist literature) attempts to ground the minimalist notion of economy in a physical basis, and the latter comes from someone who is well-known for his strong attachment to such curiosities – so strong, in fact, that he published a sizeable monograph about them (Uriagereka 1998, and for criticism, see Levine 2002). Moreover, both of these examples illustrate very clearly how venturing beyond familiar territory can open a Pandora's box of misconceptions.[5]

Fukui (1996: 51) argues for the existence of "rather unexpected fundamental connections … between the principles of language and the laws governing the inorganic world." He claims that these connections represent "a concrete interpretation of Chomsky's suggestion that language appears to show the kind of [economy] property that we expect in the core areas of the natural sciences" (Fukui 1996: 65). In particular, he claims that the principle of

economy of derivation "is exactly the linguistic version of [Hamilton's] Principle of Least Action in physics" (Fukui 1996: 67). Let us examine the validity of this claim.

To begin with, Fukui concedes that there is an important difference between Hamilton's principle and the principle of economy of derivation, namely that the former "deals with continua, whereas [the latter] is a property of a discrete system" (Fukui 1996: 56). Yet he plays down this difference by drawing attention to what he regards as remarkable similarities between the two principles. He mentions two: economy and globality. We consider these in turn.

Fukui contrasts the two principles in question and asserts that the similarity between them is "obvious; they both have the effect of minimizing the value of a function" (1996: 61). In other words, just as the principle of *least action* requires an action integral to be minimal in value, the principle of economy of derivation requires a syntactic derivation to be minimal in cost. The problem with this analogy is that it relies on an erroneous definition of Hamilton's principle. Fukui (1996: 55) defines this principle as "stat[ing] that the action integral of the difference between the kinetic energy of an object and its potential energy over the interval of time during which the motion takes place must be a minimum for the path actually chosen by nature." Observe that this definition states that the value of this function *must be a minimum*, which, if true, would lend support to Fukui's analogy and allow him to claim that "considerations of economy in physics," like Hamilton's principle, "offer a number of quite interesting implications for the design of language, if language indeed exhibits the property of economy" (1996: 52). Unfortunately, for Fukui, the matter is not so straightforward.

We need to note two points. First, as the passage from Hamilton quoted below indicates, the value of the action integral *need not* be a minimum as Fukui would have us believe. Second, as we shall see in Section 5.7, the notion of "least action" did not originate with Hamilton, but was dear to the eighteenth-century mathematician and philosopher Pierre-Louis Maupertuis, who invested it with a mystical aura. Hamilton himself refused to attach any metaphysical significance to this notion and rejected its association with the notion of "economy in the universe." He also expressed his dissatisfaction with the adjective "least," proposing instead the term "stationary." Thus, after tracking the history of minimum principles in optics and mechanics, Hamilton writes:

> But although the law of least action has thus attained a rank among the highest theorems of physics, yet its pretensions to a cosmological

> necessity, on the ground of economy in the universe, are now generally rejected. And the rejection appears just, for this, among other reasons, that the quantity pretended to be economised is in fact often lavishly expended ... In mathematical language, the integral called action, instead of being always a minimum, is often a maximum; and often it is neither the one nor the other: though it has always a certain *stationary* property ... We cannot, therefore, suppose the economy of this quantity to have been designed in the divine idea of the universe: though a simplicity of some high kind may be believed to be included in that idea. And though we may retain the name of *action* to denote the stationary integral to which it has become appropriated – which we may do without adopting either the metaphysical or (in optics) the physical opinions that first suggested the name – yet we ought not (I think) to retain the epithet *least*: but rather to adopt the alteration proposed above, and to speak, in mechanics and in optics, of the *Law of Stationary Action*. (Hamilton 1967 [1833]: 317–18)

In the light of this passage, it should be clear that the Hamilton on which Fukui relies does not exist, and, therefore, the analogy which he seeks to establish between the principles of language and the laws of physics is unfounded. We shall return later to consider the mystical implications that such an unfortunate analogy would have for the minimalist explanatory framework (Section 5.7). But now, we turn to a second similarity which Fukui alleges to exist between Hamilton's principle and the principle of economy of derivation.

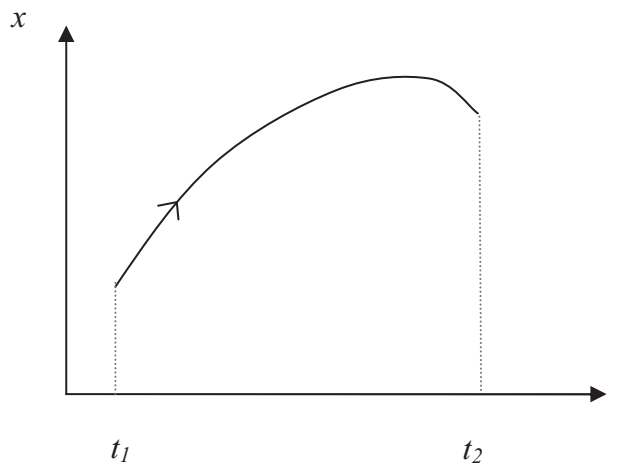
Fukui (1996: 61) claims that "both principles require some form of 'globality.'" He indirectly defines the term "global" in the case of language as follows: "We say a condition C is *local* if we can determine whether C is fulfilled or not by inspecting a single Phrase-marker; otherwise it is *global*" (1996: 61, underscoring in original). On the basis of this definition, he describes the principle of economy of derivation as a global condition, because in order "to determine whether it is satisfied or not, we have to inspect more than one Phrase-marker or perhaps even more than one derivation" (1996: 62). Fukui's views on this issue, relying on the framework of Chomsky (1995a) in which calculations of derivational economy have this character, may be outdated, but this is immaterial to the present discussion. The important point is to see just how he brings Hamilton's principle to bear on the principle of economy of derivation. Consider what he has to say:

> [T]he economy principle in physics is fundamentally "global" in nature ... For example, in order to apply Hamilton's Principle

> to ... the motion of an object starting at time t_1 and ending at time t_2, we must know the initial condition of the motion at t_1 and the final condition of the motion at t_2 ... In fact, we can say that while Newtonian mechanics approaches physical phenomena in a "local" fashion in terms of differentiation, Hamiltonian (*sic*) approach provides us with a "global" alternative for the description of physical phenomena, in terms of "action integrals." This point will become important when we consider the "global" nature of economy of derivation in language. (Fukui 1996: 55)

It is not clear why the fact that Hamilton's principle is expressed by an integral equation should have anything to do with economy of derivation, although the term "global" provides a hint as to what Fukui might have in mind. Let us try then to make the supposed analogy, which he fails to fully express, as clear as possible. We already know what it means for a condition on derivations to be global, but it remains to be specified what it means for an integral equation to be "fundamentally global in nature." The following remark from a physicist may help to clarify this point: "Particles move according to local, not global instructions; their equations of motion are *differential* equations, even though the functional they come from is a definite integral, a *global* description" (Neuenschwander 2010: 199, italics in original). We have here a correlation between a local/global distinction on the one hand, and differential/integral equations on the other, and that is all we need to understand why Fukui refers to Newtonian and Hamiltonian approaches as being local and global, respectively. Now, Fukui says that the application of Hamilton's equation to an object in motion requires prior knowledge of the initial and final conditions of its motion at t_1 and t_2, respectively. To make this more intelligible, consider a simple example from Feynman *et al*. (1964: 19):

> Suppose you have a particle ... which starts somewhere and moves to some other point by free motion – you throw it, and it goes up and comes down ... if the particle has the path $x(t)$ (let's just take one dimension [for simplicity]), where x is the height above the ground, the kinetic energy is $\frac{1}{2}\, m\, (dx/dt)^2$, and the potential energy at any time is mgx. Now I take the kinetic energy minus the potential energy ... and integrate that with respect to time from the initial time to the final time ... The actual motion is some kind of a curve – it's a parabola if we plot against the time – and gives a certain value for the integral.



(Feynman *et al*. 1964: 19-1&2)

This example illustrates what it means to apply Hamilton's integral equation to a moving object, and in the light of it we now begin to understand what Fukui means by claiming that some sort of globality is required for both Hamilton's principle and the principle of economy of derivation. Just as we cannot determine whether the action integral is a minimum by inspecting a local subsection of the path $x(t)$ (see the graph above), we cannot determine whether the computational cost is a minimum by inspecting a local domain of a derivation. The analogy is flawed, however; it turns out that, *pace* Fukui, we *can* actually determine the integral of the action by inspecting a subsection of the path of the motion, no matter how infinitesimal this subsection is. As Feynman *et al*. (1964: 19–8) put it, "if the entire integral from t_1 to t_2 is a minimum, it is also necessary that the integral along . . . every subsection of the path must also be a minimum." The authors explain that, in the latter case, if the subsection is small enough, the minimum action can be obtained by a simple *differential* equation. From this they (1964) conclude: "So the statement about the gross property of the whole path becomes a statement of what happens for a short section of the path – a differential statement." It should be clear by now that Fukui's claim that Hamilton's principle is "fundamentally global in nature" is manifestly incorrect, and, therefore, his "global" analogy is critically flawed.

Supposing for the sake of argument that the analogy can be sustained; the question now arises of how, despite the globality involved, an object in motion

chooses precisely the minimum path out of an infinite number of alternative paths between two points, and how, in the case of language, the computational system calculates the derivation with minimum computational cost out of many alternative derivations. The former question, as we shall see in the next section, has brought mysticism into physics, but Fukui seems to be unconcerned about this, for his survey of the relevant history is completely silent on this issue. As to the latter, Fukui (1996: 68) conjectures that, from the point of view of the theory of computational complexity, "economy of derivation ... is fundamentally computationally intractable." He sees his conjecture as a vindication of Chomsky's claim (1991) that language is designed for "elegance," not for use (1991: 66). But language is used after all, and to explain this fact Fukui, as already pointed out in the previous section, suggests that language is used thanks to "heuristic algorithms" or "computational tricks" embedded in economy of derivation, such as the principles of Greed and Procrastinate (see Section 2.5.3 for a description of these principles). But the conceptual and computational problems associated with these principles are well known, especially in connection with the teleological notion of "look-ahead" (see, in particular, Chomsky 2000b). For instance, in order to delay movement operations until after Spell-Out (i.e. Procrastinate), the computational system has to look ahead to see whether the delay is justified or not. Consequently, Fukui's analogy, even if it were plausible, leads to two negative consequences: an inflation of the complexity of the computational system, and a teleological explanation of the fact that language is used.

Turning now to our second example, let us consider the "entropy condition" which we have met earlier. Uriagereka (2000: 869) maintains that this condition "is comparable to the Second Law of Thermodynamics." However, this comparison rests on a misunderstanding of an established physical law. To see this, we first need to say something about this law.

The Second Law of Thermodynamics dictates that "a closed system will tend toward maximum entropy" (Chabay and Sherwood 2002: 484). A closed system is one in which no energy flows across the system boundary, such as in the case of a closed container filled with gas atoms colliding with each other and with the inner walls of the container. Entropy is a technical term one of its definitions being "the number of microstates corresponding to a particular macrostate of specified energy," that is, "the number of ways to arrange energy among a group of atoms" (Chabay and Sherwood 2002: 483). To see the Second Law at work, consider a simple and perhaps familiar example. A sugar cube is in a highly ordered state in which the number of ways to arrange the energy among sugar molecules is very small. However, when the sugar cube is placed in a

glass of hot water, the number of accessible microstates increases and, therefore, the number of atomic configurations also increases. In this case, the Second Law predicts that the state of the system will tend toward maximum disorder (i.e. high entropy) as the sugar molecules disperse in the solution – the sugar cube dissolves. The Second Law is much more complicated than this, but for our purposes, it is important to realize that the law *applies only to closed systems of atoms and molecules with a natural tendency of spontaneous change* toward maximum thermodynamic entropy. What makes this tendency spontaneous is the fact that the motions and collisions of atoms are fuelled by their internal kinetic energy.

Returning now to the supposed analogy between the "entropy condition" on syntactic derivations and the Second Law, Uriagereka's justification for it is that derivational paths may be regarded as "micro-states" and, therefore, one can speak of "the idea of the *entropy* of a derivation" (Uriagereka 2002: 29). Clearly, this sort of justification does not seem to carry us beyond the superficial level of terminology, but let us see how far the analogy can go. Consider what he has to say:

> A derivational decision d will allow further convergent steps in a number of n possibilities, whereas some other derivational decision d' will only allow a number m of possible continuations, where $m < n$. In those circumstances d induces more derivational entropy than d' does, which [the entropy condition] aims at optimizing. (Uriagereka 2002: 29)

Three observations on this passage are in order. First, Uriagereka blatantly ignores what physicists consider to be common knowledge in their field. By asserting that his entropy condition aims at *optimizing* the derivational decision d, he is assuming that *via* the analogy variational minima-maxima principles are applicable to thermodynamics (on these principles, see Section 5.7). However, physicists consider it to be "commonly agreed that thermodynamics is a branch of physics which is not adaptable to the technique of variational principles" (Yourgrau and Mandelstam 1960: 93).

Second, we have stressed that the Second Law applies only to closed systems in which atoms and molecules have a tendency of spontaneous change toward thermodynamic equilibrium (i.e. maximum entropy) and we have noted that this is a consequence of inherent kinetic energy. Now, since no one is in a position to demonstrate that the physical system underlying syntactic computations constitutes a closed system displaying continuous thermally induced motion, Uriagereka's claim does not go further than the metaphorical level. On at least one occasion, he seems to concede this, for he says that his entropy condition

"can be metaphorically linked with the idea that costly derivations are dispreferred" (Uriagereka 2002: 33). If metaphors are what he seeks, there is perhaps little to dispute, but when vague metaphors and analogies are offered in the guise of profundities, there is more cause for disquiet.

Finally, we are told that the entropy condition aims at optimizing derivational decisions that result in more convergent derivations. Thus, given two derivational decisions d and d', the former outranks the latter if it leads to more convergent derivations. Whether this condition actually enhances efficient computation need not concern us here, but we mention in passing that Lappin *et al*. (2000, 2001) argue that the entropy condition has the opposite effect of what Uriagereka seeks to accomplish, calling it "an anti-economy condition" (Lappin *et al*. 2001: 910). Uriagereka's (2001: 895) unconvincing reply is that he did not intend his entropy condition to serve as an economy condition, but rather as a condition which has indirect economy consequences; having to decide between d and d' has the effect of reducing "the class of possible derivations." What should be mentioned here, however, is that Uriagereka's entropy condition induces the teleological notion of "look-ahead"; at each point of a given derivation, the condition requires the system to look ahead in order to choose between d and d'. Besides the computational complexity it induces, the entropy condition thus clearly requires commitment to a teleological mechanism. This teleology is expected, however; since Uriagereka's condition reduces what is supposed to be an explanation based on optimal computation to an interface-based explanation (as observed in the previous section), and since the latter has a teleological character (as we argued in Section 5.4), it is only natural that the former explanation should inherit such a teleological character. We shall now see how teleology in a new guise infects the sort of argumentation we are concerned with here.

5.7 Minimalism and teleological physics

So far we have discussed some examples of the minimalist effort to ground optimal computation on physical principles, concluding that they have been abortive. In this section we shall be concerned with the status and nature of the physics to which the notion of optimal computation might be linked. Before proceeding to our main task, let us first have a brief look at how some minimalists view the supposed connection between minimalism and physics.

Fukui (1996) offers a brief survey of what he terms "economy principles in physics," referring to minimal principles such as Fermat's principle of *least time* in optics, and Maupertuis's principle of *least action* in mechanics. The former

states that, among all possible paths between two points, light "chooses" the one that requires the least time. It was proposed by Pierre de Fermat in the seventeenth century, and was later subsumed by quantum electrodynamics. As to Maupertuis's principle, it simply generalizes Fermat's principle by replacing the notion of time by the broader notion of action, the latter being understood as the product of three physical quantities (mass, velocity, and distance). As observed earlier, Fukui claims that there are fundamental connections between physical principles like these and economy principles in language.

Fukui is not alone in his enthusiasms. Uriagereka (1998) reiterates the supposed analogy between economy in language and the principle of least action in physics, describing it, through the voice of his character the Linguist, as "a nice analogy," one which indicates that "just as this principle describes a mechanical path and, say, electricity, a deeper version of it may also describe a successful linguistic computation" (Uriagereka 1998: 84). Freidin and Vergnaud (2001), following in the footsteps of Fukui, also appeal to the principles of least time and least action in arguing for a substantive link between minimalism and physics, claiming that "economy considerations contribute substantially to what constitutes the 'perfection' of the computational system in both domains" (Freidin and Vergnaud 2001: 652).[6]

In order to evaluate these claims, it is important that we have a historical perspective on minimal principles in physics. This is not the place for a detailed account of the history of these principles and their philosophical roots. Our exposition will therefore be brief and confined, in the most part, to what is relevant for our purposes. For a more detailed exposition, the reader is referred to, for example, Yourgrau and Mandelstam (1960) and Dugas (1988). The present exposition relies principally on the former work.

We begin with the notion of "simplicity." As a scientific ideal, simplicity has its origins in the millennia-old effort of philosophers and scientists to reduce the apparent complexity of the observable world to a minimum of principles or substances. Thus, the pre-Socratics Thales and Heraclitus identified the origin of all being as water and fire, respectively. Empedocles, by contrast, postulated four elements: water, fire, earth, and air. In sharp contrast to this conception of the ultimate basis of the universe in terms of material substances and chemical elements, the Pythagorean legacy had it that number ruled the universe.[7] So obsessed were the Pythagoreans with an ideal principle of form underlying natural phenomena that they coined the slogan "all is number," which so characterized their number mysticism. Owing to this preoccupation with the concept of number, their cosmology was loaded with metaphysical ideals such as simplicity, beauty, harmony, symmetry, perfection, and so on. But it is well to

remember that, for the Pythagoreans, these ideals were valued, not simply on account of some embodied aesthetic or pragmatic character, but because of an underlying presumption that they have an epistemic value, as they were believed to contribute to man's knowledge of the natural world. Such presumption has exerted a powerful influence on natural philosophers from the earliest times and, as we shall see later, was only finally overcome by the development of modern physics.

The influence of simplicity and its relations manifested itself in various ways. One is in the work of Aristotle, to which we come shortly. Another is the scholastic doctrine, according to which *simplex sigillum veri* ("simplicity is the hallmark of truth"). Among the fathers of modern science, the Pythagorean influence can be seen in the works of Copernicus, Galileo, Kepler, and Newton. Thus, Copernicus asserted that "*Natura simplicitatem amat*," and Burtt (1924: 46) observes that "[w]ith him nature's simplicity and unity was a commonplace." Galileo wrote in his dedication of the *Dialogue* to the grand duke of Tuscany that "turning over the great book of nature ... is the way to elevate one's gaze," and "that book is the creation of the omnipotent Craftsman, and is accordingly excellently proportioned" (Galileo 1967: 3). Kepler went so far as to write a poem in which he "presented his vision of a world created from number in which Copernicus was the restorer of Pythagorean truth" (Walton and Walton 1997: 48).[8] As for Newton, in the *Principia* he justified one of the so-called "rules of the study of natural philosophy" by saying: "Nature does nothing in vain, and more causes are in vain when fewer suffice. For nature is simple and does not indulge in the luxury of superfluous causes" (Newton 1999: 794).

It is on the basis of this metaphysical foundation that one has to understand the import of minimum principles in physics, for they are nothing but attempted instantiations of venerable metaphysical views. As the history of these principles demonstrates, it is only fairly recently that physics has begun to free itself from such metaphysical issues. Let us trace this history briefly.

The foundation on which the notion of "minimum principle" rests is to be found in Aristotle's physics, for here we find two postulates which we shall see in operation in all minimum principles, and which are intimately related as far as these principles are concerned. The first is that of *teleology*. The Aristotelian teleology, as is well known, is expressed by the technical term "final cause," which suggests that the relation of means to end is operative in nature. Thus, Aristotle concludes a discussion of the question of whether nature acts for an end with: "It is plain then that nature is a cause, a cause that operates for a purpose" (Arist. *Phys*. 2.8.199b32–33, in McKeon 1941). The second postulate

involves *simplicity*.[9] In *De caelo* Aristotle argues that all motions are either straight or circular, because "these two, the straight and circular line, are the only simple magnitudes" (Arist. *Cael*. 1.1.268b19–20, in McKeon 1941). Following this same simplicity criterion, he goes on to provide an explanation for the apparent circular motion of the planets in terms of a *minimum* hypothesis:

> [I]f the motion of the heaven is the measure of all movements whatever in virtue of being alone continuous and regular and eternal, and if, in each kind, the measure is the minimum, and the minimum movement is the swiftest, then, clearly, the movement of the heaven must be the swiftest of all movements. Now of lines which return upon themselves the line which bounds the circle is the shortest; and that movement is the swiftest which follows the shortest line. Therefore, if the heaven moves in a circle and moves more swiftly than anything else, it must necessarily be spherical. (Arist. *Cael*. 2.4.287a23–31, quoted in Yourgrau and Mandelstam 1960: 4)

As Yourgrau and Mandelstam (1960: 5) note, "Aristotle's minimum hypothesis ... was clearly not dictated by an appeal to quantitative measurement and was not subject to rigorous scrutiny." However, in the first century AD, Hero of Alexandria proposed what might be called the principle of *least distance*. He sought in this principle an explanation of the optical law according to which the angle of incidence and the angle of reflection are equal. This law of reflection was well known in Hero's time, but what was not understood was why the law should hold. Hero proposed that the behaviour of a ray of light reflected from a mirror could be explained by a minimum principle, namely that light traverses the *shortest* of all possible paths between one point (the light source) and another (the light receptor). In comparing Hero's principle with Aristotle's minimum hypothesis, Yourgrau and Mandelstam (1960) observe:

> Although Hero differed from Aristotle by demonstrating mathematically that his principle was in agreement with experimental data, he considered this principle to provide an "explanation" of these data. His approach was therefore akin to Aristotle's in that he deduced his results from preconceived suppositions.

Indeed, the practice of deducing certain empirical results from "preconceived suppositions" concerning the behaviour of the natural world characterizes all minimum principles in the history of physics. Thus, sixteen centuries after Hero had proposed his principle of least distance, the French mathematician Pierre de

Fermat extended Hero's principle to explain both the laws of reflection and refraction. As mentioned in the previous section, Fermat's principle assumes that, among all possible paths between two fixed points, a ray of light takes the path that requires the *least time* to traverse. A century later, Pierre-Louis Maupertuis generalized Fermat's principle by referring not to the notion of time but to a quantity he termed "action," which he believed could be expressed mathematically as the product of mass, velocity, and distance.

As already observed, a common feature of all these minimal principles is their *teleological* character. Their teleology inheres in the fact that they imply that the minimization of a certain quantity in a physical system is the *goal* to which the behavior of the system is sensitive. Put differently, the behavior of a physical system passing from one configuration to another is driven by the *purpose* of minimizing a certain quantity in the system (e.g. distance, time, action). What this really entails is that the *future* or *final* state of the system is essential for explaining its behavior.

Now, it is evident that the appeal to teleology was justified by an *a priori* maxim of simplicity as an inherent property of nature. For instance, if one were to ask why nature should behave in such a way as to minimize a certain quantity, we would be given a Copernican answer, namely "*Natura simplicitatem amat*." It is through this intertwined relationship between teleology and simplicity that minimum principles have contaminated physics with theological and mystical ideas. The principles of Fermat and Maupertuis are two examples of this. In the former, the mysticism manifests itself through light's ability to pick the quickest path between two points, an ability which requires information on all other possible paths, we might suppose; it is as if light were behaving intelligently when it "chooses" the path that requires the least time to arrive at its final destination. It comes as no surprise therefore to learn that Fermat's principle provoked outrage among Cartesians, champions of mechanical (as opposed to teleological) explanations of nature. Claude Clerselier, Descartes's friend and editor of his work, sent a letter (dated May 6, 1662) to Fermat, in which he argued that nature "acts without foreknowledge, without choice, and by a necessary determination." Two weeks later, Fermat wrote back to say:

> I have often said . . . that I do not claim and that I have never claimed, to be in the private confidence of Nature. She has obscure and hidden ways that I have never had the initiative to penetrate; I have merely offered her a small geometrical assistance in the matter of refraction, supposing that she has need of it. (quoted in Dugas 1988: 259)

The "obscure and hidden ways" of nature to which Fermat refers were later revealed by classical electrodynamics in its appeal to the wave nature of light. But it was not until the rise of quantum electrodynamics that a full explanation of Fermat's principle emerged. Richard Feynman, one of the main authors of the latter theory, showed how many of the phenomena associated with the behavior of light can be explained by a simple method of adding "arrows," where each arrow represents a possible path for light between two points. It would be well beyond the scope of this discussion to give a detailed account of Feynman's approach.[10] For our purposes, suffice it to say that Feynman describes the principle of least time as only a "crude picture of the world," and shows that a more sophisticated analysis reveals that "where the time is least is also where the time for the nearby paths is nearly the same" (Feynman 1985: 45). More specifically, he shows that light in fact traverses all possible paths, where each path is associated with a so-called probability amplitude represented as the length of a vector or an "arrow" (as Feynman calls it). By simple vector addition of arrows (or amplitudes), Feynman demonstrates how the contributions of all the arrows cancel each other out except for those arrows that are near the centre of the mirror, which "also happens to be where the total time is least" (1985: 43).

Even more important for our present purposes, modern physics tells us another fact about light which contradicts what Fermat proposed: it may sometimes travel along the path with a *maximum*, rather than minimum, travel time (cf. the quotation from Hamilton in the previous section). As Raj (1996: 161) puts it, "a number of cases are known in which the real path of light is the one for which the time taken is maximum rather than minimum."[11] This, according to Mirowski (1989: 21), should not come as a surprise once we realize that extrema (i.e. minima *or* maxima) principles are special cases of more general ones, namely the so-called variational principles.[12] While the technical details of the calculus of variations are complex and well beyond the scope of this discussion, the crucial point should be clear: minimal principles in physics do not represent anything fundamental that governs the behavior of the natural world.

Turning to Maupertuis, he firmly believed in what he called the "Economy of Nature," which he saw as providing a "proof" for God's existence.[13] As Yourgrau and Mandelstam (1960: 20) put it, his main objective behind his principle of least action was "to furnish not merely a rational but also a theological foundation for mechanics," an objective which the authors describe as "a last vestige of medieval scholasticism with its imperative to reconcile faith and reason."

The principle of least action received further elaboration in the work of, among others, Euler, Lagrange, and Hamilton, the latter being the scientist with

whom the principle has been commonly associated (see previous section). These developments helped to undermine mystical and anthropomorphic interpretations of physical theory. For instance, in contrast to Maupertuis's metaphysical interpretation of the principle of least action, the physics of the nineteenth century expressed this principle in terms of differential equations, making the appeal to apriorism in stating least principles superfluous (cf. Yourgrau and Mandelstam 1960: 174).

In light of the preceding exposition, one cannot fail to see how biased and partial some minimalists are when they describe the history of minimal principles in physics, Fukui (1996) providing a perfect example. Thus, in his brief review of the relevant history, he fails to mention the mystical origins of minimal principles and is careful not to refer to the contempt with which many physicists regarded the metaphysical and theological conception of these principles. For instance, he refers to the work of the Hungarian mathematician and physicist Cornelius Lanczos, observing that it "contains a quite readable, yet accurate, discussion of the history of [variational principles]" (Fukui 1996: 54, n. 3). Yet, Lanczos's description of the fate of minimum principles, and which Fukui fails to mention, is as follows: "The sober, practical, matter-of-fact nineteenth century – which carries over into our day – suspected all speculative and interpretative tendencies as 'metaphysical' and limited its program to the pure description of natural events" (Lanczos 1970[1949]: xxvii). That Fukui fails to refer to the work of Yourgrau and Mandelstam (1960), regarded by many physicists as one of the major works on the subject, is also noteworthy.

Moreover, Fukui (1996: 53) affirms that "[b]y the early eighteenth century, there had been a few important attempts at elaborating on the description of the nature of economy in the physical world," and he goes on to cite as an example "Huygens' elaboration of Fermat's Principle of Least Time." In this way, the reader is led to assume that the Dutch physicist and astronomer Christian Huygens, Fermat's contemporary, was committed to the belief in a metaphysical basis for Fermat's principle. However, the truth is that Huygens "found no satisfaction in [Fermat's principle] and considered it was a 'pitiable axiom'" (Bell 1947: 58). The nineteenth century French mathematician and physicist Siméon Poisson expressed a similar view about Maupertuis' principle of least action, describing "it as only a useless rule" (Yourgrau and Mandelstam 1960: 32). Yet, when Fukui turns to Maupertuis' principle, he describes it as "the next important step after Fermat"s work," one which was later "refined further by ... Lagrange [and others]" (Fukui 1996: 53). Maupertuis' theological interpretation of his own principle, together with the actual view of

nineteenth-century physicists on the status of the principle, is simply ignored in Fukui's account.

Another example comes from the analogy which Fukui sought to draw between Hamilton's principle and the principle of economy of derivation, and which we have discussed in the previous section. Fukui (1996: 55) asserts that Hamilton's principle, despite the development of physics in the twentieth century, "stands as a basic principle for many branches of physics." If by this assertion it is meant that the principle continues to be empirically adequate, it is true. But if by the assertion it is meant that the old mystical interpretation of least principles has anything to do with Hamilton's principle, it is *false*. We have seen that such an interpretation was described by some eminent physicists as a crude picture of the behavior of the natural world, one which was seen to be inappropriate in view of later developments in physics. Yet, it is precisely this crude interpretation that is required for the claimed analogy between the principles of language and the laws of physics to be sound. This is clearly evident when Fukui says that

> the common feature of "economy principles" in physics can be summarized as follows: (i) find the relevant quantity Q; (ii) then, the principle is stated in the form "minimize Q," that is, in the form of a minimum principle. If the fundamental principle of language is shown to be stated essentially in this form . . . it is a rather surprising discovery which indicates a remarkable similarity between the inorganic world and language, a similarity that is by no means expected, given the biological nature of language. (Fukui 1996: 55)

What is truly surprising, however, is that Fukui should think that the fundamental laws of nature are stated in this form, that is, in terms of optimization or some related notion (cf. the SMT). Indeed, he does not seem to acknowledge the consequence of that which he proposes, namely a mystical, teleological interpretation of the natural world, in which nature is perceived as if it were a giant chess-playing machine, where all possible moves must be checked before the best move is chosen.

The overriding concern which is emerging from this discussion is the dangers that are associated with distorting the recent history of science. A further example is provided by Freidin and Vergnaud (2001: 650) when they maintain that economy principles "have a long standing legitimacy in the physical sciences." This claim is clearly incompatible with the views of *modern* physics, as has been shown above. It is equally alarming to note one of Uriagereka's (1998: 83) fictional characters saying with full confidence that "physics didn't

give up on the idea that a law of economy derives the substantive behavior of light." The truth of the matter is that most physicists never gave up this deeply metaphysical idea precisely because it had never been regarded as providing a fundamental explanation of the behavior of light, not even by Fermat himself (see his reply to Clerselier, quoted in p. 137).

Let us now recapitulate our discussion of optimal computation and its explanatory role in minimalism. We first considered the extent to which optimal computation might constitute a level of explanation independent of the interfaces, and we arrived at the conclusion that certainly some explanations based on optimal computation are subordinated to interface-based explanations and, therefore, do not enjoy explanatory autonomy. We next turned our attention to a consideration of some of the attempts which have been made to ground optimal computation in the natural sciences. After a careful examination of these attempts, we concluded that they comprise nothing more than vague analogies which furthermore reflect serious misconceptions of some scientific concepts. Besides being wrong, such attempts, I feel, do more harm to minimalism than good, often unnecessarily exposing it to contempt and ridicule. Lastly, we considered the status of the physics that is supposed to relate to the minimalist program. By placing minimum principles in physics – to which economy principles in language have been supposed to relate – in their historico-philosophical context, we were led to two important observations. First, the history of these physical principles is not adequately portrayed in the minimalist literature. Second, those minimalists who appeal to minimum principles do not seem to acknowledge the mystical implications that such an appeal would have for their explanatory framework.

The overall conclusion, then, is that there is currently very little on offer to justify the view that optimal computation can be properly grounded in general physical principles. In other words, given current speculations, there is little empirical support for a physical basis for aspects of optimal computation. Of course, one may argue that one should not exclude the possibility that a future physics of the brain might alter this situation. This is actually what some of the advocates of minimalism hope for. Uriagereka (2002: 33), for instance, believes that his speculation concerning the entropy condition "can be fully tested only when we learn more about the physical basis of language." This same attitude of "let's wait and see" is also expressed by Freidin and Vergnaud (2001: 652), who assert that the link provided by economy considerations between linguistics and physics "will have to be determined by a future neuroscience that can validate the physical approach to complex mental structures." However, there are arguments which, while not demonstrative, raise the possibility that investigation of

explanatory links between optimal computation and general physical principles may not be merely empirically unsatisfactory in the current state of knowledge, but may not in fact obtain for principled reasons. This is the topic of the next chapter, in which we approach optimal computation from a quite different perspective to that adopted so far.

6 *Optimal computation and multiple realization*

6.1 Introduction

This chapter expands the discussion of the explanatory status of the strong minimalist thesis (SMT) developed in the previous chapter by considering optimal computation in the context of the philosophy of mind. The chapter brings Chomsky's naturalism face-to-face with Fodorian functionalism and examines the tensions that arise between the two. The most significant one is that which emerges between the minimalist thesis that optimal computation can be grounded in physical laws, on the one hand, and the functionalist thesis that the mind as a computational device is to be approached as independent from the brain as a biological device, on the other. The main aim of this chapter is to discuss this tension and its implications for the explanatory role of optimal computation in particular, and for the status of the biolinguistic approach to language in general.

The chapter is organized as follows. Section 6.2 introduces Chomsky's naturalism and Section 6.3 illustrates its connections with the work of the computational neuroscientist Christopher Cherniak. In Section 6.4, a general introduction to the philosophical doctrine of functionalism is presented, followed in Section 6.5 by an exposition of the central argument of this position, namely the so-called "multiple realization argument." Section 6.6 contrasts Chomsky's naturalism with Fodor's functionalism and identifies a number of uneasy tensions between them. Next, in Section 6.7, Chomsky's criticism of functionalism will be discussed, with a special focus on his position on the mind-body problem. This will be followed in Section 6.8 by the main argument of the chapter, which is to show that the "grounding" of optimal computation in physical principles may be implausible *in principle*. The implications that follow from this for the biolinguistic approach, together with the explanatory status of optimal computation, will be discussed in Section 6.9.

6.2 Chomskyan naturalism

Chomsky's naturalism has two aspects: methodological and substantive. The former recognizes no significant distinction between the study of language and the study of any other "natural object," and the latter aspires to eventual unification between cognitive science and neuroscience. Let us consider these two aspects in more detail.

The naturalism advocated by Chomsky and his followers considers the human mind and its products (including language) as part of the natural world, where the term "mind" is understood as denoting the mental aspects of the world, and the term "mental" is placed on a par with such terms as "chemical," "electrical," "optical," etc. As we shall see later when we discuss Chomsky's views on the mind-body problem (Section 6.7), this conception of mind and mental phenomena derives, at least partly, from a particular interpretation of Newton's work in the context of the Cartesian mechanistic approach to physics. Here suffice it to say that Chomsky subscribes to the views of eighteenth-century thinkers such as La Mettrie and Priestley, according to which mental phenomena are properties of "organized matter." What this means is that "we can only assume that those phenomena 'termed mental' are the result of the 'organical structure' of the brain" (Chomsky in Cela-Conde and Marty 1998: 21).

Chomsky (2000a: 75) insists that he uses the term "mental" "without metaphysical import and with no suggestion that it would make any sense to try to identify the true criterion or mark of the mental." He goes on to say:

> Since the brain, or elements of it, are critically involved in linguistic and other mental phenomena, we may use the term "mind" – loosely but adequately – in speaking of the brain, viewed from a particular perspective developed in the course of inquiry into certain aspects of human nature and its manifestations. There are empirical assumptions here – that the brain, not the foot, is the relevant bodily organ, that humans are alike enough in language capacity so that human language can be regarded as a natural object, and so on. (Chomsky 2000a: 76)

Thus, this conception of mind, which is clearly neurologically based, subscribes to no metaphysical distinction between mind and brain, but only to the empirical assumption that the brain is the relevant bodily organ in the study of language and mind. Accordingly, Chomsky subscribes to *methodological naturalism* as opposed to methodological dualism, and argues that the "naturalistic approach" to human language and other mental phenomena should not be submitted to constraints that would not be acceptable in other domains of rational inquiry.

Such constraints are regarded by him as "a form of harassment of emerging disciplines [such as linguistics]" (Chomsky 2000a: 77). He summarizes them by saying:

> In the study of other aspects of the world, we are satisfied with "best theory" arguments, and there is no privileged category of evidence that provides criteria for theoretical constructions. In the study of language and mind, naturalistic theory does not suffice: we must seek "philosophical explanations," delimit inquiry in terms of some imposed criterion, require that theoretical posits be grounded in categories of evidence selected by the philosopher, and rely on notions such as "access in principle" that have no place in naturalistic inquiry. Whatever all this means, there is a demand beyond naturalism, a form of dualism that remains to be explained and justified. (Chomsky 2000a: 142)

In opposition to this methodological dualism, Chomsky has pursued an approach to language and mind which he considers similar to that of the natural sciences, with theoretical physics providing the preferred model. Already in his early work, Chomsky (1957: 49) alluded to an association between linguistics and physics in terms of theory construction, saying:

> Any scientific theory is based on a finite number of observations, and it seeks to relate the observed phenomena and to predict new phenomena by constructing general laws in terms of hypothetical constructs such as (in physics, for example) "mass" and "electron." Similarly, a grammar of English is based on a finite corpus of utterances (observations), and it will contain certain grammatical rules (laws) stated in terms of the particular phonemes, phrases, etc., of English (hypothetical constructs). These rules express structural relations among the sentences of the corpus and the indefinite number of sentences generated by the grammar beyond the corpus (predictions).

This somewhat superficial analogy becomes more sophisticated in Chomsky's subsequent works, especially in connection with the "Galilean style" as understood by major intellectual figures such as Husserl and Weinberg, figures who differ fundamentally in their interests and expertise in other respects.[1] The latter, for instance, conceives of this style as an attempt to construct "abstract mathematical models of the universe to which at least the physicists give a higher degree of reality than they accord the ordinary world of sensation" (Weinberg 1976: 28–9). Such an interpretation is regarded as plausible by Chomsky (1980a: 9), who adopts the Galilean style in Weinberg's

sense to the field of linguistics. The way in which this adoption is introduced and the methodological consequences it involves are illustrated in the following rhetorical question posed by Chomsky:

> Can we hope to move beyond superficiality by a readiness to undertake perhaps far-reaching idealization and to construct abstract models that are accorded more significance than the ordinary world of sensation, and correspondingly, by readiness to tolerate unexplained phenomena or even as yet unexplained counterevidence to theoretical constructions that have achieved a certain degree of explanatory depth in some limited domain, much as Galileo did not abandon his enterprise because he was unable to give a coherent explanation for the fact that objects do not fly off the earth's surface? (Chomsky 1980a: 9–10)

Thus, according to this passage, doing linguistics *à la* Galileo would involve, on the one hand, a substantial idealization of the object of inquiry, and, on the other, a primacy of theoretical constructs over empirical data. Chomsky is careful to stress that adoption of the Galilean style does not imply the disregard of recalcitrant data, although he adds that they "simply will not be considered very important for the moment" (Chomsky 1980a: 11–12). Data that are not yet explained by some consistent theory can still be described in whatever descriptive framework one chooses.

In addition to his commitment to methodological naturalism, Chomsky seems to subscribe to what might be called *substantive naturalism*. Indeed, he appears to press for connections between the study of language (and other cognitive systems) and the hard sciences that go beyond methodological considerations to substantive links. The issue for him is not merely that explanatory theories of mind should observe the canons of methodological naturalism, but it also involves his aspiration for the eventual integration of these theories and the "core" sciences, notably physics. Already in the mid-1980s, Chomsky expressed his belief that linguistics should sooner or later disappear as our understanding of the human brain improves:

> The study of language structure as currently practiced should eventually disappear as a discipline as new types of evidence become available, remaining distinct only insofar as its concern is a particular faculty of the mind, ultimately the brain: its initial state and its various attainable states. (Chomsky 1986: 37)

In his more recent writings, Chomsky envisages this integration as a form of unification between cognitive science and neuroscience, and not necessarily as

a reduction of the former to the latter (but see Section 6.8). Presumably, what this means is that the various theoretical terms of cognitive science will not need to be systematically mapped onto those of physical theories of the brain, as would be required by reduction; rather, the latter will be transformed in scope and significance in order to incorporate the former. This seems to be what Chomsky (2000a: 82) has in mind when he says that, in the history of science, it is often the case that "the more 'fundamental' science has ... to be revised, sometimes radically, for unification to proceed."

This optimism regarding the prospects of unification underlies the efforts of some minimalists to seek a connection between the principles of language and the laws of physics – efforts of which we have seen and discussed several examples in the previous chapter. But it should be remembered that such a prospective was on the agenda long before the advent of minimalism (cf. the passage from Chomsky 1978: 201, cited in p. 159). More recently, Chomsky seems to be seeking to instantiate his substantive naturalism by his positive attitude towards Cherniak's thesis of "non-genomic nativism," an attitude the purpose of which is to ground properties of language, notably the optimality of computations, in physical law. Let us now look closely at how Chomsky's minimalist program might relate to Cherniak's work in computational neuroscience.

6.3 Optimal computation and non-genomic nativism

The main point of the following exposition is to introduce and assess the asserted relationship between the minimalist notion of "optimal computation" and the neuroscientific concept of "optimal-wiring." As a starting point, it is necessary to introduce Cherniak's work to provide the reader with an overview of its philosophical and empirical aspects. With this in place, we can proceed to see how this work might be related to the minimalist conception of language as exhibiting "optimal design."

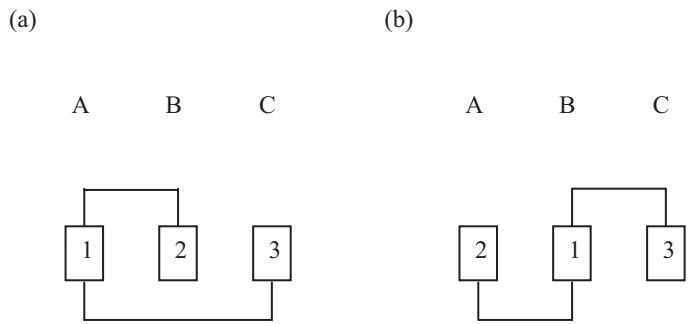
Cherniak (1990, 1994, 2005) sets out the philosophical framework upon which his work in neuroscience is based. He places his emphases in line with "the tradition of seeking simple underlying mathematical form in complex aspects of Nature, ranging from Pythagoras through D'Arcy Wentworth Thompson" (Cherniak 2005: 107). The basic idea underlying his approach is that human minds have only finite cognitive resources available in the brain to perform their functions; just as nature has limited resources, so also for humans' computational resources, i.e. we are bounded by so-called *minimal rationality* (on this, see Cherniak 1981). In this connection, Cherniak (1990) objects to the overestimation

of cognitive resources in connectionist models, in which the capacity for the development of neural connections is assumed to be virtually infinite. In contrast to this idealist view of the resources of the brain, he seeks a more realistic approach to cognitive science, where one should take seriously the motto "We do not have God's brain" (Cherniak 1994: 94).

Given that the brain is a finite device with limited resources available to it, Cherniak (1994) poses the following question: "If actual brain connections are in severely short supply, is their anatomy correspondingly optimized?" To tackle this question, he turns to the problem of "saving wire" and explains how it has been expressed and solved by formalisms developed within the framework of combinatorial network optimization theory. To illustrate this, he considers an example that has received considerable attention within the field of computer science, namely the problem of component placement optimization. This problem focuses on the question of how to plan a very large-scale integrated (VLSI) microcircuit, and Cherniak introduces it as follows:

> Given the interconnections among a set of components, find the spatial layout – the physical arrangement – of the components that minimizes total connection costs. The simplest cost-measure is length of connections (often represented as the sum of squares of the lengths); usually the possible positions for components are restricted to a matrix of "legal slots." (Cherniak 1994: 96)

To make this problem concrete, Cherniak considers a simple scenario in which components 1, 2, and 3 are to be placed in slots A, B, and C. He offers figures (1a) and (1b) as representing two of six alternative possible placements for this arrangement, where (1a) and (1b) correspond, he suggests, to the most and least total connection costs in terms of wire length, respectively:



(Cherniak 1994: 96)

Before proceeding and to allay confusion, we should note a couple of failings in Cherniak's presentation. First, he nowhere makes it explicit that component 1 must be directly connected to both component 2 and component 3. Given this assumption and the three slots, we have the 6 (= 3!) possible placements for the three components referred to above, *viz.*:

(a) ABC	(b) ABC	(c) ABC	(d) ABC	(e) ABC	(f) ABC
1 2 3	1 3 2	2 1 3	2 3 1	3 1 2	3 2 1

We can easily see that here placements (a), (b), (d), and (f) require the same total wire lengths, as do (c) and (e). Thus, what we actually have is not a single configuration exhibiting maximal wire length and another single configuration minimizing this property, as Cherniak's presentation appears to suggest, but two sets: one whose members exhibit maximal wire length by virtue of having placed component 1 on one side or the other; and the other whose members exhibit minimal wire length by virtue of having placed component 1 in the middle.

Returning to our exposition, we can observe that for *n*-components there are *n*! different possible arrangements, and just as in the simple case of three components, not all of them will differ in required wire length. However, it is immediately apparent that the greater the number of possible physical arrangements of the components, the greater the time of computation to reach the most optimal network configuration, supposing, of course, that all candidates are available for consideration. For example, solving a placement optimization problem involving only twenty components requires consideration of 20! possibilities, that is, "2.4×10^{18} layouts, more than the total number of seconds in the 20 billion year history of the Universe since the Big Bang" (Cherniak 1994: 96–7).

Cherniak next extends these considerations to the architecture of neural systems. In particular, he considers the brain as a VLSI microcircuit, with the number of its components varying depending on the level of analysis being considered. Thus at the highest level of analysis, he considers a biological one-component placement optimization problem by inquiring as to why the brain should be placed in the head rather than in any other part of the body, arguing that since the brain has more sensory/motor connections to the upper part of the body than to its lower part, it must be located in the head to minimize wire length. At a more finely grained level of analysis, he considers a fifty-component placement problem by focusing on the functional areas of the cerebral cortex. Given the expected astronomical computational cost involved in this problem, he introduces an "adjacency-rule" in investigating whether the

placement of components reflects an optimal neural network: "If components *a* and *b* are interconnected, then they are positioned contiguous to each other, other things equal" (Cherniak 1994: 98). After discussing similar problems the solutions of which seem to involve such neural optimization, Cherniak (1994: 104) concludes with the provocative question as to "whether *Deus sive Natura* can build the best of all possible brains without supernatural or magical powers."

Now, leaving metaphors aside, one point, which becomes clearer in Cherniak's subsequent work (cf. Cherniak 1995, 2005, and Cherniak *et al*. 2002, 2004, 2006), is that the hypothesis of the "best of all possible brains" is intended to provide a third alternative to the traditional nature versus nurture dichotomy (with "nature" here referring *only* to biology). To see what this might mean, let us focus on one aspect of Cherniak's work that has a more direct bearing on Chomsky's (2005) "Three Factors in Language Design." Specifically, we shall be concerned with the thesis of *non-genomic nativism*, according to which "generation of optimal brain structure appears to arise simply by exploiting basic physical processes, without the need for intervention of genes" (Cherniak 2005: 103), this "exploitation" providing the mechanism referred to at the end of the previous paragraph.

Consider first the following question: Why is neural structure the way it is? One possible answer that is consistent with Cherniak's computational approach is: a neural network is structured in an optimal way to save "wire length." Thus, Cherniak (2005: 105) posits his first hypothesis: "Optimization accounts for a significant extent of observed anatomical structure." Now, consider this question: Why is neural structure optimized? Here comes Cherniak's (2005) second hypothesis, namely: "Simple physical processes are responsible for some of this optimization." Cherniak, overlooking the force of the quantifier "some" in the latter hypothesis, combines the two hypotheses to arrive at the following schema:

Physics ➔ Optimization ➔ Neural structure

However, it should be noted that, *pace* Cherniak, this combined hypothesis is warranted only in those cases where a physical explanation is available. In other words, in cases in which wire length considerations are the only (more basic) explanation which can be offered for neural structure, we are forced to cite this design specification as our single *explanans* and end the story there; in which case the schema above would be reformulated as "Optimization ➔ Neural structure."

Cherniak (1992) and Cherniak *et al*. (1999) have defended at length the plausibility of a thesis along the lines of the combined hypothesis as illustrated

above, providing a range of evidence in support of it. For instance, it has been argued that certain types of neuron arbors are self-structuring, in the sense that they tend to minimize their total volume by obeying simple laws of fluid mechanics (Cherniak 1992: 508–9; Cherniak *et al*. 1999: 6005). Since these laws apply also to all other tree-like structures, including non-living ones such as river junctions, Cherniak (2005: 503) infers that "the brain may involve basic physical processes only: the genome seems to get the anatomy of local neural junction optimization automatically and directly from energy-minimization phenomena involving classical mechanics." Cherniak *et al*. (2005: 6008) arrive at the same conclusion: "Since river networks perform as well at topology optimization as dendrites and axons … DNA-based mechanisms do not seem to be required." Here is where the thesis of non-genomic nativism emerges: "some complex biological structure … is intrinsic, inborn, yet not genome-dependent" (Cherniak 2005: 107).

One question that arises at this point is this: granted the plausibility of this thesis, and given the experimental findings of a significant level of optimization associated with brain structure, what prompts organisms to make the most of this "free anatomy"? Cherniak's (2005: 107) answer to this question is that the brain's structure is too complex to be largely determined by the information-limited human genome, and it is most likely that much of this structure emerges directly from the routine workings of basic physical laws. One can perceive a clear parallelism here with a notable feature of the minimalist program, to wit: the structure of language is too complex to be largely determined by the recently evolved faculty of language, and it is more likely that much of this structure derives directly from third factor constraints. Indeed, just as Cherniak relies on the laws of physics to resolve the mismatch between brain structure and genome structure, so too does Chomsky rely on third factor conditions to ease the disparity between the apparent complexity of the language faculty and its relatively short evolutionary history (cf., however, Johansson's (2006) conclusions on the timing of language evolution, referred to in Section 5.2).

But the parallelism goes further. As mentioned earlier (p. 150), Cherniak considers his non-genomic nativism to be a third alternative to genetic endowment and environmental effects. This clearly matches Chomsky's (2005) postulation of three factors in language design. Moreover, Cherniak (2005: 107) concedes that brains "cannot grow like crystals," i.e. unconstrained by genetic instructions, but he argues that "life must still play by the rules of the game, subject to mathematical and physical law." Likewise, Chomsky (2000a: 22) does not deny the role of natural selection in shaping the growth and development of language, but he argues that "a belief in pure natural selection would be

totally irrational," and that there must be "a kind of a 'channel' set up by physical law." There is also an analogy between Cherniak's explanation of why the structure of the brain is the way it is and Chomsky's explanation of why the structure of language is the way it is. As indicated earlier, Cherniak's explanation has two components: (i) the brain is structured in such a way so as to minimize the total length of neural wiring, and (ii) it is the way it is simply because this is how the laws of physics work. Chomsky's explanation follows the same pattern: (i) language is designed in such a way so as to minimize the total number of derivational steps, more generally, to optimize computation, and (ii) it is the way it is simply because this is how physics works. Consequently, parallel to Cherniak's combined hypothesis schematized earlier, we might propose the following schema as an illustration of what Chomsky is striving to achieve:

Physics ➔ Optimal computation ➔ Language structure

This is what we might call "Chomsky's combined hypothesis," which asserts that (i) optimization is responsible for much of language structure, and (ii) *some* (or all?) of this optimization is a consequence of physical laws. Recall, however, that we have noted above that Cherniak's "combined hypothesis" is only valid for those cases that yield to the physical level of explanation, and it is obvious that Chomsky's is subject to the same reservation. Observe further that, taking an empirical perspective in Section 5.6, we have seen that there is little or no evidence for a physical basis of any aspect of optimal computation, so long as we take the requirements associated with such evidence seriously. In Section 6.8, I will argue to the conclusion that it is conceivable that *no* aspect of optimal computation is reducible to the "neatness" often associated with physical laws and that this non-reducibility may be principled, rather than merely empirical. As a first step to this conclusion, we need to introduce one of the major doctrines in the philosophy of mind.

6.4 Functionalism

Modern linguistics, especially in those varieties developed by Chomsky and his associates, is a discipline that takes seriously the ascription of mental states. Thus, for Chomsky (1980a: 51), "to know a language is to be in a certain mental state comprised of a structure of rules and principles (comparably, for certain other aspects of cognition)." Notions such as "mind," "mental representation," or "mental computation" are widely used in Chomsky's linguistics, although, of course, without ontological commitment other than to consider them as abstract

descriptions of yet largely unknown physical mechanisms (cf. e.g. Chomsky 1980a: 5). Moreover, the conception of language as a mental state, in the sense of a mental function that transforms a specific input into a certain output, appears to have obtained a wide currency among cognitive scientists, including linguists. It is therefore legitimate, indeed necessary, to examine whether views on the nature of such states have any implications for the sorts of issues that concern us here.

Now, it so happens that one particular view of mental states that enjoys a good deal of popularity among philosophers and others, *viz*. functionalism, is of considerable interest to the concerns of this chapter.[2] Before raising these matters, I will offer an overview of functionalism and explore some of its major features. Any philosophical doctrine has its adversaries, and functionalism is no exception.[3] Nevertheless, it continues to enjoy a pre-eminent role in the philosophy of mind, and for the purposes of the discussion that follows, I will assume its basic correctness. Much of the following exposition of functionalism relies on Fodor (2004 [1981]) and Block (2004 [1980]).

Traditional approaches to the philosophy of mind can be classified into two major categories: dualism and materialism. The distinction between the two is ontologically based; the former conceives the mind as a substance distinct from physical substance, while the latter makes no such distinction and only recognizes physical substance. The roots of this distinction can be traced to Cartesian philosophy. Descartes, as is well known, recognizes two fundamental kinds of substance: the mental (thought) and the material (extension). This substance dualism gave rise to one of the most famous problems in philosophy: the mind-body problem. As we shall see later (Section 6.7), Descartes conceived of physical causation in terms of "action-by-contact," that is, by direct "push" or "impact." Since nothing can push or make a physical impact unless it has an extension, the question arises as to how the mind, which is supposed to have no extension, can be causally efficacious with respect to material substances. This is the fundamental question underlying the mind-body problem.

Dualism and materialism are two different answers to this fundamental problem. The major failure of dualism resides in the fact that it does not provide a satisfactory explanation for mental causation. If the mind has no physical properties, it does not belong to the physical world, and thus the question arises as how the mental can influence and be influenced by the physical. Consider, for example, the simple act of raising a hand whenever one desires to do so. How could a bodily object (in this case, a human hand) be influenced by a mental state (having the desire to raise the hand)? This simple illustration, together with countless others, flies in the face of a dualist approach that postulates an

ontological distinction between mind and body and yet fails to account for their interaction.

Materialism, by contrast, posits that the physical world is "causally closed" and that the mind is essentially physical in nature. There are several varieties of materialism, but for our purposes it will suffice to consider two: logical behaviorism and identity theory. Common to these two approaches is the desire to maintain the tenets of materialism while at the same time seeking to make sense of mental causation.

Logical behaviorism sought to define the meaning of mental state ascriptions in terms of stimuli and responses. The basic idea was that each mental state ascription can, in principle at least, be expressed by a dispositional statement in the form of an *if–then* sentence. For instance, the statement "John ate an apple because he was hungry" including the "mental" state predicate "(be) hungry" can, it is supposed, be translated into the conjunction of a hypothetical statement and its antecedent, "If there were an apple available, then John would eat it, and there was an apple available," where no apparent reference to the mental survives. In this way mental causation was considered to be nothing other than a manifestation of an appropriate choice of behavioral disposition(s). As Block (2004 [1980]: 189) has put it, behaviorists "did not think mental states were *themselves* causes of the responses and effects of the stimuli," rather they "took mental states to be 'pure dispositions'" (emphasis in original).

Identity theory comes in two versions, the difference being the ontological level at which the identity relation is supposed to hold. One version is *type physicalism*, and the other is *token physicalism*. Type physicalism, standardly regarded as necessary for the correctness of *reductionism* (see below), maintains that every *mental type* is identical with its corresponding *brain type*, e.g. one might assume that the mental type "pain" is identical with the brain type "C-fibre firing." According to token physicalism, however, every *mental token* is identical to a specific *brain token*, e.g. every instance or token of the type "pain" is identical to an instance or token of some brain-event or other, which might, but need not be, a token of the brain-event type "C-fibre firing." For our purposes, it is important to bear in mind that the correctness of type physicalism entails the correctness of token physicalism, while the converse does not apply. Thus, type physicalism is stronger than token physicalism, and therefore it is more susceptible to being challenged.

As theories of mind, logical behaviorism and identity theory have both strengths and weaknesses, and, as we shall see later, it is precisely the virtue of functionalism that it inherits the strengths and overcomes the weaknesses of these two theories of mind. To illustrate this, it is necessary first to say

something about some of the advantages and disadvantages of these two predecessors of functionalism.

As regards logical behaviorism, its main advantage lies in the fact that it accounts for the relational character of mental states. For the logical behaviorist, "to have a headache," Fodor (2004 [1981]: 174) says, "is to be disposed to exhibit a certain pattern of relations between the stimuli one encounters and the responses one exhibits." As a consequence of this, the logical behaviorist is not committed to type physicalism. As Fodor (2004 [1981]) puts it: "If that is what having a headache is ... there is no reason in principle why only heads that are physically similar to ours can ache." The reason why not being committed to type physicalism is an advantage will become apparent when we consider the empirical and theoretical difficulties that this variety of identity theory gives rise to.

However, the chief weakness of logical behaviorism is its failure to account for complex causal interactions involving the mental, resulting in the untenability of the behaviorist thesis that mental causation can be avoided by reference to behavioral hypotheticals expressing behavioral dispositions. Consider, for instance, mental-to-mental causation, where mental states are said to cause other mental states. It may be argued that just as it is possible for a physical event (or state) to cause another, as when the falling of a tree is attributed to a strong wind, so too it is possible for a mental event (or state) to cause another, as in the case where my desire for drinking more coffee is occasioned by my fear of not being able to finish this book by the deadline. Since mental-to-mental causation cannot conceivably be translated into a behavioral disposition, the logical behaviorist's account of mental states is, at best, incomplete. Moreover, consider our earlier example "John ate an apple because he was hungry." As observed before, this might be translated into the conjunction "If there were an apple available, then John would eat it, and there was an apple available." However, there is certainly no reason to believe that it is always the case that when the antecedent of the hypothetical is satisfied eating is a consequence, for it is perfectly conceivable to imagine a situation in which there is an apple available, and despite John's hunger, he chooses otherwise, perhaps because he *believes* it to be poisoned.[4] What such examples entail is that mental causation is for real and cannot be eliminated in the way that the logical behaviorist proposes.

Turning now to the identity theory, it was advertised as overcoming the difficulties with which logical behaviorists were confronted. One such difficulty, noted above, has to do with the fact that some mental causes appear to be linked to other mental effects, with complex interactions of the mental resulting on occasions in specific behavioral outcomes. Now, it is not difficult to see

how identity theorists propose to deal with this by claiming that mental events are brain events. As Fodor (2004 [1981]: 172) has noted, on account of this claim one can "make sense of the idea that a behavioral effect might sometimes have a chain of mental causes; that will be the case whenever a behavioral effect is contingent on the appropriate sequence of neurophysiological events." Notice further that if mental events and processes are neurophysiological, one can also see how to overcome another difficulty in behaviorism, namely that mental causation, even if expressed in dispositional terms, might not lead to a behavioral effect. This is because if mental processes are brain processes, it follows that the former must have the causal properties of the latter. If true, then the property of not having a specific behavioral effect is *ultimately physical* in nature.

However, the main weakness in the identity theory, taking this now as subscribing to type physicalism, lies in its failure to account for the relational character of the mental. As observed, according to type-identity, every mental type is identical to a brain type. This, however, is too strong a claim, both on the empirical and the theoretical levels. At the empirical level, neuronal plasticity, a concept central to neuroscience since the work of Ramón y Cajal (1913–14), testifies against it, at least as far as the critical period of development of the brain is concerned.[5] At the theoretical level, there is no logical connection between (human) mentality and (human) brains. The empirical fact that mental states are instantiated by physical neurons may only represent a coincidence, not a logical relation. This point will become clearer in due course (cf. the multiple realization argument in the next section).

As a response to these and other weaknesses of its predecessors, and closely linked to developments in various fields within cognitive science, including linguistics, psychology, artificial intelligence, and the theory of computation, functionalism has emerged. Two features are common to these domains of inquiry: their object of study involves information processing systems, and their approach to such systems is conducted at a certain level of abstraction. As will become clear by the end of this discussion, functionalism can be viewed as an attempt to accommodate this level of abstraction. For now, let us see how functionalism seeks both to overcome the drawbacks of logical behaviorism and identity theory, and to capture the best of their features.

Functionalism defines a mental state in terms of its *causal role* in relation to perception, behavior, and other mental states. More specifically, what makes a mental state the kind of state it is, and which determines its essence, is the functional/causal relations it enters into with sensory inputs, behavioral

outputs, and other mental states. To be in pain, for example, is to be in a mental state which, *inter alia*, causes a disposition of calling a doctor in people who believe doctors are able to relieve their pain, causes a feeling of anxiety and discomfort, often causes someone to say "ouch" or something similar, and is brought on by different kinds of stimuli. It is important to see how this individuation of mental states by reference to their causal role allows functionalism to account for the causal and relational aspects of the mental – two aspects that logical behaviorism and identity theory respectively fail to explain.

The logical behaviorist may content himself with viewing mental states as dispositions – in the sense of defining them in terms of hypothetical statements – but he must concede that they are not *behavioral* dispositions (cf. Fodor 2004 [1981]: 175). By taking seriously the causal role of mental states the functionalist can avoid the main difficulty which the logical behaviorist encounters, namely that the consequences of stimulus inputs are not specified solely in terms of behavioral outputs. To return to our earlier example about John and the apple, we have observed that satisfaction of the antecedent in "If there were an apple available, then John would eat" may not necessarily result in the exercise of the relevant behavior; rather, the relevant consequences may also refer to mental states. While this observation is inconsistent with the behaviorist account, it is consistent with a functionalist view that ascribes a causal role to mental states.

With respect to the type physicalist, his account of mental states is constrained by the type of the underlying material from which these states are obtained; for him, without brain states there can never be mental states, as he can only conceive of the latter in terms of the former. By contrast, the functionalist is committed to the belief that mental states are *functional* states. This allows him to avoid the difficulties encountered by the type physicalist, namely: the empirical observation concerning the plasticity of the brain, and the illegitimate commitment to an identity relation between the mental and the neurological. Thus, unlike the type physicalist, the functionalist is not vulnerable to neurological findings about plasticity. If a certain area of the brain can take over the mental function of another area, this should not be surprising from the functionalist point of view, for mental states are functional, *not* neurological, states. Moreover, the functionalist is comfortable with the possibility that all kinds of different systems, physical and non-physical, might have mental states. As Fodor (2004 [1981]: 169) puts it, from "the functionalist view the psychology of a system depends not on the stuff it is made of (living cells, metal or spiritual energy) but on how the stuff is put together." Thus, if brain states (or events) turn out to be the *only* stuff with the functional properties of mental states, both

the type physicalist and the functionalist will be correct. But, if this turns out not to be the case, only the type physicalist need despair.

"It is no wonder," says Fodor (2004 [1981]: 175), "that functionalism has become increasingly popular." Indeed, as should be clear from the above discussion, functionalism seems to succeed in preserving the strengths of its predecessors and eschewing their limitations. We now turn to an exposition of the central argument of this philosophical doctrine.

6.5 The multiple realization argument

We have just observed that functionalism does not exclude the possibility that all kinds of different systems, physical and non-physical, might have mental states. This follows logically from the multiple realization argument (henceforth MRA), which states that every mental kind is (or can be) multiply realizable by different physical kinds.[6] The argument dates back to the 1960s when it was introduced by Putnam as part of his criticism of proponents of type physicalism (so-called "brain states theorists").

Putnam's (1967) MRA is a deductive argument: Premise 1, all psychological kinds are multiply instantiated by different physical kinds; Premise 2, if a particular psychological kind is multiply instantiated by different physical kinds, it follows that this specific psychological kind cannot be identified with any particular physical kind; Conclusion: no psychological kind can be identified with any specific physical kind (cf. Bickle 2008). Fodor puts the same argument in the language of computer science:

> The problem with type physicalism is that the psychological constitution of a system seems to depend not on its hardware, or physical composition, but on its software, or program. Why should the philosopher dismiss the possibility that silicon-based Martians have pains, assuming that the silicon is properly organized? And why should the philosopher rule out the possibility of machines having beliefs, assuming that the machines are correctly programmed? If it is logically possible that Martians and machines could have mental properties, then mental properties and neurophysiological processes cannot be identical, however much they may prove to be coextensive. (Fodor 2004 [1981]: 173)

A provocative way to pursue this perspective is to apply the computer metaphor to which Fodor alludes to the following passage from Chomsky (1978), in which he expresses his hope of finding evidence for the physical basis underlying universal grammar *qua* "mental" program:

> Ultimately, we hope to find evidence concerning the physical mechanisms that realize the program; it is reasonable to expect that results obtained in the abstract study of the program and its operation should contribute significantly to this end (and, in principle, conversely; that is, information regarding the mechanisms might contribute to understanding of the program). (Chomsky 1978: 201)

Given the functionalist view and the computer metaphor that seeks to make it more concrete, the question arises as to whether Chomsky is justified in considering it reasonable to expect that our knowledge of the program should contribute to our knowledge of the physical system that instantiates it. Given the MRA, there seems to be no justification for this. Thus, utilizing the computer metaphor, on the assumption that what we might call "universal grammar software" is instantiated by physically different hardwares, we might argue that one cannot deduce anything about the hardware from theoretical knowledge of the software. Put in more direct terms, there seems to be no reason for the suggestion that, from our studies of UG, we should expect insight into the physical mechanisms that realize it; for if UG can, in principle at least, be instantiated by different physical systems made of different materials (neurons, silicon chips, or even my old car parts), we cannot, *contra* Chomsky, infer anything about the neural representation of UG from what we know at the computational level. Observe in passing that Chomsky also runs the inference the other way when he says that "information regarding the [physical] mechanisms might contribute to understanding of the program." At first sight, the MRA might not seem to prohibit such an inference. However, as we shall see in Section 6.8, there are good reasons to regard such an inference with caution. For now, we continue our discussion of the MRA.

Fodor (1975: 11) suggests that every science has its predicates and is largely individuated by them. Now, the concepts of "software" and "hardware" represent two theoretical spaces in which the psychologist and the neurologist construct their respective theories. Thus the question arises as to whether the two spaces *coincide* with each other, that is, whether the generalizations of psychology can be expressed as neurological generalizations or, at an even more basic level, generalizations of physics.

Building on Putnam's ideas, Fodor (1974) uses the MRA to argue against the prospects of reducing the special sciences, including psychology, to physics. He starts by emphasizing a distinction between token physicalism and reductionism. The former is seen by him as "simply the claim that all the events that the sciences talk about are physical events," and the latter as "token physicalism

with the assumption that there are natural kind predicates in an ideally completed physics which correspond to each natural kind predicate in any ideally completed special science" (Fodor 1974: 100). The classical reductionist argument has it that, since all events which the special sciences describe are physical events, they should be captured by the generalizations of physics.

To challenge the reductionist view, Fodor (1974) argues, first, that if we assume every science is individuated by reference to its typical predicates, where these are what appear in its laws and empirical generalizations, it follows that not all predicates are physical predicates. Now suppose that $S_1x \rightarrow S_2y$ is a psychological law with "$\rightarrow$" being read as "causes." For reduction to proceed there has to be a law $P_1 \rightarrow P_2$ in the more basic science and "bridge laws" underwriting a mapping between psychological predicates and predicates in the more basic science. Such bridge laws will have the form $S_1x \Leftrightarrow P_1x$ and $S_2y \Leftrightarrow P_2y$, and the connective they contain is different from that which appears in other laws. The fact that this connective establishes a symmetric relationship between the two kinds of predicates indicates that such a relationship does not signify causation but a species of *identity* (cf. Fodor 1975: 20). Crucially, Fodor (1975: 19) now argues that there is no reason to believe that psychological natural kind predicates are nomologically co-extensive with physical natural kind predicates. "What seems increasingly clear," he goes on, "is that even if there are such coextensions, they cannot be lawful" (cf. Block and Fodor 1972: 163). That, we saw earlier, follows from the MRA; for "there is an open empirical possibility that what corresponds to the kind predicates of a reduced science may be a heterogeneous and unsystematic disjunction of predicates in the reducing science" (Fodor 1975: 20). If this is true, then an implementation of reduction will result in statements in the more basic science containing (possibly massive) disjunctive terms in its generalizations, and the latter would not comprise laws of that science. To forcefully illustrate this point, Fodor (1975: 21) observes that the statement "(either sunlight or friction) causes (either photosynthesis or heat)" is *not* a law of physics.

Before we conclude, one point is worth stressing: functionalism, and its compatibility with multiple realization, does not preclude physicalism. Indeed, if the argument from multiple realization is correct, it is *plausible* that creatures or entities, whose *physical* composition is distinct from ours, might share with us (some of) our psychological states. For this reason, functionalism is perfectly compatible with token physicalism. However, the claim is crucially different when we consider type physicalism, the correctness of which requires that psychological *properties* should "translate" as coherent physical properties

across a range of distinct physical systems. The implausibility of this latter is what is important for the present discussion.

Perhaps the most significant conclusion to emerge from the MRA is that the level of abstraction at which psychological generalizations are made might be *principled*. What this means is that, even on the assumption of an ideal physics, where all physical laws are known, we might still not be able to preserve the significance of psychological generalizations at the level of physics.

The ramifications of the above matters for Chomsky's naturalism and, more specifically, for optimal computation and the biolinguistic approach to language, are the subject of the remaining sections of this chapter, and it might be useful to offer some pointers before we embark on the task of making explicit what we argue to be a tension between the tenets of Chomskyan naturalism and those of psychological/computational functionalism. We will examine this tension at two different levels. At a more general level, several conflicting points emerge. These include issues relating to Chomsky's position on the mind-body problem, his skepticism about the prospects of reconciling commonsense with cognitive science, and his optimism concerning the prospects of unification in science (Section 6.6). We pursue this discussion further in Section 6.7, where we take a closer look at Chomsky's views on functionalism. At a more specific level, we focus on the central thesis of this chapter, namely that the MRA poses a serious challenge to the minimalist goal of going "beyond explanatory adequacy," at least as far as the explanatory notion of "optimal computation" is concerned. This thesis will be developed in Section 6.8, where we examine the nature of optimal computation against the background of the MRA, arguing that it cannot straightforwardly be seen as providing the basis for the sort of "principled explanation" that Chomsky is anxious to provide. The implications of this conclusion for Chomsky's biolinguistics will then be discussed in Section 6.9.

6.6 Functionalism and naturalism: uneasy bed partners

One way to observe the tension between Chomsky's naturalism and Fodorian functionalism is to consider Chomsky's position on the mind-body problem. He asserts, for instance, that "[t]here seems to be no coherent doctrine of materialism and metaphysical naturalism, no issue of eliminativism, no mind-body problem" (Chomsky 2000a: 91). This rather radical position is based on a particular reading of the impact of Newton's work on Cartesian dualism. Chomsky argues that the mind-body problem made sense only in the pre-Newton era, but after Newton's introduction of "action at a distance," the Cartesian concept of "body" became devoid of coherence, and, consequently, the

mind-body problem became unformulable. It will become apparent as the discussion proceeds that this view of the mind-body problem has direct implications for the topic of this chapter, and we shall return to it in the next section. For the moment, we need to see how Chomsky's position on this metaphysical problem creates the tension we are interested in here.

Chomsky (2000a: 84) seems to draw three conclusions from his preferred interpretation of the Newtonian impact on the Cartesian relationship between mind and body. These are: (1) mental phenomena are properties of organized matter; (2) we can no longer expect the world to be as intelligible as it was once thought to be; and (3) there is currently no coherent notion of "body." Not surprisingly, all these conclusions have favorable implications for Chomsky's naturalism. However, the first conclusion, if stated non-categorically and with some qualifications, is perfectly compatible with functionalism, whereas the second and third conclusions lead to certain implications that are in conflict with some of the basic tenets of functionalism. Let me explain in some detail what I mean by this.

Take the first conclusion. It clearly echoes one of the basic assumptions of Chomsky's methodological naturalism, according to which one "can only assume that those phenomena 'termed mental' are the result of the 'organical structure' of the brain" (Chomsky in Cela-Conde and Marty 1998: 21). Now, if mental phenomena, including language, are properties of organized matter, it follows that it may be possible to study these properties in the same way as any other properties of organized matter (e.g. electrical, chemical, optical, etc.), which has been and still is the standard view within Chomsky's "naturalistic approach" to language. Ontological claims about the mind aside, this approach is compatible with some functionalist efforts to "naturalize" the mind, i.e. to provide an account of (at least some) mental phenomena with whatever means are available to natural science. We will come back to this point shortly, but for the moment it suffices to say that, insofar as the task is to find a convincing "story" of how mental phenomena might fit into the natural world, there seems to be no conflict between Chomsky's naturalism and Fodorian functionalism. Conflicts do arise, however, when we turn to the second and third conclusions noted above.

As with the first conclusion, Chomsky sees in the second conclusion a justification for his methodological naturalism. From Chomsky's perspective, if Newton did indeed show that the world was not as intelligible as Cartesian scientists had thought it to be, then the only hope available to us in understanding the world is by constructing intelligible theories. This would involve eschewing our common sense intuitions and relying instead on a more abstract level of rational inquiry, namely the "Galilean style":

> Newton essentially showed that the world itself is not intelligible, at least in the sense that early modern science had hoped, and that the best you can do is to construct theories that are intelligible, but that's quite different. So, the world is not going to make sense to common sense intuitions. There's no sense to the fact that you can move your arm and shift the moon, let's say. Unintelligible but true. So, recognizing that the world itself is unintelligible, that our minds and the nature of the world are not that compatible, we go into different stages in science. Stages in which you try to construct best theories, intelligible theories. So that becomes another part of the "Galilean style." (Chomsky 2000a: 4)

Clearly, Chomsky considers common sense intuitions to be of dubious epistemic value if taken as a basis for rational inquiry in the natural sciences. Indeed, he believes that "natural science quickly departs from folk theories, and it is presumably on to something when it does so" (Chomsky 2003a: 262). By contrast, Fodor does try to make a strong case for the principles of folk psychology. Consider, for instance, Fodor's computational/representational theory of mind, in which he expresses his commitment to *intentional realism*, according to which entities that are like propositional attitudes are psychologically real and causally efficacious. We should also recall Fodor's attempt to demonstrate the compatibility between his intentional realism and physicalism. As Fodor (1987: 16) put it, there is "no reason to doubt that it is possible to have a scientific psychology that vindicates commonsense belief/desire explanation." Now, while many philosophers of mind consider Fodor's effort as part of a wider project directed towards the "naturalization of mind," Chomsky (2003a: 262) considers such endeavors to be nothing more than an "indication that Functionalism has taken the wrong course," that is, "mistaking ethnoscience as the natural science of the mind." For he takes it that common sense intuitions fall under the domain of ethnoscience, and have nothing to do with natural science.[7] Evidently, we can see a conflict here between Fodor's optimism regarding the prospects of reconciling commonsense conceptions of folk psychology with cognitive science on the one hand, and Chomsky's skepticism regarding the status of such an enterprise on the other.

As noted above, the third conclusion Chomsky draws from the impact of Newton's work on the mind-body problem refers to the claim that there is no longer a coherent notion of "body" (or "physical," "material," etc.). Unlike the two previous conclusions, this one is in tune with the substantive (rather than the methodological) aspect of Chomsky's naturalism. This is because if the notion

of "physical" lacks coherence, then what we call the physical "world simply offers a loose way of referring to what we more or less understand and hope to unify in some way" (Chomsky 2000a: 84). In other words, Chomsky seems to base his optimism regarding the prospects of unification in science on the belief that, since there is no coherent notion of "physical," there should be no (metaphysical) reason to mark a dividing line between the physical and the mental, except when such a division is considered methodologically convenient. In this respect, his views would seem to conflict with the basic tenets of functionalism, for this philosophical doctrine is based on a non-trivial distinction between the mental and the physical and is committed to the autonomy of the former from the latter – an autonomy which must be understood in this context as one that prohibits the identification of a mental property with a physical property.

In fact – and this is a point we shall return to later – Chomsky's position on the mind-body problem seems to conflict not only with functionalism, but also with the whole of experimental (cognitive) psychology, a field which operates largely with the ascription of functionally-construed mental states. Insofar as modern cognitive psychology can be construed as the empirical investigation of the properties of mental states, we might justly regard its activities as an attempt to make sense of that part of the world we call "mental," and in this sense modern cognitive psychology is on a par with Chomsky's own framework. Yet, not only does Chomsky seem to imply that modern cognitive psychologists are wasting their time and energies, but also that post-Newtonian philosophers such as Hume and Kant, and modern philosophers such as Russell and Popper, along with countless other philosophers who approached the mind-body problem from different perspectives, have somehow been missing the point all along!

More importantly, as far as the purposes of this chapter are concerned, if the mind-body problem is indeed unformulable post-Newton, and if as a consequence of this the dividing line between the physical and the mental should be considered as merely methodologically convenient, then the MRA will also be unformulable. This is perhaps why, as we shall see shortly, Chomsky is skeptical about the force of the MRA.

6.7 Chomsky's case against functionalism

Consider what Chomsky has to say on functionalism:

> The "Cognitive functionalist" approach seems to me to draw from Cartesianism the wrong property: the dualism that made sense as a

> scientific hypothesis when Descartes formulated it, but that cannot be sustained, as Newton showed. *Cognitive functionalism reconstructs a dualistic perspective in a form that is methodologically useful as a way of investigating the world in the present state of our understanding . . . But it should not be regarded as anything more than a temporary convenience*, in my opinion, and surely not invested with any metaphysical claim. (Chomsky in Cela-Conde and Marty 1998: 21, my italics)

To evaluate the content of this passage in a way that does justice to functionalism, we should first be clear about what species of dualism Chomsky is attributing to the "cognitive functionalist approach." More specifically, we should be careful not to equate, as Chomsky seems to, Cartesian dualism with the "dualistic perspective" that he attributes to functionalism. Failure to do so will lead to the assumption, implicit in the passage above, that functionalism considers Cartesian dualism to be in principle unavoidable, but nothing can be farther from the truth. Functionalism is ontologically neutral, and it is as compatible with dualism as it is with materialism or even idealism (cf. the discussion of token physicalism in Section 6.4). With this point clear, let us turn our attention to two claims made in the passage above: (1) Newton showed that Cartesian dualism *qua* scientific hypothesis made no sense; (2) functionalism inherited from Cartesianism the "wrong property" – the ontological dualism between mind and body. We examine these two claims in turn.

As noted, Chomsky maintains that the mind-body problem is a dualistic hypothesis that made sense only when it was formulated in pre-Newtonian terms, but after Newton had demolished the mechanical philosophy, by introducing into his mechanics the "mysterious" force of "action at a distance," the Cartesian notion of "body" lost its coherence and has never been replaced by a more coherent notion; consequently, the mind-body problem has become unsustainable ever since Newton. As Chomsky (2002: 53) put it: "Mind-body dualism is no longer tenable, because there is no notion of body." Now, I contend that this reading of the history of natural philosophy is mistaken. To fully explain why would carry us too far from our theme, so suffice it to bring forward a few observations from the history of science. These observations will, I maintain, establish the plausibility of the claim that Chomsky's reading is indeed misguided.

First, the notion of action at a distance refers to a form of interaction whereby two bodies act on each other without coming into actual contact.[8] Whether and how such a species of action is possible is, as the physicist James Clerk Maxwell put it, "a question which has been raised again and again ever since men began

to think" (Maxwell 2003 [1876]: 311). Indeed, action at a distance is as old as magic itself; the Greek atomists rejected it and suggested instead that all bodies (or, more accurately, atoms flying through empty space) act, and are acted upon, by *touch* (O'Keefe 2005: 80–1). Second, Descartes followed in the footsteps of the Greek atomists and defended a notion of action by *contact*. These two observations alone indicate that "action at a distance" constituted a problem which Descartes believed himself to have overcome. But how did Descartes justify his notion of action by contact? To answer this question we need to see how he defined "matter" in the first place. As noted in Section 6.4, Descartes considered *extension* to be the essence of "bodyhood" or "materiality." He rejected the "void" or empty space, "for geometrical space was extension and thus the very essence of body or matter" (Popper and Eccles 1977: 178). Thus, for Descartes space and matter are identical, and the distinction between them is only an illusion; space is "a matter just as real and as 'material' ... as the 'gross' matter of which trees and stones are made" (Koyré 1957: 75). Now, this identification of matter and geometrical space (i.e. extension) has two consequences. The first is Cartesian dualism; the mind is the only entity that has no extension and, therefore, it must be regarded as a substance different from matter. The second is Cartesian causation; the assumption of action by contact or push is "the only kind of causal action which Descartes had permitted, since only push could be explained by the essential property of all bodies, extension" (Popper 1963: 143). Thus, Cartesian extension offers an essentialist explanation of both the notion of action by contact and the exclusion of mind from the material world.

Now, to say that the Cartesian notion of "body" (i.e. an extended substance) lost its coherence as a result of Newton's introduction of action at a distance is misleading; it would be just as correct to say that the centuries-old notion of "action at a distance" renders the Cartesian notion of "body" incoherent. The fact of the matter is that Cartesians had good reason not to worry about "action at a distance," Newtonian or otherwise; their fundamental notion of "extension" prohibits such a mode of physical causation. Why then did they feel offended by Newton's "spooky" action at a distance? The reason for this is that, unlike Cartesian causation, Newton's mode of causation lacks an essentialist ground for legitimization. It should be borne in mind that this was the very reason why Newton himself did not feel satisfied with the principle of "action at a distance." Newton, like Descartes, was an essentialist; he believed in an ultimate explanation of natural phenomena (cf. Popper and Eccles 1977: 192). More important for our discussion, Newton's seeking an essentialist explanation could *not* have been because he felt his conception of gravity in terms

of action at a distance was at odds with the Cartesian conception of "body." There are two reasons for this.

First, Newton's definition of "body" differs from that of Descartes (cf. Chomsky in Bricmont and Franck 2010: 105, where he seems to believe otherwise). He did not believe that extension *alone* was the defining property of matter; he added hardness, impenetrability, mobility, and inertia to the list of so-called primary qualities of matter (Koyré 1957: 127). In addition, he believed in the "void" and rejected the Cartesian identification of the essence of matter with extension (Koyré 1957).

Secondly, had Newton accepted Descartes' definition of "body," his theory would have been inconsistent; for as we have observed, the only physical causation permitted by an extension-based conception of matter is action by contact, excluding action at a distance. But Newton, far from feeling his theory inconsistent, felt the need to attach an essentialist explanation to the latter mode of causation. His problem was: how can one explain an "action at a distance" in terms of an "action by contact" without assuming the validity of the (Cartesian) maxim that there is no void? That this was indeed the problem with which Newton was faced seems to be indicated by what he says in a letter to Bentley: "It is inconceivable that inanimate brute matter should, without the mediation of something else, which is not material, operate upon and affect other matter without mutual contact" (Newton 2004: 102). It is worth noting that Newton here formulates the action-at-a-distance problem in terms of a preconceived notion of what matter is (cf. the list of primary qualities of matter referred to above). Thus, it is hardly the case that Newton was troubled by the notion of "body," Cartesian or otherwise, and this is further clear from the fact that he committed himself to describing the notion of "body" in such a way "that we can hardly say that it is not body" (Newton 2004: 27). Indeed, he was guided, not by the question of what constitutes "material bodies," but by the question of how to bridge the gap between them in such a way as to account for their interaction. He explored different possible solutions, but we will not consider them here.

In view of these observations, we can now see that Chomsky's interpretation of the impact of Newton's work on Cartesian dualism is misguided in at least two respects. In the first place, Newton's action at a distance was in conflict, not with the Cartesian essential property of body (i.e. *extension*), but with the Cartesian principle of causation (i.e. action by contact or *push*). Kant captures this point when he credits Newton with being "the first one who suspended the mechanical mode of explanation" by attributing to "matter a power of *attraction* ... which does not at all depend on the shape of the

matter" (Kant 1997: 32, emphasis in original). Second, the Cartesian notion of "body" was not the only notion available at that stage of the history of science and, therefore, the dualism between mind and body does not stand or fall with Cartesian dualism; Descartes' notion of body might have been incoherent, but the relationship between mind and body was as mysterious in the time of Descartes as it is now.

Before leaving this point, I wish to stress that it is one thing to claim that Newton's notion of "action at a distance" was at odds with the (Cartesian) mechanical mode of explanation, and it is quite another to claim that it was in conflict with the notion of "body." Chomsky seems to conflate these two claims when he concludes, from the impact of Newton's work on Cartesian dualism, that the mechanical philosophy was demolished *and* that the notion of body became incoherent.

Turning now to Chomsky's second claim (*viz*. functionalism inherited the "wrong" dualistic perspective from Cartesianism), first I think it is fair to say that the passage quoted above (p. 164) tells us more about Chomsky's own naturalism than the philosophical doctrine he criticises. Consider, for example, the import of the italicized portion of the passage, which seems to suggest that the level on which explanatory theories of mind rest should be regarded as "a temporary convenience" which may not stand up to further scrutiny at a more fundamental level, say that of neurology. In effect, what we are witnessing here is Chomsky's "wait and see" attitude towards the future of science and the prospects of unifying cognitive science with brain science.

Now, the "dualistic perspective" which Chomsky mentions in the passage above, if intended to identify a position adopted by modern cognitive psychologists, refers to nothing Cartesian. Rather, it amounts to the thesis that no matter how sophisticated our knowledge of the brain and its activity may become, there are good reasons to suppose that we shall continue to be unable to translate our explanatory theories of the mental into the material mode *in such a way as to preserve explanatory adequacy*. This thesis is a straightforward consequence of the MRA and it is important to realise that, if this thesis is plausible, it follows that Chomsky's optimism on the prospects of unification in science is misplaced at best, and plain wrong at worst. But, clearly, Chomsky remains sceptical, for he says:

> Though it is possible and sometimes useful to study certain properties of a system X in abstraction, it would be an unacceptable form of dogmatism, in my opinion, to reject insights into the properties that derive from other ways of studying the system X . . . Suppose we have

> two theories of cognitive function, and it is discovered that only one of them is compatible with brain structures. *It would make little sense to disregard this evidence on the grounds that we are investigating mental functions in abstraction from brain structures*. (Chomsky in Cela-Conde and Marty 1998: 20, my italics)

There are two points to consider here. First, given the italicized portion of this passage, Chomsky seems to miss the point about what the real objection is. What is at stake here is not that the evidence should be dismissed because "we are investigating mental functions in abstraction from brain structures," but that the evidence cannot even be considered relevant without assuming type-identity between mental states and brain states.[9] Second, it is clear that Chomsky considers the level of "abstraction" at which theories of cognition are formulated to be only "temporary" and perhaps "useful," which suggests that he thinks it possible that this "abstraction" will be removed at some future time. At this point, the conflict between his naturalism and classical functionalism ceases to be merely methodological and becomes substantive. Thus, on the one hand, we have Chomsky's (2006: 185) "realistic prospect of moving significantly beyond explanatory adequacy to principled explanation," and, on the other hand, we have Fodor's (2007: 9) assertion that "[e]ven if basic physical laws are true of everything, they don't explain everything." In other words, while Chomsky (2010) seems to believe that linguistic "laws" and principles will ultimately be explicated by general physical laws, Fodor (1974, 1997) maintains that, even if all events are physical events and describable in the language of physics, there is no reason to suppose that the relevant "laws" and principles will fall under physical laws.

As we have just noted, Chomsky accuses functionalism of inheriting the "wrong property" of dualism from Cartesianism, but we have tried to show that this charge cannot be sustained. On the other hand, Chomsky's naturalism might well be guilty of a similar charge, namely that of inheriting the "wrong property" of type-physicalism from the "old fashioned" identity theories and the classical project of the unity of science. Now, Chomsky (2003a: 261) maintains that "there is no interest in taking 'mental types' to be non-biological, any more than there would be in defining 'chemical' or 'optical types' that share some properties of chemical and optical aspects of the world." I do not wish to question the interest of Chomsky's biolinguistic approach, but what I do question is his claim that "computational theories of language . . . require no identity theory" (Chomsky 2003a: 260). Thus, insofar as we are able to demonstrate that the minimalist program, in the perspective it adopts on the nature of true

explanation presupposes something approximating type identity between computational states and neurological states, the tension between this aspect of the program and functionalism will become apparent. As we shall see from the discussion that follows, such a tension indicates what appears to be a serious challenge to this aspect of the minimalist program.

6.8 Optimal computation versus multiple realization

Let us briefly remind ourselves of one aspect of the parallelism we have drawn between Cherniak's neurological work and Chomsky's minimalist program. As noted in Section 6.3, Cherniak's notion of "optimality" is defined in terms of the biological property of total length of neural "wire" connections; the shorter wiring a neural structure has, the more optimal it is. Correspondingly, the Chomskyan notion of "optimality" can be understood in terms of, for example, the principle of economy of derivation, which we have already met (Sections 2.5 and 5.5). Although we focus here on this principle, we do not mean to imply that other aspects of optimal computation (e.g. minimal search, no-tampering condition, etc.) should be excluded from more comprehensive discussion. Rather, our choice is made for the sake of simplicity of exposition and because economy of derivation can intuitively be linked to minimizing wire length in a more transparent way than other aspects of optimal computation. Thus, corresponding to the biological property of total length of neural wire connections, we have the computational property of the number of derivational steps; the fewer derivational steps a linguistic structure has, the more optimal it is. Now, this could have been an innocent parallelism were it not for the fact that Chomsky has, on occasions mentioned Cherniak's wire length speculations as providing support for the status of optimal computation and its role in providing genuine explanations. Presumably, then, we can say that, by invoking Cherniak's neurological work, Chomsky appears to identify mental (linguistic) states in which computation is optimized with brain states in which wire length is optimized, and it is here where the charge of subscribing to a position that has some resemblance to an acceptance of type identity between the mental/computational and the physical has at least *prima facie* force.

Of course, one might object to the above by pointing out that there is no reason to assume that the kind of identity involved here must be associated with *type*s, for it is possible that what we are confronting involves no more than *token*s. Thus, rather than saying that each type of computational (linguistic) state (or process) will be identical with a given type of brain state (or process), we might suggest instead that every token of a linguistic computational state

(or process) will be identical with a token brain state (or process). But notice that if this is indeed the case, then reference to Cherniak's neurological work is not clearly legitimate in seeking support for the "true explanatory" status of optimal computation. Let us see why this is so.

An assertion of type identity is intended to provide a basis for explanation. In the first place, if observed cases testify to the fact that there is a one-to-one correlation between a psychological type P and its corresponding neurological type N, then the identity relation between P and N constitutes an explanation for P in terms of a more basic science. For instance, if whenever any individual is in pain it is found that their brain is in a state in which C-fibres are firing, then it is reasonable to propose that C-fibres firing provides an explanation for pain. But notice that this strategy is justified only on the assumption that the relevant type identity is plausible. For if we assume, instead, that *only* a specific token of the mental type "pain" can be identified with a token of the physical type "C-fibres firing," other tokens of "pain" not having this characteristic, it would follow that the neurologist's explanation of "pain" would no longer be available. Simply put in terms of the law of transitivity, we have:

Premise I:	the reduction of psychology to neuroscience is at least in principle possible if and only if the strong hypothesis of type identity is correct, namely that psychological types map one-to-one onto neurological types.
Premise II:	reference to neurological evidence to explain a particular psychological type is warranted if and only if reductionism (in the sense of Premise I) is possible.
Conclusion:	reference to neurological evidence to explain a particular psychological type is warranted if and only if type identity is correct.

From this it follows that Chomsky's appeal to Cherniak's concept of "wire length," an appeal which he sees as underwriting the minimalist reliance on optimal computation as providing a contribution to "true explanation," is not justified unless the correctness of something akin to a type identity is assumed.

So far, we have been concerned with demonstrating that Chomsky's appeal to Cherniak's work presupposes the correctness of identity between the linguistic/computational type of "shortest derivation" and the neurological type of "minimal wire length." We are now in a position to explore the tension between the SMT's view of optimal computation and the views on the mind that we have been exploring here. At the heart of this tension is the MRA, and the difficulties it raises for the minimalist standard of "true" or "principled" explanation.

Let us begin by considering how the notion of "explanation" is applied in the context of Cherniak's non-genomic nativism concerning the properties of neural structure (Section 6.3). Cherniak's thesis can be summarized as follows:

Initial conditions + Laws of Physics ➔ Physical optimization of a system (in this case neural structure) in which *wire length* is *minimized* (here "wire length" is a biological property)

This thesis would become a concrete proposal if we specified the content of what we have on the left side of the arrow. But for the purposes of our discussion, we will assume that these are available. The important thing to realize here is this: if we can deduce what we have on the right-hand side of the arrow from the content of what we have on the left, we will then be warranted in believing that we have provided a causal explanation for *physical optimization* of neural structure (i.e. an explanation of why neural structure exhibits the property of minimal wire length). If we now turn to examine Chomsky's thesis, we find the same basic pattern of explanation:

Initial conditions + Laws of Physics ➔ *Computational optimization* of a system (in this case linguistic structure)

Notice, however, that although Cherniak's thesis and Chomsky's thesis share the same pattern of explanation, they do not refer to the same kind of *optimization*; the former concerns the physical optimization of neural structure, and the latter psychological/computational optimization of linguistic structure. Now, assuming that we have reasons to suppose that physical optimization of neural structure is a necessary consequence of physical laws – in the sense that these laws produce only "neat" physical/neural systems – we could see optimal computation as resulting from, and being ultimately explained by, physical laws if psychological/computational systems exhibiting optimal computation always *correspond* to neural systems displaying the attractive consequences of having been shaped by physical laws, in this case, minimal wire length. But it is precisely this notion of "correspondence" that the MRA calls into question. Let us see how this argument might apply to Chomsky's thesis as sketched above.[10]

According to the MRA, a computational system with this or that property can be built out of different physical systems with different kinds of physical properties. In particular, we might suppose that a computational system, which displays the property of optimal computation (understood here in terms

of the smallest number of derivational steps), can be built out of two distinct systems: one which exhibits physical optimization with minimal wire length, and the other which exhibits physical non-optimization with non-minimal wire length. Thus, the psychological or computational notion of optimization is, in principle, compatible with both physical optimization and non-optimization. But observe that, if computational optimization is indeed compatible with physical non-optimization, it follows that the "fact" that the laws of physics can be recruited to account for physical optimization can hardly be cited as an explanatory principle for computational optimization.

In fact, these considerations lead to a stronger conclusion. Suppose, instead of having an optimal computational system, we had a system which displays *non-optimal* computation, that is, a system in which computations with the smallest number of derivational steps do not obtain. The MRA again tells us that this can, in principle, be instantiated by two different systems: one of which exhibits physical optimization by displaying minimal wire length, and the other which does not. Now, given that Chomsky appeals to Cherniak's thesis as outlined above to explain optimal computation, consistency would require him to have a "real" explanation for non-optimal computation, in which case his position collapses.

The upshot of the argument is that we have no reason to suppose that properties that we might regard as somehow analogous (e.g. length of derivation vs. wire length) are preserved as we move from one level to another. Rather, it is a possibility that what corresponds to the computational property of "length of derivation" may vary from token to token so that the type can only be seen as corresponding to an open-ended disjunction of physical properties. And what is true for these properties is also true for the predicates that designate them, which would mean that the generalizations into which these predicates enter will not survive the journey between the computational level and the neurological level. However, it is not our purpose to press this point, for it is not at all clear that generalizations about syntactic derivations constitute "laws." This is why our discussion has focused on applying the MRA to a set of properties rather than a set of laws, but it should be clear from our discussion of functionalism (Section 6.4) that this argument is equally forceful in both cases.

6.9 Implications for the biolinguistic approach

Our discussion suggests a number of implications for Chomsky's biolinguistic approach to language in general and for the explanatory status of

optimal computation in particular. For one thing, it raises the question as to whether the "bio" in "biolinguistic" is really significant. The MRA suggests that it is not; but if that is indeed the case, one should not see this as undermining in any way the scientific status of linguistics or even of cognitive science at large. If a more or less plausible computational account of the language faculty can be realized in *non-biological* terms, and if this is the most we can hope for, no one would find the computational account in the absence of biology lacking in scientific character unless he thought that biology was somehow necessary to stamp linguistics with "the seal of science." Indeed, the scientific merit of any discipline should be based not on its public relations with other disciplines, but rather on its own results which, while they may not be somehow legitimized by results from other disciplines, may nevertheless be found satisfactory at a certain level of understanding. There are certainly many cognitive scientists who do not feel that the computational character of their discipline should force them to ground their speculations about mental states in an account of this or that bit of the brain working in this or that way. Rather, they seem to have followed Marr's (1982) advice in keeping their inquiry focused on the computational and algorithmic levels without worrying about the implementation level.

True, it may be argued that Marr considered the distinction between his three levels of analysis as methodologically convenient rather than conceptually sound (cf. Marr 1982: 28). It may also be argued that even a zealous functionalist such as Block (1995), who wrote that "the computer model of the mind is profoundly *unbiological*," felt it necessary to add that "cooperation between the biological and computational approaches is vital to *discovering* the program of the brain" (Block 1995: 390, his emphases). Yet, what cannot be denied is that the evidence that we have received to date from the brain sciences remains unclear in its significance. So unclear in fact that a recent work on neurolinguistics asserts that attempting "to understand how the brain processes language may always lie just beyond the realm of scientific feasibility" (Ingram 2007: 5). Given the MRA, the question arises as to whether this might be more than a matter of fact; that is, is it possible that it is a matter of principle that examining brains does not help in our attempts to understand minds? Chomsky will certainly answer this question in the negative, even though he describes findings from the brain sciences as "something of a curiosity" (Chomsky 2000a: 117). In a 1999 on-line interview Chomsky says:

> I don't see any principled way to distinguish linguistics ... from neurolinguistics, any more than one can distinguish chemistry from physical chemistry in principle. These may be useful distinctions for temporary purposes, but one looks forward to erosion of such boundaries as understanding progresses. (Chomsky 1999)

But if history keeps telling us that we are not making any progress, is it not the case that our lack of progress is perhaps itself principled in some way? Has Descartes been right all along, at least in assigning to the mind a special status? Are we to declare the relation of mind and body a "mystery," rather than a "problem" in Chomsky's (1980a: 6–7) terms? Chomsky and his followers hope the answer is "no" to all these questions, but considerable optimism might be required to sustain such a hope, and, given the MRA, it may not be warranted.

But let us suppose for the sake of argument that progress is in fact being made and that the boundaries between linguistics and neurobiology are overcome. The question that immediately arises, then, is what sort of theory are we to expect? If one insists, as Chomsky (2000a: 77) does, that "the place to look for answers is where they are likely to be found: in the hard sciences," then at best the ultimate theory (whatever that is) will most likely be unintelligible to any but a few scientists working at the cutting edge of a yet-to-be-discovered physics. At worst, the MRA tells us that this "grand unified theory" can hardly be called a "theory" in the usual sense. This is because, as observed earlier, such a theory will most likely include generalizations employing vast open-ended disjunctive terms that cannot be viewed as expressing genuine laws. But if one resists the bias towards the hard sciences, one can see the "positive" side of the MRA, namely that whatever the ultimate theory may be, it will be likely to be essentially *computational* in nature. As Block (1995: 391) has put it: "If we can create machines in our computational image, we will naturally feel that the most compelling theory of the mind is one that is general enough to apply to both them and us, and this will be a computational theory, not a biological theory."

Chomsky (2000c: 22), not unexpectedly, does not regard this as "a wise course." What is important, however, is the (rather extreme) implication he draws from Block's position, namely that such a position implies that "cognitive science is nonnaturalistic, not part of the natural sciences in principle" (Chomsky 2000c: 21). Unless Chomsky holds "nonnaturalistic" and "unbiological" to be synonymous, his interpretation of Block's position does not seem to me to be accurate (cf. Hinzen and Uriagereka 2006: 77–9, who

make similar inaccurate assumptions about functionalism). As noted, Block concedes – perhaps too readily in my opinion – that cooperation between biology and psychology is "vital" for understanding how the brain functions. Moreover, although he maintains that the "right" theory of mind will be computational rather than biological, he nevertheless suggests that a biological theory of the *human* mind would have a "complementary advantage," because it "will encompass us together with our less intelligent biological cousins, and thus provide a different kind of insight into the nature of human intelligence" (Block 1995: 391–2). More importantly, as observed in Section 6.6, functionalist efforts to provide an account of mental phenomena with whatever means are available to natural science constitute a *naturalization* of the mind that is perfectly compatible with Chomsky's naturalistic approach.

Now, although inaccurate, this implication which Chomsky conceived as lurking in Block's views is significant in that it illustrates just how costly Chomsky's position can be. Indeed, as observed in Section 6.6, the cost seems to be so high as to involve the wholesale rejection of modern cognitive psychology on the grounds that it is not "biological" in the proprietorial sense. The tension which we have described earlier between Chomsky's naturalism and Fodor's functionalism crystallises here; where Chomsky seeks to deduce a "neat mind" from a "tidy brain," Fodor insists that we should leave our brains alone (Fodor 1999). Interestingly, Fodor seems to express lack of enthusiasm for the minimalist program itself. When asked what he thinks of Chomsky's optimism in this regard, Fodor (p.c.) replied that he did not see any reason for it; and he adds, "I don't even think there's even any reason to assume a 'well-designed brain' for that matter. I'm not in much sympathy with these functionalist moves of Noam's."[11] By contrast, Chomsky (p.c.) remains as optimistic as ever: "I don't expect to see [the deduction of a neat mind from a well-designed brain] in my lifetime, but I think the day may come, contrary to Jerry's expectations."

Who is going to be right? Only time will tell: but if the past offers insight, I personally would not put too much money on Chomsky's optimism. More importantly for the purposes of this chapter, if the jury is still out on this issue, then the jury is still out on optimal computation providing the basis for "true explanations" in the way Chomsky envisages, or whether it needs to be taken as a *primitive*. If the latter turns out to be the case – as the MRA seems to indicate – it follows that one fundamental aspiration of the minimalist program would be shown to not be realizable; optimal computation would be "contingent" and not explained in any fundamental way. Of course, this conclusion does not in itself call into question the explanatory role of optimal

computation, but it would now have to be taken as a primitive and not reducible to physical "neatness." Perhaps minimalists should start thinking seriously about this conclusion which appears to lead to a new dualism – and indeed there are signs that some have already began to do so (cf. Hinzen and Uriagereka 2006: 77).

7 *Conclusion*

In this book, I have sought to explore the nature of the strong minimalist thesis and develop a detailed evaluation of its plausibility from conceptual and empirical perspectives. In what follows, I give a summary of the book and its main findings and conclude with some general remarks whose main purpose is to look forward on how the reflections in this book may provide insight into how to address some of the major difficulties that we have encountered throughout the previous chapters.

After a short introduction outlying the scope and objectives of the book, I have attempted to clarify some of the misconceptions about the development of Chomskyan linguistics, and suggested that a problem-directed approach – as opposed to the goal-directed approach favored by Boeckx and Hornstein (2010) – may offer a way of avoiding these misconceptions. The discussion then turned to examine the nature of the shift to the minimalist program, where I have maintained that this is not as sharp as some seem to believe. I have shown that the crucial inference in pre-minimalist thinking is from innateness to genetic endowment, an inference the general applicability of which is questioned by the more explicit recognition in minimalism of the role of *non-genetic* nativism. After illustrating the impact of the minimalist program on both the theoretical role of universal grammar (UG) and the design of the Faculty of Language (FL), I have posed the question of what has driven the shift to this program, and suggested that insights from the fields of biology and neuroscience may have lain behind its emergence.

Having taken a broad view of the minimalist program, we turned to the strong minimalist thesis (SMT) and took a closer look at its content. I have argued there for the existence of three different sorts of emphases in Chomsky's work, each linked with a distinct formulation of the SMT at different stages in his writings. The first formulation suggests that nothing is special to language and, I have argued, is clearly incompatible with Chomsky's long-standing claim that something must be special to language or otherwise language acquisition is a miracle. I have tried to resolve this incompatibility by suggesting that it may have been

due to Chomsky's inconsistent use of the phrase "virtual conceptual necessity," a phrase that has led to much confusion which I have attempted to expose and overcome. The second formulation of the SMT emerges out of Chomsky's approach to UG relying on an "imperfection strategy," and suggests that nothing is imperfect in language. Reviewing this strategy, I have drawn attention to some of its major limitations, notably the inconsistent ontological status of the operation Merge and the insensitivity of the SMT to falsifiability. The third and last formulation arises in the context of the "three factors" framework, and it indicates that the SMT should be interpreted as maintaining that language is the result of Merge operating under the conditions of interface-legibility and computational optimality. I concluded with an initial examination of the contrast between Chomsky's linguistic and interdisciplinary discourses and suggested caution in equating them.

Following our clarification of the nature of the SMT, we embarked on a comprehensive evaluation of its plausibility. We have started this evaluation by focusing on one aspect of this minimalist thesis, namely the Merge-only hypothesis. To properly evaluate this hypothesis, I have compared it with the recursion-only hypothesis, arguing that, contrary to a widespread assumption in the literature, the two hypotheses are not equivalent. My argument was based on several indications that suggest, *inter alia*, that the notion of recursion as employed in Hauser *et al*. (2002) and Fitch *et al*. (2005) is much more general and inclusive than Merge. I have drawn several consequences from this result, most notably that the two hypotheses have different empirical content. Taking these consequences into consideration, I have evaluated the Merge-only hypothesis and indicated several conceptual and empirical difficulties with it. Perhaps more importantly, I have also pointed to an inconsistency in Chomsky's views on the language-specificity of Merge, an inconsistency that gives rise to an unacceptable circularity. This circularity stems from the "instruction-to-use-Merge" proposition, an unsatisfactory proposition that Chomsky may have no option but to adopt.

Next, the remaining two aspects of the SMT, interface conditions and optimal computation, have been considered. With respect to the former, I have reviewed some arguments that have been given for and against the appeal to interface-based explanations, concluding that, given the lack of consensus in the language evolution literature concerning the timing of language emergence, and given our ignorance of what the language-external systems are, caution in this respect is necessary. As for the plausibility of interface-based explanation, I have argued that such an explanation suffers from two weaknesses, namely: tautology and teleology. The remaining part of our evaluation has focused on optimal

computation, first examining the extent to which optimal computation may constitute a mode of explanation *sui generis*, and I have offered several arguments suggesting that optimal computation lacks explanatory autonomy. Next, turning attention to several attempts to ground optimal computation in physical law, I have argued that such attempts are abortive and do more harm to the minimalist program than good. Furthermore, in attending to the model of physics on which these attempts have depended, two conclusions have been reached: that the history of minimum principles in physics is not adequately portrayed in the minimalist literature, and that the appeal by some minimalists to these physical principles risks infecting the MP with a mystical view of the natural world that has long been rejected by modern science.

Moving away from the empirical and towards the conceptual, we took the discussion of optimal computation to a different level: the explanatory status of optimal computation from the point of view of the philosophy of mind, beginning with an overview of Chomsky's naturalism, and I have spelled out its connection with Cherniak's non-genomic nativism. The discussion then introduced functionalism and its core argument, the multiple realization argument (MRA), and I have explored the tension that arises between Chomsky's naturalism and Fodorian functionalism. It has been argued that the conflict between the two manifests itself in two ways. First, in Chomsky's optimism and Fodor's scepticism concerning the prospects of unification in science. And, second, in Chomsky's scepticism and Fodor's optimism regarding the prospects of reconciling commonsense conceptions of folk psychology with cognitive science. Focusing on Chomsky's criticism of the functionalist doctrine, I have proposed both that Chomsky's reading of the mind-body problem is misguided, and that the species of dualism he attributes to this philosophical doctrine refers to nothing Cartesian. Turning to the question of whether the concept of "true explanation" within the SMT entails a commitment to something resembling type physicalism, and therefore whether it is subject to the MRA, I have argued that this is the case and, consequently, applying the MRA to optimal computation, leads to the conclusion that there is reason to believe the latter is implausible as a basis for the sort of "principled explanation" that Chomsky is keen to provide. Finally, I have pointed out the implications of this outcome for the biolinguistic approach in general and for the explanatory status of optimal computation in particular, casting doubts on the importance of biology in relation to the study of mind, and raising the prospect of a variety of Minimalism in which optimal computation is viewed as a primitive, physically-irreducible notion.

I believe that the reflections and analyses in this book contain many insights into the strengths and, perhaps more importantly, the weaknesses of the MP.

Major among these are: (1) a clarification of the content of the SMT, especially in connection with the much confused notion of "virtual conceptual necessity"; (2) a proper appreciation of the distinction between a human language and "language as such" and its implications for the foundations of minimalism; (3) a synthesis of Chomsky's linguistic and interdisciplinary discourses, providing insight into their similarities and differences; (4) an assessment of the consequences of the naïve enthusiasm displayed by some linguists when the question of the relationship between principles of language and the laws of physics is raised; and (5) an analysis of the notion of optimal computation from conceptual, empirical, and philosophical perspectives.

Inevitably, several issues were either left unresolved or simply skipped over entirely. But if I were to single out one problem that badly needs to be addressed and resolved, it is the lack of falsifiability of the SMT as discussed in Sections 3.6 and 5.3. Lamentably, minimalists do not seem to take this issue seriously, and when it is brought up, it is often dismissed as a commitment to the notion of "naïve falsification" which overlooks the importance of notions like "simplicity," "uniformity," "harmony," etc., for scientific inquiry. Let us discuss a few examples of this skepticism toward falsifiability before we proceed to suggest a way of overcoming the problem that concerns us here.

Boeckx (2006: 113) complains that "[n]aïve Popperian empiricism pays no attention to such methodological principles like the principle of simplicity." Such a complaint is surprising since Popper (1959 [1935]) dedicated a full chapter to his arguments for rejecting simplicity as a methodological criterion for evaluating theories, and defending falsifiability as a better alternative. Boeckx discusses none of Popper's arguments. Instead, he dismisses the criterion of falsifiability as "an idol of the theatre, an illusionary or fairytale account of reality that obscures our understanding of the latter" (Boeckx 2006: 89). Since he does not supply a single citation from Popper on the question of falsification, he can hardly expect us to take his claim seriously.

Boeckx's unfounded claim is echoed by Hornstein (2013) in his blog post titled "Falsifiability," who adds that "a candidate theory's main problem initially concerns not falsification but verification," that is, the "relevant question is not whether there is counter-evidence but whether '*there is any interesting evidence in its favor*!'" (emphasis and exclamation mark in original). Hornstein seems to be unaware of the fact that verificationism and dogmatism go hand in hand. As Popper (1976: 39) has remarked, a genuine scientific attitude, that is, one that expresses an eagerness to consider evidence contrary to a cherished theory, "was utterly different from the dogmatic attitude which constantly claimed to find 'verifications' for its favourite theories."[1] Indeed, it is ultimately against

dogmatism that Popper proposes his criterion of falsifiability. More specifically, he believed it necessary to protect science by exposing dogmatic pseudo-science, and, acting on this belief, he attempted to tackle "the problem of demarcating science from pseudoscience" (Popper 1976: 42). Now since, as far as I am aware, no adherent of the minimalist program would consider themselves either dogmatic or a pseudoscientist, it is not clear why Popper's falsifiability demarcation criterion should be a target for criticism by some minimalists. This, in view of the absence of specific criteria for falsifying the central thesis of minimalism, may turn out to be almost a rhetorical question.

More importantly, Chomsky himself is dismissive when Popper's criterion is pointed out:

> People talk about Popper's concept of falsification as if it were a meaningful proposal to get rid of a theory: the scientist tries to find refuting evidence and if refuting evidence is found then the theory is given up. But nothing works like that. (Chomsky 2002: 124)

What Chomsky says here is problematic, because it suggests that his position on the issue of falsification is contradictory. In contrast to what he says above, Chomsky had previously made a remark which Popper no doubt would have applauded, namely that a theory which has been refuted must be credited for its ability to allow us to refute it. In his own words, he says:

> To say that we refute the theory is to make a positive comment about it, that is, this theory was presented in a clear enough way so that it was possible to determine whether or not it is correct, or at least on the verge of being correct. It is a merit of a theory to be proved false. Proposals that do not allow such a determination, or the determination of whether or not evidence bears on them, do not have that merit. (Chomsky in Piattelli-Palmarini 1980: 111)

Thus, according to Chomsky's own standard, the SMT (*qua* empirical thesis) does not have "that merit" that he values here, for it does not suggest a clear enough way by which it could be refuted. Indeed, when asked what sort of empirical discovery would refute the SMT, Chomsky (2002: 124) suggests that all linguistic phenomena "appear to refute it," adding the caveat of "whether it is a real refutation." Nothing in his answer suggests what constitutes a "real refutation," which is precisely what the question asks about. Now, it is not my intention here to discern inconsistencies in Chomsky's views. Rather, I make this point simply because it seems to lend support to a conclusion reached in Chapter 3, namely that there appear to be no independent grounds for falsifying

the SMT (see Section 3.6). One aspect of this, which we have explored in Section 5.3, is that interface-based explanations are rendered uninformative, in the sense that the only evidence given for their *explanantia* (i.e. interface constraints) is their own *explananda* (i.e. linguistic phenomena). This aspect of the problem is particularly worrying from an epistemological point of view and, I believe, should initiate a serious research program.

One way to think about the form that such a program might take is to pursue the goal of expanding the evidential basis for interface constraints beyond the linguistic domain. If the external systems evolutionarily predate the language faculty, then we should expect the interface conditions to have analogues in children's prelinguistic cognitive abilities and in animal cognition and communication in general. Ideally, such analogues would provide us with a basis for developing criteria for what should constitute an interface condition on the language faculty, thus allowing us to both establish (extralinguistic) independent grounds for falsifying the SMT and to overcome the circularity that infects current minimalist reasoning (see Section 5.3). Of course, proper examination of these issues will involve detailed assessment of comparative and prelinguistic human cognition, with a view to developing appropriate criteria for the formulation of contentful interface conditions and to emphasise the need for more sophisticated experimental research into non-linguistic properties of the external systems.

Perhaps a concrete example of how this proposal may be implemented would be useful here. Consider the linguistic distinction between count nouns and mass nouns. The distinction is not always clear-cut (*some beer*, *two beers*, etc.), but we need not be concerned about this for our illustrative purposes. Suppose now that our best theory of the language faculty tells us that the count/mass distinction must be represented and we speculate that it can be linked to an ontological distinction between objects that can be individuated and substances. To put it more pointedly, we consider the status of an interface-based explanation for the presence of the linguistic feature [+/−count] that relies on a non-linguistic cognitive distinction between individuated and unindividuated entities. Now, if we are to prevent this justification for the presence of [+/−count] from being circular, we should seek independent evidence for the postulated cognitive distinction; that is, evidence independent of the linguistic feature in question. Such evidence has in fact been provided by several studies on preverbal human infants (e.g. Carey 1994, 2001; Imai and Gentner 1997; Huntley-Fenner *et al.* 2002) and non-human animals (e.g. Mahajan *et al.* 2009).

No doubt the requirement of expanding the evidential basis for interface constraints in this way is an extremely difficult one, but it is made even more

difficult by the tendency of some minimalists to engage in an unrestrained speculation about the nature of interfaces and how to satisfy their legibility conditions. An interdisciplinary approach to cognition, if taken seriously, will not only induce minimalists to be more reserved in their speculations, but will also provide a solid ground for interface-based explanations. Indeed, Chomsky (2000c: 26) himself acknowledges that interface conditions "can no longer simply be taken for granted," and that their investigation should not be limited to the field of linguistics. But I suspect that there are some who, when reading Chomsky's work, would, if pressed, admit to skipping over paragraphs that have no direct bearing on the details of linguistic analysis. Such selectivity not only misses the point of Chomsky's interdisciplinary work, but encourages free-riding on our ignorance of the nature of the interface systems and infecting the minimalist program with specious explanations.

Whatever the outcome of this and other future research striving to understand the nature of human language, the attitude towards the minimalist program I expressed in my introduction to this book remains with me in my conclusion: it is hard to see how any linguist can fail to be interested in it.

Notes

Introduction

1. This opening sentence is borrowed, with minor modification, from W. V. O. Quine's (1950: vii) famous aphorism on Gottlob Frege's revolutionary work in logic: "Logic is an old subject, and since 1879 it has been a great one."
2. On biolinguistics, see Jenkins (2000); see also Boeckx and Grohmann (2007) for a brief historical background.

The minimalist program

1. Chomsky (2007a) is his response to Boden (2006) and Boden (2008) is her reply to this response. I am grateful to Noam Chomsky for drawing my attention to these papers and for other helpful comments.
2. Some, such as Wilks (1971), have tried to strengthen the analogy by suggesting that the "all-and-only" requirement of grammaticality in Chomsky (1957) can be associated with the two logical notions of "completeness" and "decidability," in that the former can be linked to the generation of all and only grammatical sentences, and the latter to a formal criterion by which the generative system can determine the (un)grammaticality of an arbitrary string of English words. However, the analogy is flawed. The logical notion of "completeness" involves an association between the syntactic concept of "derivation" and the semantic concept of "validity," and as far as the theory of phrase structure grammars is concerned, there is nothing corresponding to the notion of "validity" in logical systems. The discussion that follows will elaborate this point.
3. Different kinds of phrase structure grammar (PSG) impose different conditions on X, Y, and the relations between them, e.g. the difference between context-free and context-sensitive PSGs, but we shall not be concerned with this detail here.
4. For instance, Chomsky (1957: 14) maintains that "we are not only interested in particular languages, but also in the general nature of language."
5. That is, a grammar should conform to (i) a condition of generality, and (ii) an external condition of adequacy (Chomsky 1957: 49–50). I will come back to this shortly.
6. I am aware of the reason given by Chomsky (1975b) for not developing cognitive themes and focusing on formal generative issues in *Syntactic Structures*, namely the

importance of securing publication in an intellectual climate dominated by behaviorist ideology. However, this does not mean that *Syntactic Structures* was completely silent on psychological themes, and whether or not we accept Chomsky's explanation, there seems to be no reason to suggest a conceptual discontinuity between Chomsky (1957) and Chomsky (1965).

7. The "Argument from Poverty of the Stimulus" originates in Chomsky (1980a: 34). For a criticism of this argument, see among others Vallauri (2004), Pullum and Scholz (2002). For a response to the latter critique, see Legate and Yang (2002).
8. A statement from *Syntactic Structures* that supplements this passage is this: Any grammar of a language will *project* the finite and somewhat accidental corpus of observed utterances to a set (presumably infinite) of grammatical utterances. In this respect, a grammar mirrors the behavior of a speaker who on the basis of a finite and accidental experience with language, can produce or understand an indefinite number of new sentences (Chomsky 1957: 15, italics in original).
9. Of course, B&H may object to this by saying that, although the cognitive theme was not absolutely absent from Chomsky's early writings, it was only in *Aspects* that this theme became explicit and that "the generative program became indeed 'biolinguistic'" (p. 120). Frankly speaking, I don't understand why *Aspects* should be regarded as the starting point of "biolinguistic" and cognitive themes in Chomsky's work. For such themes were no more explicit in *Aspects* than they were in Chomsky's (1959) famous review of B. F. Skinner's *Verbal Behavior*. As another example, two years before the publication of *Aspects*, Chomsky and Miller (1963: 275) state that the question of "[h]ow an untutored child can so quickly attain full mastery of a language poses a challenging problem for learning theorists." In this same paper, the authors also stress the importance of the question of how language arose in the individual, and they refer briefly to "the genetic issue" (Chomsky and Miller 1963: 272).
10. In fact, what the authors say here is, literally, false of both *Aspects* and *Syntactic Structures*, for there is a mismatch in their use of the terms "internally" and "externally." This mismatch in terminology does not seem accidental, however, for they assert that "[e]xternal justification hinges on outlining explanatory adequate grammatical theories, theories embedded in accounts of how the grammars postulated could have arisen" (Boeckx and Hornstein 2010: 123).
11. Not surprisingly, germs of these formulations can be located in Chomsky (1955), a text on which *Syntactic Structures* was based. Thus Chomsky (1955: 12) says: "It appears then that there are two factors involved in determining the validity of a grammar, the necessity to meet the external conditions of adequacy, and to conform to the general theory."
12. The nearest anyone got to measuring simplicity was to count symbols; consequently, if G_1 involved fewer symbols than G_2, G_1 was preferred. Insofar as progress was made in this regard, it was with respect to partial analyses of specific phenomena and not whole grammars.
13. There are numerous models in the acquisition literature that deal with the setting of parameters in phonology or syntax (e.g. Dresher and Kaye 1990, Gibson and Wexler 1994, Clark and Roberts 1993), but for our purposes this is beside the point. It can be

easily argued, for instance, that all these models involve, in some form or another, an evaluation measure. Thus the idea of associating each parameter with a *cue* that facilitates its ordering with respect to other parameters (Dresher and Kaye 1990) or, for that matter, the concept of a *global fitness metric* that determines the "fittest" grammar among other competing alternatives (Clark and Roberts 1993), can clearly be seen as representing some form of an evaluation measure.

14. It is interesting to note quite a change in Boeckx's (2011) views on this point. He now seems to admit that the P&P model is far from being successful, and that Plato's problem is far from being solved.
15. For an indirect criticism of the way in which parameters are postulated in the literature, see Smith and Law (2009).
16. Simplicity and generality go hand-in-hand and are the cornerstones of any theory construction. Chomsky and Miller make this point clear when they say:

> Since a grammar is a theory of a language and simplicity and generality are primary goals of any theory construction, we shall naturally try to formulate the theory of linguistic structure to accommodate rules that permit the formulation of deeper generalizations. (Chomsky and Miller 1963: 287)

17. This should not be surprising, however, since Chomsky (1995a: 131) seems to be supportive of the idea of restricting parameters to the lexicon (cf. the so-called Borer-Chomsky Conjecture). To be sure, in Chomsky (2001: 26–36) there is a long discussion of the Object Shift Parameter in Scandinavian languages, but it seems to have had little or no subsequent impact.
18. One example comes to mind; Chomsky's use of François Jacob's biological ideas to introduce his P&P approach (see Chomsky 1980a: 67).
19. This is perfectly in line with Chomsky (1955: 714–15): Before we have constructed a linguistic theory we can only have certain vaguely formulated questions to guide the development of the theory. A simple and natural theory, once established, determines the precise formulation of the questions that originally motivated it, and leads to the formulation and resolution of new problems that could not previously have been posed.
20. We will see an example of this in Section 2.6, where we will observe some parallels between Chomsky's approach to language acquisition and his views on the evolution of language.
21. By "core aspects" I mean rules and principles, theoretical artefacts, special mechanisms and whatever has been attributed to UG in the pre-minimalist era with the purpose of contributing to a solution to the problem of language acquisition.
22. For language acquisition, this conclusion translates into an empirical problem of devising a "hypothesis rich enough to account for the acquisition by the child of the grammar that we are, apparently, led to attribute to him" (Chomsky 1968: 86).
23. Reference to "biological necessity" can also be found in Chomsky (1975c: 60) and Chomsky (1980a: 28).
24. For discussion of the differences between Fodor's and Chomsky's notions of "modularity," see Smith (2004: 15–25).
25. As an example, consider: "It is, of course, an empirical question whether the properties of the 'language faculty' are specific to language or are merely a particular case of

much more general mental faculties (or learning strategies)" (Chomsky 1968: 86–7). Open-mindedness is of course consistent with the view that the available scientific evidence in 1980 appeared to support the assumption of modularity.

26. For a detailed discussion of the SMT and its evolution in Chomsky's work, see Chapter 3.
27. Boeckx continues by saying that "minimalism forces researchers to look for uniformity across cognitive systems, and in fact even more broadly across complex systems, at a level that is much more abstract and refined than the Piagetian perspective" (Boeckx 2006: 149). One wonders whether what Boeckx has in mind is that Piaget was nearly right but not quite, and that minimalism has succeeded in ameliorating the Piagetian conception of language. It is noteworthy that, except for the reference to "the Piagetian view," this passage is strikingly similar to what we find in Rizzi (2004: 340): "The Minimalist program naturally leads research to look for uniformity across cognitive systems, and in fact even more broadly, across complex systems."
28. As Chomsky and Miller (1963: 277) put it, "[a]fter all, stupid people learn to talk, but even the brightest apes do not."
29. It must be noted, however, that no explicit mention of Merge was made in Hauser *et al*. (2002). We will return to this point in Section 4.3.
30. As to the phonological system, Chomsky (1980a: 61) poses the following question, the answer to which he regards as an empirical matter: "To what extent, for example, does the organization of sound properly belong to the system of language rather than to other systems? Here there are real and significant empirical questions concerning perceptual categorization and its possible special relation to language." Interestingly, decades later Hauser *et al*. (2002: 1572) refer to "categorical perception" as a typical example in which a trait was originally considered to be uniquely human only for it to be discovered later that it is actually shared with other species.
31. Chomsky's speculation later becomes the topic of research carried out by Fitch and Hauser (2004), in which they attempt to demonstrate that certain species of monkeys are incapable of processing hierarchical phrase structures.
32. This is evident from their preference of the kind of passages they cite to make their point. See, for instance, Pinker and Jackendoff (2005: 4–5).
33. This is particularly clear in Chomsky's (2007b: 5) suggestion that "unbounded Merge is not only a genetically determined property of language, but also unique to it." We shall come back to this in Sections 4.4 and 4.5.
34. The term "standard theory" was actually coined later by Chomsky (1972) to refer to the theory of transformational grammar presented in Katz and Postal (1964) and Chomsky (1965).
35. One may object to this reference to the complexity of UG, especially in light of the earlier claim that the P&P approach has removed the complications associated with traditional rules and constructions from the theory thereby contributing to the simplification of UG. Indeed, Chomsky (1995a: 29) himself acknowledges this simplifying role of the P&P approach. However, it is well to remember that the complexity of UG is *not* exhausted by the specific rules of early generative grammar and in the P&P approach considerable complexity continues to reside in the rich

system of general principles and various modules of GB, and also in the number of levels of representation (see Section 2.5).

36. See Section 3.7 for a more detailed discussion of these three factors.
37. Johnson and Lappin (1999: 133) acknowledge that the MP involves a deduction of the properties of UG, but they see in this deduction a threat to scientific integrity, calling it "a transcendental deduction of universal grammar," for it involves an attempt "to derive the essential properties of human language from a set of first principles that are adopted without regard to evidence." My response to this is that it has been recognized since the rise of twentieth-century physics that the "first principles" of any scientific inquiry need not be directly confirmed by empirical evidence, but rather that the theorems deduced from them are the ones which need to be so confirmed (see on this point Frank 1961: 246–7). On a different matter, I agree with Johnson and Lappin's critique of the idea that there is a plausible analogy between economy in language and economy in physics, especially in relation to the so-called principle of "least action." I will come back to this issue in Section 5.6.
38. I follow here Chomsky's tendency to not discriminate between mind and brain, although I will later call this tendency into question. For more on this, see Section 6.9.
39. On the foundations of optimization and its philosophical roots, see Beightler *et al.* (1979).
40. This task can be thought of as the technical counterpart of the theoretical shift in UG from an *explanans* to an *explanandum* (see Section 2.4).
41. On government and binding theory, see Chomsky (1981). For an introduction to the theory, see Haegeman (1994).
42. While many researchers embrace this argument without reservation (Boeckx 2006, Chandra 2007, Hornstein 2009, just to mention a few examples), others resist it (for example, Brody 2006, Koster 2007).
43. More specifically, agreement takes place between a probe P and a goal G iff (a) G is active, (b) the φ-features on P match the φ-features on G, (c) P c-commands G, and (d) there is no second goal G′ such that (i) G′ satisfies the above conditions on G, and (ii) G′ c-commands G, i.e. G′ is "closer" to P than G is (cf. Chomsky's (1995a: 311) formulation of the minimal link condition).
44. The detailed mechanics of the operation Agree have been the subject of considerable debate among minimalist syntacticians over the last decade. For instance, Hiraiwa (2005) has argued that the operation should be able to target more than one goal leading to the postulation of Multiple Agree while Chomsky (2007b, 2008a) has sought to identify probes with phase-heads in his phasal approach to syntactic derivation. Such details as these will not be relevant to the issues raised in this book.
45. The reference to "partially" here acknowledges that in Chomsky (2007b, 2008a) we find acceptance of a species of movement directly triggered by what Chomsky calls an edge feature that does not presuppose a token of Agree. One might wonder whether this innovation, while increasing descriptive coverage, represents a retreat from minimalist principles. We shall not try to pursue this matter here.
46. More recently, we have the *no-tampering condition* (Chomsky 2008a), which states that merging X and Y must leave the two syntactic objects intact. Put another way,

objects that serve as an input of Merge should come out unchanged in the output of this operation.

47. For a concise history of subjacency and other constraints, see Yoshimoto (2001). For a review of the language-specific properties suggested by Chomsky in his debate with Piaget and their fate within the minimalist program, see Al-Mutairi (2005).
48. In fact, Chomsky did not merely distance himself from this "random mutation" view; he went so far as to ridicule it in his later work (see Section 5.3, p. 116).
49. For a detailed review of this field, see Schwartz (1990).
50. We shall discuss Cherniak's work and its connection to the basic tenets of minimalism in Section 6.3.

The strong minimalist thesis (SMT)

1. However, see Atkinson (2009). Whether the positions we are about to review are properly regarded as substantively distinct is a difficult question. However, I believe that treating them as distinct has virtues in the light it throws on the interpretation of the SMT.
2. See, for example, Boeckx (2006: 73), who suggests that the reason for the modifier "virtual" is that "in science we must always be ready to be proven wrong," and Postal (2003: 599), who thinks that "[t]he hedging with 'virtually' is a . . . clue that something is amiss." Chomsky (1995a: 212, n. 2) merely notes that the necessity involved is "[n]ot literal necessity."
3. Atkinson's (2005a) lectures notes grew out of and developed Atkinson (2000), which itself subsequently appeared as Atkinson (2005b).
4. Alongside the differences that we are now going to discuss, there is a very general matter that also points to a fundamental difference between the strong minimalist thesis and laws in the special sciences, namely that there is no sense in which the former, with or without its qualification, *is* a law; rather, it is perhaps better thought of as constituting a high-level constraint on the status of statements in linguistic theory as explanatory (or not).
5. Fodor's (1991: 22) proposal could not be more urgent: "[W]e should do what we can to provide a clear account of the truth conditions of hedged laws."
6. Of course this is a bit of handwaving but it can be made precise. For instance, the language-related gene FOXP2 (though it remains controversial) may offer a way of *envisaging* what reference to "the vocabulary of genetics" might involve.
7. I am grateful to Martin Atkinson for drawing my attention to this passage.
8. Perhaps it is desirable to qualify this statement to avoid some objections that have been raised against this conception of Merge (cf. Postal 2003). Thus, we may say that the language system we are referring to is one which we take to be derivational, and in which the lexicon and the computational system are separated from each other (cf. Atkinson 2005b: 213–14).
9. I should add, however, that – owing, perhaps, to Chomsky's (1995b [1994]) failure to be more explicit about the third dimension of conceptual necessity, and his positive attitude in Chomsky (2000b) to Fodor's language of thought hypothesis – Atkinson (2005a) attempted to account for the presence of Merge in terms of legibility conditions.

10. Cf. Berwick and Chomsky (2011: 32), who maintain that there is "a conflict between computational efficiency and interpretive-communicative efficiency," and argue that "languages resolve the conflict in favor of computational efficiency," suggesting "that language evolved as an instrument of internal thought, with externalization a secondary process."
11. As mentioned in several places above (especially in Sections 2.3 and 2.6), Chomsky subscribes to the view that Merge/recursion may be the only property that is unique to humans and to language. This view is further explored in the next section and throughout Chapter 4.
12. For recent discussion of this question, see Atkinson and Al-Mutairi (2012).
13. It is well to remember, however, that Chomsky himself is more cautious in this regard. Contrary to what Grohmann (2006) seems to believe, Chomsky does not suggest an immediate rejection of what might escape the force of the SMT. Rather, he proposes a "close examination, to see if [the imperfection] is really justified" (Chomsky 2008a: 135). However, see the next paragraph of the text.
14. If we take into account Chomsky's suggestion that language imperfections arise from the mapping between the syntax and phonology (see Section 3.5), then my statement above concerns only imperfections arising from interpretation (i.e. from the mapping between the syntax and the semantics).
15. It seems to me that this extension of the content of Factor I is precisely what Yang (2010) argues for when he says "that not asking UG to do too much doesn't mean asking UG to do too little" (Yang 2010: 1174).
16. There are, of course, many who do not share Chomsky's biolinguistic perspective on language and, therefore, would probably see a false dichotomy here (*viz*. that either Factor I is non-empty or language acquisition is a "miracle"). Thus, they may be prepared to expose this fallacy by challenging the inference from "language is an exclusive human property" to "Factor I must be non-empty," perhaps on the grounds that the relevant differences between humans and animals can be explained from a non-genetic point of view. But this is an issue that does not concern us here.
17. We explore this hypothesis in more detail in Section 4.2.
18. Chomsky (2005: 11–12) speculates that a slight mutation might have caused a rewiring of the human brain and, therefore, provided the operation Merge.
19. However, there is a lack of congruence between Fitch *et al*.'s suggestion that FLN may be empty and Chomsky's assertion that Factor I must be non-empty. This is perhaps sufficient to indicate the non-identity of these two constructs, at least in principle.

The SMT in an evolutionary context

1. Kinsella (2009) is an exception, for she explicitly attempts to differentiate between recursion and Merge, although she ends up by equating the two notions. We will come back to this in Section 4.3.
2. In response to this charge, Pinker and Bloom (1990: 711) maintain that "adaptationist proposals are testable in principle and in practice," suggesting that by "[s]upplementing the criterion of complex design, one can determine whether putatively adaptive structures are correlated with the ecological conditions that make them useful." As

we shall see in the next chapter (Section 5.3), the charge of telling "just-so-stories" is as applicable to minimalist explanation as it is to Darwinian explanation.

3. Of course, all recursion is syntactic in the sense that it depends upon the *form* of the objects to which it applies. However, syntactic recursion, as understood here, has the additional property of being responsible for the *construction* of the infinite array of discrete expressions.
4. For discussion of Pirahã and whether it exhibits recursion, see Everett (2009); Nevins *et al*. (2009).
5. Fitch *et al*. (2005: 179) accuse their opponents of misinterpreting their hypothesis by blurring the distinction between FLN and FLB. They emphasize that their hypothesis does not concern language as a whole, but only FLN. Accordingly, the authors dismiss many of the arguments in Pinker and Jackendoff (2005) as irrelevant to their core hypothesis. However, this is hardly a counter-argument against Pinker and Jackendoff's position, for what Fitch *et al*. consider "irrelevant" to their hypothesis is precisely what is at stake in the debate. To be sure, the arguments of their opponents can only be irrelevant if both sides of the debate agree on what should be included in FLN, but they obviously do not, this being the reason why there is a debate in the first place.
6. In Samuels (2011: 33) this rather strong statement is significantly qualified to read: "The relation of Hauser *et al*.'s claims to the Minimalist Program is somewhat controversial . . . "
7. That the first statement is more indicative of the position Hauser *et al*. support becomes clear in the authors' second article. There, Fitch *et al*. (2005: 189–90) say:

   > FLN may include more than the computations subserving recursion and mappings to the interfaces to SM and CI, as we suggest in several places in [Hauser *et al*. 2002]. If so, our Hypothesis 3 can simply be restated as specific to the recursive machinery and associated mappings, rather than FLN in full, and all the same considerations will apply. But in either case our hypothesis concerns a specific subset of linguistic mechanisms, not "language" in a broad sense.

8. Cf. Scheer (2004: xliv), who interprets the hypothesis of Hauser *et al*. as saying that "FLN is made of Merge and Phase."
9. The fact that Chomsky entertains such a speculation may explain why he appears to have seen no anomaly in regarding Merge both as an indispensable mechanism in any language-like system and as a "lucky" event in the course of evolution (that is, to put it in terms that we have been led to employ in Section 3.4, as both a "perfection" and an "imperfection"). For if it could be maintained that Merge-like operations in domains other than language are all derivative from *linguistic* Merge, then there should be no incongruity in asserting the general and indispensable presence of Merge in any language-like system while at the same time recognizing its language-specificity and path-dependent evolutionary history.
10. Fitch *et al*. (2005: 203) make a similar point by asserting that "[t]here are no unambiguous demonstrations of recursion in other human cognitive domains, with the only clear exceptions (mathematical formulas, computer programming) being clearly dependent upon language."

11. Given any real number, say 1, no matter what value we choose for the real number immediately succeeding 1, there will be infinitely many real numbers between these two numbers. This is due to what mathematicians call "the density of real numbers," a property that indicates the impossibility of counting the real numbers.
12. Hinzen's position on the relationship between language and arithmetic is unclear. On the one hand, he seems to suggest that syntactic Merge and mathematical Merge are two special cases of a more general recursive mechanism underlying both language and arithmetic. Yet, on the other hand, he subscribes to Chomsky's view that arithmetic is an evolutionary offshoot of language. Nowhere in his article does he provide any reason as to why the instantiation of Merge in one domain should be an evolutionary offshoot of its instantiation in another domain, rather than the other way round.
13. Hinzen's view goes further than this, contending not just that Merge alone is insufficient to yield the richness of language, but also that reliance on this operation alone has misled minimalists to shift the burden of explanation from syntax to the interfaces, which he believes results in vacuous explanations. We will return to discuss this feature of Hinzen's position in Section 5.2.
14. As Martin Atkinson (p.c.) has pointed out, on the assumption that we can observe in animal behavior signs of recursion (say, chimpanzees who can "embed" one plan inside another to form a complex plan), we have no justification for saying that this behavior is bounded; "the fact that they never manage more than *x* degrees of embedding in practice is surely irrelevant as no human has even managed to produce a linguistic structure with greater than *y* degrees of embedding for suitably large *y*."
15. Cf. Hauser (2009: 49) and Fitch (2010: 22).

The SMT as an explanatory thesis

1. It is perhaps worth remembering here the reason why Chomsky (2001: 1) believes that the "SMT cannot be seriously entertained." As indicated at the end of Section 3.3, he thinks that it would be too extraordinary for a biological system such as language to be completely efficient in using its resources to link sound and meaning.
2. Whether such reduction has been well founded is an issue we consider in Sections 5.3 and 5.4.
3. Chomsky seems to have tried hard to avoid the teleology that is implied in the notion of "look-ahead." He says, for instance, that "[t]hough motivated at the interface, interpretability of a feature is an inherent property that is accessible throughout the derivation" (Chomsky 2001: 4). But this looks like an *ad hoc* stipulation for which there is no evidence – indeed, what evidence for this might comprise is totally opaque.
4. Indeed, adaptive advantages have been postulated for both functions. As observed in Section 4.2, Pinker and Bloom (1990: 714) argue that communication of propositional structures would have adaptive advantage as a means to communicate one's knowledge to others. On the other hand, Chomsky (2002: 148) claims that the private use of language by a creature could have enormous adaptive advantages, as it "could think, could articulate to itself its thoughts, could plan, could sharpen and develop thinking as we do in inner speech, which has a big effect on our lives."

5. Uriagereka tells us that his 1998 book, *Rhyme and Reason*, took from "Fukui's (1996) original paper ... the provocative thought that comparing derivational alternatives resembles Least Action in physics" (Uriagereka 2000: 869). Freidin and Vergnaud (2001) make claims similar to those of Fukui. We shall return to discuss these views in Section 5.7.
6. Freidin and Vergnaud also seek to establish a methodological parallelism between linguistics and physics by numerous citations and references to works of prominent physicists, including Einstein, Dirac, and Feynman. After reading Freidin and Vergnaud's interpretations of these citations, one is led to wonder whether some minimalists have actually inherited from Einstein his search for a unified theory, from Dirac his mathematical methods, and from Feynman his sense of the glory of science. This is because the name of Einstein is invoked, *inter alia*, to make the point that minimalism may turn out to be premature "in much the same way that Einstein's search for a unified field theory was premature" (Freidin and Vergnaud 2001: 650, n. 21). Dirac's authority is cited in the claim that "the recent developments within MP must be viewed ... as Dirac's mathematical procedure (method) at work within linguistics" (Freidin and Vergnaud 2001: 647). As for Feynman, he is presented as someone who would have endorsed both the methodology and substance of minimalism, someone whose view on Fermat's principle of least time, according to the authors (p. 651), "extends to all economy considerations developed in the natural sciences" – including, one gathers, linguistics.
7. In the *Timaeus* (as translated by Jowett 1961), which presents Plato's cosmogony, and in which God is conceived of as a geometer or "architect of the world," we find an attempt to derive the concrete "four elements" conception of nature from Pythagorean numerical abstractions (see *Timaeus* 53b). Elsewhere in the same dialogue, Plato states that "God desired that all things should be good and nothing bad," and that "out of disorder he brought order, considering that this was in every way better than the other" (Plat. *Tim*. 30a). This metaphysical optimism will later find its highest expression in Leibniz's "best of all possible worlds." As we shall see in Section 6.3, the computational neurologist Christopher Cherniak yields echoes of Leibniz when he talks of "the best of all possible *brains*," a concept to which Chomsky adverts favorably. It is tempting to suggest that not only Plato's epistemology but also his metaphysics can be discerned in one form or another in Chomsky's work.
8. Kepler's poem appears in *Mysterium Cosmographicum*, trans. A. M. Duncan (1981).
9. Yourgrau and Mandelstam (1960), in their history of so-called variational principles (of which minimum principles are special cases), argue that Aristotelian simplicity should be distinguished from Occam's Razor because Aristotle "held that nature possess an immanent tendency to simplicity, whereas Ockham demanded that in *describing* nature one should avoid unnecessary complications" (1960: 6–7). It should be noted though that, in his *Analytica Posteriora*, Aristotle provides a formulation of how to compare theoretical proposals similar to Occam's when he says: "We may assume the superiority *ceteris paribus* of the demonstration which derives from fewer postulates or hypotheses" (Arist. *APo*. 1.25.86a33-34, in McKeon 1941). It is more accurate, then, to say that simplicity as understood by Aristotle has both

a methodological and a substantive character, a particularly revealing characterization in the context of the minimalist program.

10. For this approach I refer the reader to Feynman (1985), which is fairly readable for the non-specialist. For linguists, Johnson and Lappin (1999: 129–31) provide a useful and clear exposition of the main ideas of Feynman's book.
11. One such case is that in which a source of light is positioned in the centre of an ellipsoidal mirror. For a simple illustration of this case, see Nahin (2004: 133).
12. Berdichevsky (2009: xvii) defines a *variational principle* as "an assertion stating that some quantity defined for all possible processes reaches its minimum (or maximum, or stationary) value for the real process. Variational principles yield the equations governing the real processes." For an introduction to the subject, see Sagan (1969).
13. This theological inclination made him an easy target for Voltaire's satirical pen (see, for instance, Hankins 1985: 36).

Optimal computation and multiple realization

1. An earlier version of this chapter included a detailed discussion of the "Galilean style" in linguistics and the problems it raised; reasons of space and questionable direct relevance led to its omission from the final version. However, see Al-Mutairi (2007, 2008).
2. In this chapter, whenever we refer to functionalism, we mean Putnamian/Fodorian functionalism and not other varieties such as *analytical functionalism* (Armstrong 1968; Lewis 1972).
3. Putnam, an early supporter of functionalism, was later critical of the doctrine. Interestingly, Putnam (1988) shows how the multiple realization argument (see Section 6.5) can be applied to functionalism itself or, more specifically, to computational states. It should be noted, however, that Putnam's criticism in this case does not undermine the significance of the argument as a refutation of reductionism, an issue of central concern later in this chapter. For a criticism of functionalism in general and the multiple realization argument in particular, see, among others, Wilson (1985), Kim (1998), Sober (1999), Batterman (2000), and Shapiro (2004).
4. Of course, a behaviorist might respond to this by trying to incorporate into his hypothetical statement a *ceteris paribus* clause as a way of restricting the conditions under which his proposed causal model holds, e.g. "Other things being equal, if there were an apple available, then John would eat it, and there was an apple available." But this won't work, for "the phrase 'other things being equal' is behavioristically illicit, because it can only be filled in with references to *other mental states*" (Block 2004 [1980]: 189, italics in original). Down this route, an infinite regress beckons. Observe, further, that the necessity for *ceteris paribus* clauses extends beyond the doctrine of logical behaviorism, infecting the special sciences in general (see Section 3.3 for relevant discussion). There is something in this that creates further difficulties for the logical behaviorist: the general assumption is that the content of *ceteris paribus* clauses can be spelled out in the vocabulary of a more basic science. Thus, given the example above, it is tempting to conceive of the study of mind as

being more basic than the study of behavior, an uncomfortable perspective for the logical behaviorist, to say the least.

5. On the history of the concept of neuronal plasticity and its significance in neuroscience, see Stahnisch and Nitsch (2002).
6. The multiple realization argument is one of the most powerful arguments against reductionism. Kim (1992: 1) describes it as "part of today's conventional wisdom in philosophy of mind," and LePore and Loewer (1989: 179) refer to it as "practically received wisdom among philosophers of mind." In fact, the significance of the argument extends beyond the sphere of philosophy of mind. For instance, Hull (1972) and Kitcher (1984) have applied it to the biological sciences in arguing against the reduction of Mendelian genetics to molecular genetics.
7. Chomsky (2003b: 268) makes a distinction between ethnoscience and other naturalistic enterprises as follows: "Ethnoscience is an empirical pursuit that seeks to discover how commonsense understanding, in various cultures and settings, seeks to make some sense of how the world works; different naturalistic enterprises seek to discover how the world actually works."
8. For a history of this notion, see Hesse (1961).
9. Of course, one might argue that neurological evidence can be considered relevant if we suppose that token mental states are identical to token brain states. However, notice that, in this case, it is unlikely that any proposed explanation of cognitive functions would achieve the same generality as whatever psychological formulation we are starting from. This point will be clarified in detail in Section 6.8.
10. The argument that follows was suggested to me by Martin Atkinson, and it was in fact one of the main reasons why I wrote this book.
11. Of course, Fodor is using the word "functionalist" *not* in the philosophy of mind sense but rather in the sense we have seen in previous chapters, namely the sense in which the language faculty has an internal *function* with respect to the external systems with which it interacts.

Conclusion

1. For sharp criticism of the "dogmatic attitude" referred to here, see Popper (1945, 1961).

References

Al-Mutairi, F. R. (2005), "From Language-specific to General Principles: the Development of Chomskyan Thought," University of Essex, MA dissertation.

(2007), "Linguistics and the Natural Sciences," University of Essex, Mres dissertation.

(2008), "Minimalism and the Road towards the Natural Sciences," in Jenset, G. B., Heggelund, O., and Cardona, M. D. (eds.), *Linguistics in the Making: Selected Papers from the Second Scandinavian PhD Conference in Linguistics*, 97–114.

(2011), "The Strong Minimalist Thesis: its Nature and Plausibility," University of Essex, PhD dissertation.

Armstrong, D. M. (1968), *A Materialist Theory of Mind*, London: Routledge and Kegan Paul.

Atkinson, M. (2000), "Minimalist Visions," *Essex Research Reports in Linguistics* **34**: 1–55.

(2005a), Unpublished lecture notes, University of Essex, UK.

(2005b), "Minimalist Visions," in Saleemi, A. P., Bohn, O.-S., and Gjedde, A. (eds.), *Search of a Language for the Mind-Brain*, Denmark: Aarhus University Press, 200–284.

(2009), "Recursive Musings," Unpublished ms., University of Essex, UK.

Atkinson, M. and Al-Mutairi, F. R. (2012), "UG or not UG: Where is Recursion?", *Iberia* **4**(1): 35–60.

Ayala, F. J. (2007), *Darwin's Gift to Science and Religion*, Washington, DC: Joseph Henry Press.

Ayoun, D. (2003), *Parameter Setting in Language Acquisition*, London: Continuum.

Batterman, R. W. (2000), "Multiple Realizability and Universality," *British Society for the Philosophy of Science* **51**: 115–45.

Beightler, C. S., Philips, D. T., and Wilde, D. J. (1979), *Foundations of Optimization*, Englewood Cliffs, NJ: Prentice-Hall.

Bell, A. E. (1947), *Christian Huygens and the Development of Science in the Seventeenth Century A. E*, London: Edward Arnold and Co.

Berdichevsky, V. L. (2009), *Variational Principles of Continuum Mechanics*, New York: Springer.

Berwick, R. C. (1998), "Language Evolution and the Minimalist Program: the Origins of Syntax," in Hurford, J. R., Studdert-Kennedy, M., and Knight, C. (eds.), *Approaches to the Evolution of Language*, Cambridge University Press, 320–40.

Berwick, R. C. and Chomsky, N. (2011), "The Biolinguistic Program: the Current State of its Development," in Di Sciullo, A. M. and Boeckx, C. (eds.), *The Biolinguistic*

Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty, Oxford University Press, 19–41.

Bickerton, D. (2003), "Symbol and Structure: a Comprehensive Framework for Language Evolution," in Christiansen, M. H. and Kirby, S. (eds.), *Language Evolution*, New York: Oxford University Press, 77–93.

Bickle, J. (2008), "Multiple Realizability," in Zalta, E. N. (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), http://plato.stanford.edu/archives/fall2008/entries/multiple-realizability.

Block, N. (2004[1980]), "What is Functionalism?", in Heil, J. (ed.), *Philosophy of Mind: a Guide and Anthology*, Oxford University Press, 183–99.

(1995), "The Mind as the Software of the Brain," in Smith, E. E. and Osherson, D. N. (eds.), *An Invitation to Cognitive Science: Thinking* (2nd edn), Vol. **III**, Cambridge, MA: MIT Press, 377–425.

Block, N. and Fodor, J. (1972), "What Psychological States are Not," *Philosophical Review* **81**: 159–81.

Bloom, P. and Markson, L. (1998), "Capacities Underlying Word Learning," *Trends in Cognitive Sciences* **2**: 67–73.

Boden, M. A. (2006), *Mind as Machine: A History of Cognitive Science*, 2 vols., Oxford University Press.

(2008), "Odd Man Out: Reply to Reviewers," *Artificial Intelligence* **172**: 1944–64.

Boeckx, C. (2006), *Linguistic Minimalism*, Oxford University Press.

(2011), "Approaching Parameters from Below," in Di Sciullo, A. M. and Boeckx, C. (eds.), *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, Oxford University Press, 205–21.

Boeckx, C. and Grohmann, K. K. (2007), "The Biolinguistics Manifesto," *Biolinguistics* **1**: 1–8.

Boeckx, C. and Hornstein, N. (2010), "The Varying Aims of Linguistic Theory," in Bricmont, J. and Franck, J. (eds.), *Chomsky Notebook*, New York: Columbia University Press, 115–41.

Boeckx, C., and Piattelli-Palmarini, M. (2005), "Language as a Natural Object: Linguistics as a Natural Science," *The Linguistic Review* **22**: 447–66.

Bouchard, D. (2006), "Beyond Descriptivism: Exaptation and Linguistic Explanation," in Nishida, C. and Montreuil, J. P. (eds.), *New Perspectives on Romance Linguistics*, John Benjamins B. V., Vol. **I**, 27–50.

Bricmont, J. and J. Franck (eds.) (2010), *Chomsky Notebook*, Columbia University Press: New York.

Brody, M. (2006), "'Move' in Syntax: Logically necessary or Undefinable in the Best Case?", *Lingua* **116**: 1553–60.

Burtt, E. A. (1924), *The Metaphysical Foundations of Physical Science: A Historical and Critical Essay*, London: Routledge and Kegan Paul.

Cangelosi, A., Smith, A., and Smith, K. (eds.) (2006), *The Evolution of Language: Proceedings of the 6th International Conference (EVOLANG6)*, Singapore, World Scientific.

Carey, S. (1994), "Does Learning a Language Require the Child to Reconceptualize the World?", *Lingua* **92**: 143–67.

(2001), "Whorf versus Continuity Theories: Bringing Data to Bear on the Debate," in Bowerman, M. and Levinson, S. (eds.), *Language Acquisition and Conceptual Development*, Cambridge University Press, 185–214.

Carnap, R. (1937), *The Logical Syntax of Language*, London: K. Paul, Trench, Trubner and Co.

Carnie, A. and Denton-Mendoza, N. (2003), "Functionalism is/n't Formalism: an Interactive Review of Darnell *et al*. (1999)," *Journal of Linguistics* **39**: 373–89.

Cela-Conde, C. J. and Marty, G. (1998), "Noam Chomsky's Minimalist Program and the Philosophy of Mind: an Interview," *Syntax* **1**: 19–36.

Chabay, R. W. and Sherwood, B. A. (2002), *Matter and Interactions*, New York: John Wiley and Sons.

Chandra, P. (2007), "(DIS)AGREE: Movement and Agreement Reconsidered," University of Maryland, PhD dissertation.

Cherniak, C. (1981), "Minimal Rationality," *Mind* **90**: 161–83.

(1990), "The Bounded Brain: Toward Quantitative Neuroanatomy," *Journal of Cognitive Neuroscience* **2**: 58–68.

(1992), "Local Optimization of Neuron Arbors," *Biological Cybernetics* **66**: 503–10.

(1994), "Philosophy and Computational Neuroanatomy," *Philosophical Studies* **73**: 89–107.

(1995), "Neural Component Placement," *Trends in Neuroscience* **18**: 522–7.

(2005), "Innateness and Brain-Wiring Optimization: Non-Genomic Nativism," in Zilhao, A. (ed.), *Cognition, Evolution, and Rationality*, London: Routledge, 103–12.

Cherniak, C., Changizi, M., and Kang, D. W. (1999), "Large-scale Optimization of Neuron Arbors," *Physical Review E* **59**: 6001–9.

Cherniak, C., Mokhtarzada, Z., and Nodelman, U. (2002), "Optimal-wiring Models of Neuroanatomy," in Ascoli, G. (ed.), *Computational Neuroanatomy: Principles and Methods*, Totowa, NJ: Humana.

Cherniak, C., Mokhtarzada, Z., and Rodriguez-Esteban, R. (2006), "Neural Wiring Optimization," in Kaas, J. (ed.), *Evolution of Nervous Systems*, Vol. **I**, Academic Press, 269–72.

Cherniak, C., Mokhtarzada, Z., Rodriguez-Esteban, R., and Changizi, K. (2004), "Global Optimization of Cerebral Cortex Layout," *Proceedings of the National Academy of Science* **101**: 1081–6.

Chomsky, N. (1955), *The Logical Structure of Linguistic Theory*, MIT Humanities Library, Microfilm.

(1956), "Three Models for the Description of Language," in *IRE Transactions on Information Theory*, Vol. **IT-2**, no. 3, 113–24.

(1957), *Syntactic Structures*, The Hague: Mouton.

(1959), "Review of Skinner's *Verbal Behaviour*," *Language* **35**: 26–58.

(1965), *Aspects of the Theory of Syntax*, Cambridge, MA: MIT Press.

(1966), *Cartesian Linguistics*, New York: Harper and Row.

(1968), *Language and Mind*, New York: Harcourt Brace and World, Inc.

(1971), "The Case Against B. F. Skinner," *The New York Review of Books*, December 30.

(1972), *Studies on Semantics in Generative Grammar*, The Hague: Mouton.

(1973), "Conditions on Transformations," in Kiparsky, P. and Peters, S. (eds.), *Festschrift for Morris Halle*, The Hague: Mouton.

(1975a), *Reflections on Language*, New York: Pantheon Books.
(1975b), *The Logical Structure of Linguistic Theory*, New York: Plenum Press.
(1975c), "Language as a Key to Human Nature and Society," in Otero, C. P. (ed.), *Chomsky on Democracy and Education*, RoutledgeFalmer, 58–64.
(1978), "On the Biological Basis of Language Capacities," in Miller, G. A. and Lenneberg, E. (eds.), *Psychology and Biology of Language and Thought: Essays in Honor of Eric Lenneberg*, New York: Academic Press, XII, 199–220.
(1980a), *Rules and Representations*, New York: Columbia University Press and Oxford: Basil Blackwell Publisher.
(1980b), "On Cognitive Structures and their Development: a Reply to Piaget," in Piattelli-Palmarini, M. (ed.), *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*, New York: Routledge and Kegan Paul, 35–52.
(1981), *Lectures on Government and Binding: The Pisa Lectures*, Holland: Foris Publications.
(1986), *Knowledge of Language: its Nature, Origin, and Use*, New York: Praeger Publishers.
(1988), *Language and the Problems of Knowledge: The Managua Lectures*, Cambridge, MA, and London: MIT Press.
(1991), "Linguistics and Cognitive Science: Problems and Mysteries," in Kasher, A. (ed.), *The Chomskyan Turn*, Cambridge, MA: Basil Blackwell, 26–53.
(1995a), *The Minimalist Program*, Cambridge, MA: MIT Press.
(1995b[1994]), "Bare Phrase Structure," in Webelhuth, G. (ed.), *Government and Binding and the Minimalist Program: Principles and Parameters in Syntactic Theory*, Oxford: Blackwell, 383–439 (originally distributed as *MIT Occasional Papers in Linguistics* 5, 1994).
(1996), *Powers and Prospects: Reflections on Human Nature and the Social Order*, London: Pluto Press.
(1999), "Interview with Noam Chomsky: on the Nature of Pragmatics and Related Issues," available at: http://pi.sfedu.ru/rspu/structure/university_departments/chairs/russian_language/library/Homskiy.doc.
(2000a), *New Horizons in the Study of Language and Mind*, Cambridge University Press.
(2000b), "Minimalist Inquiries: the Framework," in Martin, R., Michaels, D., and Uriagereka, J. (eds.), *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, Cambridge, MA: MIT Press, 89–155.
(2000c), "Linguistic and Brain Science," in Marantz, A., Miyashita, Y., and O´Neil, W. (eds.), *Image, Language, Brain*, Cambridge, MA: MIT Press, 13–28.
(2001), "Derivation by Phase," in Kenstowicz, M. (ed.), *Ken Hale: a Life in Language*, Cambridge, MA: MIT Press, 1–52.
(2002), *On Language and Nature*, Cambridge University Press.
(2003a), "Reply to Lycan," in Antony, L. M. and Hornstein, N. (eds.), *Chomsky and his Critics*, Blackwell Publishing, 255–63.
(2003b), "Reply to Egan," in Antony, L. M. and Hornstein, N. (eds.), *Chomsky and his Critics*, Blackwell Publishing, 268–74.
(2004a), "Beyond Explanatory Adequacy," in Belletti, A. (ed.), *Structures and Beyond*, Oxford University Press, 104–31.

(2004b), *The Generative Enterprise Revisited: Discussions with Riny Huybregts, Henk van Riemsdijk, Naoki Fukui and Mihoko Zushi*, Berlin and New York: Mouton de Gruyter.
(2005), "Three Factors in Language Design," ms., Cambridge, MA: MIT Press.
(2007a), Review of Boden (2006), *Artificial Intelligence* **171**: 1094–103.
(2007b), "Approaching UG from Below," ms., Cambridge, MA: MIT Press.
(2008a), "On Phases," in Freidin, R., Otero, C. P., and Zubizarreta, M.-L. (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, Cambridge, MA: MIT Press, 133–66.
(2008b), Public lecture at Google Talk, 25 April, available at: www.youtube.com/watch?v=rnLWSC5p1XE.
(2010), "Some Simple Evo Devo Theses: How True Might they be for Language?", in Larson, R., Deprez, V., and Yamakido, H. (eds.), *The Evolution of Human Language: Biolinguistic Perspectives*, Cambridge University Press, 45–62.
(2013), "Problems of projection," *Lingua*: **130**: 33–49.
Chomsky, N. and Lasnik, H. (1993), "The Theory of Principles and Parameters," in Jacobs, J., Stechow, A., Sternefeld, W., and Vennemann, T. (eds.), *Syntax: an International Handbook of Contemporary Research*, Vol. **I**, Berlin: Walter de Gruyter, 506–69 (reprinted in Chomsky 1995a).
Chomsky, N. and Miller, G. A. (1963), "Introduction to the Formal Analysis of Natural Languages," in Luce, R. D., Bush, R. R., and Galanter, E. (eds.), *Handbook of Mathematical Psychology*, Vol. **II**, New York: John Wiley and Sons, 269–321.
Chomsky, N., Hauser, M. D., and Fitch, W. T. (2004), "Appendix: the Minimalist Program," unpublished ms., Harvard University.
Christiansen, M. H. and Kirby, S. (eds.) (2003), *Language Evolution*, New York: Oxford University Press.
Clark, R. and Roberts, I. (1993), "A Computational Model of Language Learnability and Language Change," *Linguistic Inquiry* **24**: 299–345.
Cowper, E. A. (1992), *A Concise Introduction to Syntactic Theory: the Government-binding Approach*, University of Chicago Press.
Darwin, C. (2003[1859]), *On the Origin of Species: a Facsimile of the First Edition*, Cambridge, MA, and London: Harvard University Press.
Dresher, B. E. and Kaye, J. D. (1990), "A Computational Learning Model for Metrical Phonology," *Cognition* **34**: 137–95.
Dugas, R. (1988), *A History of Mechanics*, Dover Science Books.
Duncan, A. M. (trans.) (1981), *Kepler: Mysterium Cosmographicum: Secret of the Universe*, New York: Abaris.
Eldredge, N. and Gould, S. J. (1972), "Punctuated Equilibria: an Alternative to Phyletic Gradualism," in Schopf, T. J. M. (ed.), *Models in Paleobiology*, San Francisco: Freeman, Cooper and Company, 82–115.
Epstein, S. D. and Hornstein, N. (1999), *Working Minimalism*, Cambridge, MA: MIT Press.
Epstein, S. D. and Seely, T. D. (2002), "Introduction: on the Quest for Explanation," in Epstein, S. D. and Seely, T. D. (eds.), *Derivation and Explanation in the Minimalist Program*, Blackwell, 1–18.

Evans, N. and Levinson, S. C. (2009), "The Myth of Language Universals: Language Diversity and its Importance for Cognitive Science," *Behavioural and Brain Sciences* **32**: 429–92.

Everett, D. (2005), "Biology and Language: a Consideration of Alternatives," *Journal of Linguistics* **41**: 157–75.

(2009), "Pirahã Culture and Grammar: a Response to Some Criticisms," *Language* **85**(2): 405–42.

Feynman, R. P. (1985), *QED: the Strange Theory of Light and Matter*, Princeton and Oxford: Princeton University Press.

Feynman, R. P., Leighton, R. B., and Sands, M. (1964), *The Feynman Lectures on Physics: Mainly Electromagnetism and Matter*, Vol. **II**, Boston, MA: Addison-Wesley.

Fitch, W. T. (2010), *The Evolution of Language*, Cambridge University Press.

Fitch, W. T. and Hauser, M. D. (2004), "Computational Constraints on Syntactic Processing in a Nonhuman Primate," *Science* **303**: 377–80.

Fitch, W. T., Hauser, M. D., and Chomsky, N. (2005), "The Evolution of the Language Faculty: Clarification and Implications," *Cognition* **97**: 179–210.

Fodor, J. (1974), "Special Sciences: or the Disunity of Science as a Working Hypothesis," *Synthese* **28**: 97–115.

(1975), *The Language of Thought*, New York: Thomas Y. Crowell.

(1983), *Modularity of Mind: an Essay on Faculty Psychology*, Cambridge, MA: MIT Press.

(1987), *Psychosemantics: the Problem of Meaning in the Philosophy of Mind*, Cambridge, MA: MIT Press.

(1991), "You can Fool Some of the People All of the Time, Everything Else Being Equal; Hedged Laws and Psychological Explanations," *Mind* **100**: 19–34.

(1997), "Special Sciences: Still Autonomous after All These Years," *Philosophical Perspectives* **11**: 149–63.

(1999), "Let Your Brain Alone," *London Review of Books*, Vol. **XXI**, no. 19, September.

(2004[1981]), "The Mind-Body Problem," in Heil, J. (ed.), *Philosophy of Mind: a Guide and Anthology*, Oxford University Press, 168–82.

(2007), "Why Pigs Don't Have Wings," *London Review of Books*, Vol. **XXIX**, no. 20, October.

Frank, P. (1961), *Modern Science and its Philosophy*, Collier Books: New York.

Freidin, R. (2007), *Generative Grammar: Theory and its History*, Routledge: New York.

Freidin, R. and Vergnaud, J. R. (2001), "Exquisite Connections: Some Remarks on the Evolution of Linguistic Theory," *Lingua* **111**: 639–66.

Fukui, N. (1996), "On the Nature of Economy in Language," *Ninti Kagaku* [*Cognitive Studies*] **3**: 51–71.

Galileo, G. (1967), *Dialogue Concerning the Two Chief World Systems* (2nd edn), Drake, S. (ed. and trans.), Berkeley: University of California Press.

Gallistel, C. R., Gelman, R., and Cordes, S. (2006), "The Cultural and Evolutionary History of the Real Numbers," in Levinson, S. C. and Jaisson, P. (eds.), *Evolution and Culture*, Cambridge, MA: MIT Press, 247–74.

Gambell, T. and Yang, C. (2003), "Scope and Limits of Statistical Learning in Word Segmentation," ms., New Haven, CT: Yale University.

Giblin, I. (2008), "Music and the Generative Enterprise," doctoral dissertation, University of New South Wales.
Gibson, E. and Wexler, K. (1994), "Triggers," *Linguistic Inquiry* **25**: 407–54.
Golumbia, D. (2010), "Minimalism is Functionalism," *Language Sciences* **32**: 28–42.
Gould, S. J. and Lewontin, R. C. (1979), "The Spandrels of San Marco and the Panglossian Paradigm: a Critique of the Adaptationist Program," *Proceedings of the Royal Society of London* **205**: 581–98.
Grohmann, K. K. (2003), "Symmetries in Locality," available at: www.punksinscience.org/kleanthes/papers/fest.pdf.
(2006), "The Road to PF," available at: www.punksinscience.org/kleanthes/papers/istal17.pdf.
Haegeman, L. (1994), *Introduction to Government and Binding Theory*, Oxford: Blackwell.
Hamilton, W. R. (1967[1833]), "On a General Method of Expressing the Paths of Light, and the Planets, by the Coefficients of a Characteristic Function," in Halberstam, H. and Ingram, R. E. (eds.), *The Mathematical Papers of Sir William Rowan Hamilton*, Vol. **III**, New York: Cambridge University Press, 311–32.
Hankins, T. L. (1985), *Science and the Enlightenment*, Cambridge History of Science Series, University of Cambridge: The Press Syndicate.
Hauser, M. D. (2009), "The Mind," *Scientific American*, September, 44–51.
Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002), "The Faculty of Language: What is it, Who has it, and How did it Evolve?", *Science* **298**: 1569–79.
Hesse, M. (1961), *Forces and Fields: the Concept of Action at a Distance in the History of Physics*, London: Nelson.
Hinzen, W. (2006a), *Mind Design and Minimal Syntax*, Oxford University Press.
(2006b), "Minimalist Foundations of Language: on the Question of Why Language is the Way it is," in Cangelosi, A., Smith, A., and Smith, K. (eds.), *The Evolution of Language: Proceedings of the 6th International Conference (EVOLANG6)*, Singapore, World Scientific, 115–22.
(2009), "Hierarchy, Merge, and Truth," in Piattelli-Palmarini, M., Uriagereka, J., and Salaburu, P. (eds.), *On Minds and Language*, Oxford University Press, 123–40.
Hinzen, W. and Uriagereka, J. (2006), "On the Metaphysics of Linguistics," *Erkenntnis* **65**: 71–96.
Hiraiwa, K. (2005), "Dimensions of Symmetry in Syntax: Agreement and Clausal Architecture," doctoral dissertation, Cambridge, MA: MIT Press.
Hofstadter, D. R. (1999[1979]), *Gödel, Escher, Bach: an Eternal Golden Braid*, Basic Books.
Hornstein, N. (1995), *Logical Form: from GB to Minimalism*, Oxford: Blackwell.
(2009), *A Theory of Syntax: Minimal Operations and Universal Grammar*, Cambridge University Press.
(2013), "Falsifiability," *Faculty of Language*, available at: http://facultyoflanguage.blogspot.com/2013/07/falsifiability.html.
Hornstein, N., Nunes, J., and Grohmann, K. K. (2005), *Understanding Minimalism: an Introduction to Minimalist Syntax*, Cambridge University Press.
Hull, D. (1972), "Reduction in Genetics – Biology or Philosophy?", *Philosophy of Science* **39**: 491–9.

Huntley-Fenner, G., Carey, S., and Solimando, A., (2002), "Objects are Individuals but Stuff doesn't Count: Perceived Rigidity and Cohesiveness Influence Infants' Representations of Small Groups of Discrete Entities," *Cognition* **85**: 203–21.

Imai, M. and Gentner, D. (1997), "A Cross-linguistic Study of Early Word Meaning: Universal Ontology and Linguistic Influence," *Cognition* **62**: 169–200.

Ingram, J. (2007), *Neurolinguistics: an Introduction to Spoken Language Processing and its Disorders*, Cambridge Textbooks in Linguistics, Cambridge University Press.

Jackendoff, R. (1983), *Semantics and Cognition*, Cambridge, MA: MIT Press.

(2008), "Alternative Minimalist Visions of Language," available at: www.summer10.isc.uqam.ca/page/docs/readings/JACKENDOFF_Ray/AMV.pdf.

Jackendoff, R. and Pinker, S. (2005), "The Nature of the Language Faculty and its Implications for Evolution of Language," reply to Fitch *et al*. 2005, *Cognition* **97**: 211–25.

Jenkins, L. (2000), *Biolinguistics: Exploring the Biology of Language*, Cambridge University Press.

Johansson, S. (2006), "Constraining the Time When Language Evolved," in Cangelosi, A., Smith, A., and Smith, K. (eds.), *The Evolution of Language: Proceedings of the 6th International Conference (EVOLANG6)*, Singapore, World Scientific, 152–9.

Johnson, D. and Lappin, S. (1997), "A Critique of the Minimalist Program," *Linguistics and Philosophy* **20**: 273–333.

(1999), *Logical Constraints vs. Economy*, Stanford, CA: Center for Study of Language and Information.

Jowett, B. (trans.) (1961), "Plato: *Timaeus*," in Hamilton, E. and Cairns, H. (eds.), *The Collected Dialogues of Plato*, New York: Bollingen.

Kant, I. (1997), *Lectures on Metaphysics*, Ameriks, K. and Naragon, S. (trans. and ed.), Cambridge University Press.

Katz, J. J. and Postal, P. M. (1964), *An Integrated Theory of Linguistic Descriptions*, Cambridge, MA: MIT Press.

Kim, J. (1992), "Multiple Realization and the Metaphysics of Reduction," *Philosophy and Phenomenological Research* **52**: 1–26.

(1998), *Mind in a Physical World: an Essay on the Mind-Body Problem and Mental Causation*, Cambridge, MA: MIT Press.

Kinsella, A. R. (2009), *Language Evolution and Syntactic Theory*, Cambridge University Press.

Kitcher, P. (1984), "1953 and All That: a Tale of Two Sciences," *Philosophical Review* **93**: 335–73.

Koster, J. (2007), "Structure Preservingness, Internal Merge, and the Strict Locality of Triads", in Karimi, S., Samiian, V., and Wilkins, W. K. (eds.), *Phrasal and Clausal Architecture: Syntactic derivation and interpretation* [Linguistik Aktuell/ Linguistics Today 101], Amsterdam: John Benjamins, 188–205.

Koyré, A. (1957), *From the Closed World to the Infinite Universe*, Baltimore, MD: The Johns Hopkins Press.

Lanczos, C. (1970[1949]), *The Variational Principles of Mechanics*, University of Toronto Press.

Langendoen, D. T. (2003), "Merge," in Carnie, A., Harley, H., and Willie, M. (eds.), *Formal Approaches to Function in Grammar*, Amsterdam: John Benjamins Publishing Co., 307–18.

Lappin, S., Levine, R., and Johnson, D. (2000), "The Structure of Unscientific Revolutions," *Natural Language and Linguistic Theory* **18**: 665–71.

(2001), "The Revolution Maximally Confused," *Natural Language and Linguistic Theory* **19**: 901–19.

Larson, R., Deprez, V., and Yamakido, H. (eds.) (2010), *The Evolution of Human Language: Biolinguistic Perspectives*, Cambridge University Press.

Lees, R. B. (1957), "Review of Chomsky (1957)," *Language* **33**: 375–408.

Legate, J. and Yang, C. (2002), "Empirical Re-assessment of Stimulus Poverty Arguments," *The Linguistic Review* **19**: 151–62.

Lenneberg, E. (1967), *Biological Foundations of Language*, New York: Wiley.

LePore, E. and Loewer, B. (1989), "More on Making Mind Matter," *Philosophical Topics* **17**: 175–91.

Levine, R. (2002), "Review of Uriagereka (1998)," *Language* **78**: 325–30.

Lewis, D. (1972), "Psychophysical and Theoretical Identification," *Australasian Journal of Philosophy* **50**: 249–67.

Mahajan, N., Barnes, J., Blanco, M., and Santos, L. (2009), "Enumeration of Objects and Substances in Non-human Primates: Experiments with Brown Lemurs (*Eulemur Fulvus*)," *Developmental Science* **12**: 920–8.

Mandelbrot, B. (1982), *The Fractal Geometry of Nature*, New York: W. H. Freeman.

Marantz, A. (1995), "The Minimalist Program," in Webelhuth, G. (ed.), *Government and Binding Theory and the Minimalist Program*, Cambridge, MA: Blackwell, 349–82.

Markson, L. and Bloom, P. (1997), "Evidence Against a Dedicated System for Word Learning in Children," *Nature* **385**: 813–15.

Marr, D. (1982), *Vision: a Computational Investigation into the Human Representation and Processing of Visual Information*, New York: W. H. Freeman and Company.

Martin, R. and Uriagereka, J. (2000), "Some Possible Foundations of the Minimalist Program," in Martin, R., Michaels, D., and Uriagereka, J. (eds.), *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, Cambridge, MA: MIT Press, 1–30.

Maxwell, J. C. (2003[1876]), "On Action at a Distance," in Niven, W. D. (ed.), *The Scientific Papers of James Clerk Maxwell*, Vol. **II**, Dover Phoenix Editions, 311–23.

Maynard Smith, J., Burian, R., Kauffman, S., Alberch, R., Campbell, J., *et al*. (1985), "Developmental Constraints and Evolution," *The Quarterly Review of Biology* **60**: 265–87.

Mayr, E. (1982), *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*, Cambridge, MA, and London: Harvard University Press.

McKeon, R. (ed.) (1941), *The Basic Works of Aristotle*, New York: Random House, Inc.

Mirowski, P. (1989), *More Heat than Light: Economics as Social Physics, Physics as Nature's Economics*, Cambridge University Press.

Nahin, P. J. (2004), *When Least is Best*, Princeton University Press.

Narita, H. (2009), "How Syntax Naturalizes Semantics: Review of Uriagereka (2008)," *Lingua* **119**(11): 1767–75.

Neuenschwander, D. (2010), *Emmy Noether's Wonderful Theorem*, Baltimore, MD: The Johns Hopkins University Press.

Nevins, A., Pesetsky, D., and Rodrigues, C. (2009), "Pirahã Exceptionality: a Reassessment," *Language* **85**(2): 355–404.

Newmeyer, F. J. (2006), "A Rejoinder to 'On the Role of Parameters in Universal Grammar: a Reply to Newmeyer' by Ian Roberts and Anders Holmberg," ms., University of Washington.

Newton, I. (1999), *The Principia: Mathematical Principles of Natural Philosophy*, Cohen, I. and Whitman, A. (trans.), Berkeley: University of California Press.

Newton, I. (2004), *Philosophical Writings*, Janiak, A. (ed.), Cambridge University Press.

O'Keefe, T. (2005), *Epicurus on Freedom*, Cambridge University Press.

Ott, D. (2007), "Reverse-engineering the Language Faculty: Origins and Implications of the Minimalist Program," *Harvard Working Papers in Linguistics* **12**: 77–90.

Penrose, R. (1989), *The Emperor's New Mind*, Oxford University Press.

Perruchet, P. and Rey, A. (2005), "Does the Mastery of Center-embedded Linguistic Structures Distinguish Humans from Nonhuman Primates?", *Psychonomic Bulletin and Review* **12**: 307–13.

Piattelli-Palmarini, M. (ed.) (1980), *Language and Learning: the Debate between Jean Piaget and Noam Chomsky*, New York: Routledge and Kegan Paul.

(1998), "Foreword," in Uriagereka, J. (ed.), *Rhyme and Reason: an Introduction to Minimalist Syntax*, Cambridge, MA: MIT Press.

Piattelli-Palmarini, M., Uriagereka, J., and Salaburu, P. (eds.) (2009), *Of Minds and Language: a Dialogue with Noam Chomsky in the Basque Country*, Oxford University Press.

Pinker, S. and Bloom, P. (1990), "Natural Language and Natural Selection," *Behavioral and Brain Sciences* **13**: 707–84.

Pinker, S. and Jackendoff, R. (2005), "The Faculty of Language: What's Special about it?", ms., Harvard University and Brandeis University. Modified version published in: *Cognition* **95**: 201–36.

Popper, K. (1945), *The Open Society and Its Enemies*, 2 vols., London: Routledge.

(1959[1935]), *The Logic of Scientific Discovery*, London: Hutchinson.

(1961), *The Poverty of Historicism* (2nd edn), London: Routledge.

(1963), *Conjectures and Refutations*, New York: Routledge and Kegan Paul.

(1976), *Unended Quest: an Intellectual Autobiography*, London: Fontana.

Popper, K. and Eccles, J. (1977), *The Self and its Brain: an Argument for Interactionism*, New York: Springer.

Post, E. (1943), "Formal Reductions of the General Combinatory Decision Problem," *American Journal of Mathematics* **65**: 197–215.

Postal, P. (2003), "(Virtually) Conceptually Necessary," *Journal of Linguistics* **39**: 599–620.

Progovac, L. (2010), "When Clauses Refuse to be Recursive: an Evolutionary Perspective," in van der Hulst, H. (ed.), *Recursion and Human Language*, Berlin and New York: Mouton de Gruyter, 193–211.

Pullum, G. and Scholz, B. (2002), "Empirical Assessment of Stimulus Poverty Arguments," *The Linguistic Review* **19**: 9–50.

(2010), "Recursion and Infinitude Claim," in van der Hulst, H. (ed.), *Recursion and Human Language*, Berlin and New York: Mouton de Gruyter, 113–38.

Putnam, H. (1967), "The Nature of Mental States," in Capitan, W. and Merrill, D. (eds.), *Arts, Mind, and Religion*, Pittsburgh: University of Pittsburgh Press.

(1988), *Representation and Reality*, Cambridge, MA: MIT Press.

Quine, W. V. O. (1950), *Methods of Logic*, New York: Holt.

Raj, M. (1996), *Fundamentals of Optics*, New Delhi: Anmol.

Ramón y Cajal, S. (1913–14), *Estudios Sobre la Degeneración y Regeneración del Sistema Nervioso*, Madrid: Moya.

Ringo, J. (1991), "Neuronal Interconnection as a Function of Brain Size," *Brain, Behavior, and Evolution* **38**: 1–6.

Rizzi, L. (1990), *Relativized Minimality*, Cambridge, MA: MIT Press.

(2004), "On the Study of the Language Faculty: Results, Developments, and Perspectives," *The Linguistic Review* **21**: 323–44.

Sagan, H. (1969), *Introduction to the Calculus of Variations*, New York: McGraw-Hill.

Sakel, J. and Stapert, E. (2010), "Pirahã – in need of recursive syntax?", in van der Hulst, H. (ed.), *Recursion and Human Language*, Berlin and New York: Mouton de Gruyter, 3–16.

Samuels, B. D. (2009), "The Structure of Phonological Theory," Harvard University, PhD dissertation.

(2011), *Phonological Architecture: a Biolinguistic Perspective*, Oxford University Press.

Samuels, B. D. and Boeckx, C. (2009), "What Emerges from Merge in Phonology?", paper presented at *OCP 6*, University of Edinburgh, 22–4 January.

Scheer, T. (2004), *A Lateral Theory of Phonology: What is CVCV, and Why Should it Be?*, Berlin and New York: Mouton de Gruyter.

Schreuder, M. and Gilbers, D. (2004), "Recursive Patterns in Phonological Phrases," in Bel, B. and Marlien, I. (eds.), *Proceedings of Speech Prosody 2004*, Nara, Japan: ISCA (SproSig), 341–4.

Schwartz, E. L. (ed.) (1990), *Computational Neuroscience*, Cambridge, MA: MIT Press.

Scott-Phillips, T. C., Tamariz, M., Cartmill, E. A., and Hurford, J. R. (eds.) (2012), *The Evolution of Language: Proceedings of the 9th International Conference (EVOLANG 2012)*, New Jersey, World Scientific.

Semenza, C., Delazer, M., Bertella, L., Granà, A., Mori, I., *et al*. (2006), "Is Math Lateralised on the Same Side as Language? Right Hemisphere Aphasia and Mathematical Abilities," *Neuroscience Letters* **406**: 285–8.

Shapiro, L. A. (2004), *The Mind Incarnate*, Cambridge, MA: MIT Press.

Siegal, M. L. and Bergman, A. (2006), "Canalization," in Fox, C. W. and Wolf, J. B. (eds.), *Evolutionary Genetics: Concepts and Case Studies*, New York: Oxford University Press, 235–51.

Smith, N. (2000), "Foreword," in Chomsky, N., *New Horizons in the Study of Language and Mind*, Cambridge University Press.

(2004), *Chomsky: Ideas and Ideals* (2nd edn), Cambridge University Press.

Smith, N. and Law, A. (2007), "Twangling Instruments: Is Parametric Variation Definitional of Human Language?", *UCL Working Papers in Linguistics* **19**: 1–28.

(2009), "On Parametric (and Non-Parametric) Variation," *Biolinguistics* **3**/4: 332–43.

Sober, E. (1999), "The Multiple Realizability Argument Against Reductionism," *Philosophy of Science* **66**: 542–64.

Stahnisch, F. W. and Nitsch, R. (2002), "Santiago Ramón y Cajal's Concept of Neuronal Plasticity: the Ambiguity Lives On," *Trends in Neuroscience* **25**: 589–91.

Tomalin, M. (2007), "Reconsidering Recursion in Syntactic Theory," *Lingua* **117**: 1784–1800.

Uriagereka, J. (1998), *Rhyme and Reason: an Introduction to Minimalist Syntax*, Cambridge, MA: MIT Press.

(2000), "On the Emptiness of 'Design' Polemics," *Natural Language and Linguistic Theory* **18**: 863–71.

(2001), "Cutting Derivational Options," *Natural Language and Linguistic Theory* **19**: 891–900.

(2002), *Derivations: Exploring the Dynamics of Syntax*, London: Routledge.

(2008), *Syntactic Anchors: on Semantic Structuring*, Cambridge University Press.

Vallauri, E. L. (2004), "The Relation between Mind and Language: the Innateness Hypothesis and the Poverty of Stimulus," *The Linguistic Review* **21**: 345–87.

Van der Hulst, H. (2010), "A Note on Recursion in Phonology," in van der Hulst, H. (ed.), *Recursion and Human Language*, Berlin and New York: Mouton de Gruyter, 301–41.

Varley, R., Klessinger, N., Romanowski, C., and Siegal, M. (2005), "Agrammatic but Numerate," *Proceedings of the National Academy of Sciences*, USA, **102**: 3519–24.

Waddington, C. H. (1940), *Organisers and Genes*, Cambridge University Press.

(1942), "Canalization of Development and the Inheritance of Acquired Characters," *Nature* **150**: 563–5.

Walton, M. T. and Walton, P. J. (1997), "The Geometrical Kabbalahs of John Dee and Johannes Kepler: the Hebrew Tradition and the Mathematical Study of Nature," in Theerman, P. H. and Parshall, K. H. (eds.), *Experiencing Nature: Proceedings of a Conference in Honor of Allen G. Debus*, the Netherlands: Kluwer Academic Publishers, 43–59.

Webelhuth, G. (1995), "Introduction," in Webelhuth, G. (ed.), *Government and Binding Theory and the Minimalist Program*, Cambridge, MA: Blackwell, 1–14.

Weinberg, S. (1976), "The Forces of Nature," *Bulletin of the American Academy of Arts and Sciences* **29** (4): 13–29.

Whitehead, A. N. (1997[1925]), *Science and the Modern World*, New York: Pelican Mentor Books.

Wilks, Y. (1971), "Decidability and Natural Language," *Mind* **80**: 497–520.

Williams, E. (1987), "Introduction," in Roeper, T. and Williams, E. (eds.), *Parameter Setting*, Dordrecht: D. Riedel.

Wilson, M. (1985), "What is This Thing Called 'Pain'? The Philosophy of Science Behind the Contemporary Debate," *Pacific Philosophical Quarterly* **66**: 227–71.

Yang, C. (2010), "Three Factors in Language Variations," *Lingua* **120**: 1160–77.

Yoshimoto, Y. (2001), "A Review of the History of Constraints in the Pre-Minimalist Transformational Generative Grammar," *Scripsimus* **10**: 143–70.

Yourgrau, W. and Mandelstam, S. (1960), *The Variational Principles in Dynamics and Quantum Theory*, London: Pitman and Sons.

Index

ϕ-features, 38

abstract mathematical models, 145
abstraction, 4, 156, 161, 168, 169
acalculia, 104
acceptability, 11, 12, 13
  vs. grammaticality, 11–13
acquisition model, 16
acquisition process, 17, 20
action at a distance, 161, 165, 166, 167, 168
action by contact, 153, 166, 167, *see also* Cartesian causation
action integral, 127, 130
active memory, 124
adaptation, 45, 84, 85, 86, 87, 104
adaptationist explanations, 110, 116
adaptationist stories, 116
adaptive computation, 121
adaptive function, 86
adjacency-rule, 149
agrammatic aphasia, 104
Agree, 38, 40, 64, 65, 66, 67, 97, 106
  ontology of, 41
alterative paths, 131
alternative derivations, 131
ambiguity, 4, 92, 98
anarcho-syndicalism, 44
anatomy, 48, 117, 148, 151
angle of incidence, 136
angle of reflection, 136
animal communication, 45, 86, 87, 88, 106
animal navigation systems, 108
animal non-communicative systems, 88
anthropomorphic interpretations, 139
anti-economy condition, 133
aphasia, 104
approach
  adaptationist, 116
  biolinguistic, 1, 4, 5, 143, 161, 169, 173, 180
  bottom-up, 35
  cognitive functionalist, 164, 165
  comparative empirical, 2
  computational, 150
  dualist, 153
  functionalist, 125
  GB, 36
  goal-directed, 6, 21, 178
  interdisciplinary, 184
  mechanistic, 144
  minimalist, 27, 72, 73, 116
  modular, 24
  naturalistic, 1, 162, 176
  P&P, 16, 17, 19, 32, 46, 49
  physical, 141
  problem-directed, 21, 178
  realistic, 148
  top-down, 35
apriorism, 139
argument from design, 86, 87
argument from ideology, 44
Aristotelian teleology, 135
Aristotle, 135, 136
Aristotle's physics, 135
arithmetic, 90, 102, 103, 104, 107
arithmetic operations, 104
arrows, 138
articulatory-perceptual system, 33, 80
artificial intelligence, 156
artificial systems, 63, 71
asymmetry, 101
Atkinson, M., 34, 54, 55, 56, 57, 60, 66, 67, 68, 70, 97, 98

atomic configurations, 132
atoms, 37, 131, 132, 166
axiomatic system, 9
axioms, 7, 8, 9, 10
axons, 151
Ayala, F. J., 119
Ayoun, D., 18

B&H, 6, 7–12, 13–14, 15, 16, 18, 20–1
bar levels, 37
bare output conditions, 53
beauty, 134
behavior of light, 138, 141
behavioral disposition, 154, 155
behavioral hypotheticals, 155
behavioral outcomes, 155
behavioral outputs, 157
Bell, A. E., 139
Bergman, A., 74
Berwick, R. C., 97
best of all possible brains, 150
beyond explanatory adequacy, 74, 161
Bickerton, D., 90
Bickle, J., 158
Big Bang, 149
binding theory, 36
biolinguistics, 2, 4, 5, 161
biological device, 143
biological necessity, 24, 59
biological structure, 84, 151
biological system, 48, 50, 60, 70, 124
biological world, 41
biology, 7, 13, 44, 47, 48, 50, 59, 84, 125, 150, 174, 176, 178, 180
blind watchmaker, 119
Block, N., 153, 154, 160, 174, 175, 176
blood pressure, 122
Bloom, P., 84, 85, 90, 91, 110, 115, 120, 121
Boden, M. A., 6
bodyhood, 166
Boeckx, C., 2, 6, 16, 17, 26, 34, 53, 95, 126, 178, 181
book of nature, 135
Bouchard, D., 90
boundary conditions, 14, 111
Boyle's law, 122
brain connections, 148
brain science, 4, 168, 174
brain states theorists, 158
brain structure, 151
brain type, 156
bridge laws, 160
Burtt, E. A., 135

calculus of variations, 138
canalization, concept of, 74
Cangelosi, A., 83
Carey, S., 183
Carnap, R., 7, 49
Carnie, A., 120
Cartesian causation, 166, 167
Cartesian dualism, 161, 164, 165, 166, 167, 168, 169
Cartesian philosophy, 153
Cartesianism, 164, 165
Cartesians, 137, 166
case assignment, 106
case theory, 36
causal role, 156, 157
c-command, 40
Cela-Conde, C.J., 144, 162, 165, 169
cerebral cortex, 47, 149
ceteris paribus clauses, 57, 58
C-fibre firing, 154
Chabay, R. W., 131
channel, 152
chemical elements, 134
chemistry, 125, 175
Cherniak, C., 48, 49, 143, 147, 148, 149, 150, 151, 173, 180
  and a combined hypothesis, 150, 152
  and minimal rationality, 147
  and non-genomic nativism, 147, 150, 172
  and physical optimization, 172
  and the adjacency rule, 150
  and the MP, 151, 152, 170
  and the notion of optimality, 170
  work in neuroscience, 147
chess-playing machine, 140
Chomsky, N.
  and Cherniak's work. *See* Cherniak N. and the MP
  and epistemological nativism, 46
  and Fodor, J., 176
  and Fodorian functionalism, 161–4
  and linguistic and interdisciplinary discourses, 3, 52, 77, 93, 94, 102, 108, 179, 181
  and modern cognitive psychology, 164, 176

Chomsky, N. (cont.)
and naturalism, 5, 34, 54, 143, 144, 145, 147, 161, 162, 163, 168, 169, 176, 180
and post-Newtonian philosophers, 164
and the definition of UG, 105
and the Galilean style, 145, 146, 163
and the genetic instruction to use Merge, 107
and unbounded Merge, 101, 106
appeal to Cherniak's work, 170, 171
Aspects, 13, 15, 16, 21, 32
Cartesian Linguistics, 120, 121
comparison between linguistic and interdisciplinary discourses, 78–81
influence of Carnap and Post, 7, 49
influence of Husserl and Weinberg, 145
influence of Jacob, 49
Lectures on Government and Binding, 18, 32
on biological necessity, 24
on Functionalism, 164, 170
on Popper's concept of falsification, 182
on the definition of recursion, 89
on the history of generative grammar, 6
on the language-specificity of Merge, 105
on the mind-body problem, 143, 144, 161, 164, 165, 180
on the P&P approach, 17
on the relationship between language and mathematics, 102
Syntactic Structures, 1, 9, 10, 15
The Minimalist Program, 1
"Three Factors in Language Design," 150
views on politics, 44
Christiansen, M. H., 83
Church, A., 89
Circular explanations, 115
circularity, 106, 107, 179, 183
circulation, 122
classical electrodynamics, 138
Clerselier, C., 137
closed system, 131, 132
coextensions, 160
cogency, 52, 72
cognition, 1, 10, 13, 22, 24, 25, 26, 27, 55, 80, 90, 96, 107, 152, 169, 184
animal, 183
cognitive abilities, 91
prelinguistic, 183
cognitive autonomy, 31
cognitive revolution, 44
cognitive science, 4, 58, 85, 144, 146, 148, 156, 161, 163, 168, 174, 175, 180
cognitive systems, 24, 27, 32, 33, 35, 36, 54, 55, 81, 103, 106, 120, 146
co-linearity thesis, 112, 117, 125
combinatorial network optimization theory, 148
combinatorial system, 65, 67, 73
combined hypothesis, 150, 152
common knowledge, 132
common sense intuitions, 162, 163
commonsense, 161, 163, 180
communication, function of, 44
communicative system, 85
communicative use, 120
compatibility, 160, 163
competence, 11, 12, 15, 104, 107
complementary advantage, 176
complex causal interactions, 155
component placement optimization, 148
computational complexity, 131, 133
computational cost, 130, 131, 149
computational device, 35, 143
computational economy, 124
computational efficiency, 34, 46, 78, 121
computational level, 159
vs. neurological level, 173
computational mechanisms, 86, 87, 92, 97, 98
computational neuroscience, 48, 49
computational operations, 38, 64, 95, 122
computational optimization, 172, 173
computational resources, 147
computational system, 27, 30, 33, 36, 37, 38, 48, 61, 63, 85, 95, 97, 98, 105, 113, 114, 115, 124, 131, 134, 172, 173
computational tricks, 125, 131
computer metaphor, 158, 159
computer model of the mind, 174
computer science, 89, 148, 158
conceptual system, 29, 30
conceptual-intentional system, 33, 69, 79, 80, 85
configuration, 80, 137, 149
connectionist models, 148
conscious design, 119
consistency, 66, 173
lack of, 70
continua, 127
continuity, 22, 29, 86
convergence, 114, 117, 119, 122, 125
convergent derivations, 38, 122, 123, 133
Copernican answer, 137

Copernicus, 135
copy, 39
correspondence, notion of, 172
cosmology, 134
cotton-top tamarins, 90
count nouns vs. mass nouns, 183
countable magnitudes, 103
covert movement, 114
Cowper, E. A., 18
crash, 114, 117, 119, 123
crashed derivations, 123
crashed structure, 124
critical period, 156

Darwin, C., 83, 119
  Darwinian explanation, 44, 45, 117, 119
  neo-Darwinian explanation, 115
data analysis, 73, 75
deeper level of explanation, 5, 32, 74
default hypothesis, 88, 102
Deletion, 97
demarcation criterion, 182
dendrites, 151
Denton-Mendoza, N., 120
derivational decisions, 133
derivational economy, 128
derivational juncture, 125
derivational options, 125
derivational paths, 132
derivational steps, 152, 170, 173
derivational system, 114
Descartes, R., 29, 137, 153, 165, 166, 167, 168, 175
  definition of "matter," 166
descended larynx, 91
descriptive adequacy, 15, 32
design conditions, 40, 41, 65
design of language, 2, 6, 34, 37, 68, 118, 119, 123, 124, 127
design specifications, 62, 111, 119
design without designer, 119
developmental biology, 49
developmental constraints, 48
differential equations, 129, 139
differential statement, 130
differential/integral equations, 129
differentiation, 129
discrepancy, 14, 23, 31
discrete infinity, 86, 87, 89, 90, 103
  definition of, 89
discrete numbers, 103
discrete system, 127
discretely infinite space, 104
disjunction, 160, 173
disjunctive terms, 160, 175
dislocation property, 40, 64
disparity, 31, 35, 151
dispensability, 65
dispositional statement, 154
distance, 137
diversity, 31, 105
DNA-based mechanisms, 151
dogmatic attitude, 181
dogmatic pseudoscience, 182
dogmatism, 26, 168, 181
domain-general mechanisms, 91
D-structure, 36, 62
dualism, 145, 153, 164, 165, 168, 169, 180
  methodological. *See* methodological dualism
  new form of, 177
  vs. materialism, 153–4
Dugas, R., 134, 137
dynamics, 122

Eccles, H. C., 166
economics, 57, 58
economy, 42, 53, 54, 122, 123, 124, 126, 127, 128, 133, 134, 140, 141
  conditions, 43
  fundamental principles of, 43
  nature of, 139
  notion of, 41
economy considerations, 123, 124, 134, 141
economy of derivation, 41, 125, 127, 128, 129, 130, 131, 140, 170
economy of nature, 138
economy of representation, 41, 125
efficient computation, 73, 75, 79, 80, 133
Einstein, A., 20
Eldredge, N., 84
electricity, 134
electron, 145
elegance, 18, 19, 21, 53, 54, 131
eliminativism, 161
embedding hierarchies, 107
emergent properties, 47
Empedocles, 134
empirical coverage, 36

empirical justification, 67, 73
empty space, 166
ends justify the means, 119
energy, 129, 131, 151, 157
engineering, 7
entropy, 131, 132, 141
  notion of, 124
entropy condition, 131, 132, 133
environmental effects, 151
epiphenomena, 43, 46
epistemic value, 135, 163
Epstein, S. D., 34, 126
essentialist explanation, 166, 167
ethnoscience, 163
Euler, L., 138
evaluation measure, 10, 16, 17
evaluation procedures, 16
Evans, N., 90
Everett, D. L., 90
evidential basis, 183
evolution of language, 2, 44, 45, 50, 68, 69, 76, 77, 83, 84, 85, 102, 113
evolutionary constraints, 48
evolutionary fable, 119
evolutionary force, 86
evolutionary history, 50, 70, 110, 151
evolutionary psychology, 44, 45
evolutionary theory, 48, 84, 93
evolutionary time frame, 110
exaptation, 86
excess baggage, 62, 71
experimental intervention, 117
explanation
  mechanical, 137
  teleology-free form of, 119
explanatoriness, 19
explanatory adequacy, 16, 17, 19, 31, 32, 75, 168, 169
explanatory autonomy, 141, 180
explanatory basis, 121
explanatory power, 36, 75, 121, 124
explanatory status, 4, 108, 109, 117, 121, 123, 143, 173, 180
expression of thought, 44, 45, 68, 121
expressive power, 32, 44
extended substance, 166
extension, 1, 153, 166, 167
extensional equivalence, 62
external conditions, 15, 117
External Merge, 40
fact learning, 91
faculty of language, 22, 25, 28, 32, 33, 35, 36, 55, 61, 74, 78, 85, 95, 96, 101, 104, 118, 151
faith and reason, 138
false dichotomy, 30
falsifiability, 179, 181, 182
  lack of, 92, 181
falsification, 71, 92, 100, 117, 181, 182
  naïve, 181
  vs. verification, 181
feature valuation, 106
Fermat, P., 134, 137, 138, 139
Fermat's principle, 133
Feynman, R. P., 20, 129, 130, 138
FI, 41, 114, 119, 122
final cause, 135
finite cognitive resources, 147
finite device, 148
Fitch, W. T., 2, 3, 27, 52, 76, 78, 79, 80, 82, 83, 85, 90, 91, 92, 93, 95, 96, 98, 99, 100, 101, 179
FLB, 28, 29, 43, 79, 85, 86, 87, 92, 93
  and Factor III, 80
  and FLN, 78, 79, 96
  complexity of, 86
FLN, 28, 29, 82, 85, 86, 87, 91, 92, 93, 95, 98, 102, 106, 121
  and Factor I, 80
  and FLB, 78, 79, 96
  and narrow syntax, 95, 97, 99
  and phonology, 96
  and the SMT, 79, 93
  and UG, 80–1, 101
  apparent complexity of, 94
  content of, 83, 94, 95, 96, 98, 99, 100
FLN/FLB distinction, 92
fluid mechanics, 151
Fodor, J. A., 116, 143, 153, 159, 160, 163, 169, 180
  and *ceteris paribus* laws, 57–9
  and modularity, 25
  and the computer metaphor, 158
  computational/representational theory of mind, 163
  on functionalism, 158
  on identity theory, 156
  on logical behaviorism, 155, 157
  on minimalism, 176
  on reductionism, 160

Fodorian functionalism, 4, 5, 143, 161, 162, 180
folk psychology, 163
foreknowledge, 137
formal deductive system, 8
formal languages, 62, 63, 64
formal semantics, 95, 96, 99, 100
fossil record, 83
four elements, 134
Freidin, R., 4, 16, 34, 114, 115, 126, 134, 140, 141
Fukui, N., 34, 125, 126, 133, 134
  and Hamilton's principle, 126–31
  and the history of minimal principles, 139–40
functional/causal relations, 156
functionalism, 143, 153, 154, 156–8, 160, 161, 162, 164, 165, 169, 173, 176

Galilean style, 145, 146, 162
Galileo, 135, 146
Gallistel, C. R., 103
Gambell, T., 74
GB, 36, 37
general intelligence, 25
general principles, 1, 2, 22, 25, 33, 46, 55, 56, 59, 73, 75, 98, 122
generative grammar, 7, 9, 13, 21, 53, 84, 93, 99
generative program, 6
genes, 22, 150
genetic basis, 24, 32
genetic code, 117
genetic component, 3, 30, 42, 70, 76, 101, 105, 107
genetic constraints, 22, 55
genetic determination, 76
genetic endowment, 1, 22, 23, 24, 31, 33, 46, 50, 56, 60, 73, 110, 151, 178
genetic factor, 23
genetic instruction, 81, 107, 108
genetic make-up, 117
genetics, 59, 66, 110
genome structure, 151
genotype, 24
Gentner, D., 183
geological record, 84
geology, 58
Giblin, I., 90
Gilbers, D., 90
giraffe embryo, 117
giraffes, 116
global condition, 128
globality, 127, 128, 130
Gödel, K., 89
Golumbia, D., 120
good design, 40, 56, 121
Gould, S. J., 84, 85, 116
government and binding model, 35
gradualist view, 83
grammatical sentences, 10, 11
grammaticality, 7, 15
  vs. acceptability, 11–13
grand unified theory, 175
gravity, 166
Greed, 41, 131
Greek atomists, 166
Grohmann, K. K., 17, 19, 53, 54, 72

Hamilton, W. R., 127, 128, 129, 130, 138, 140
Hamilton's equation, 129
handwaving, 113
hard sciences, 4, 146, 175
hardness, 167
hardware, 159
  vs. software, 158
harmony, 134, 181
Hauser, M. D., 2, 3, 27, 28, 30, 43, 52, 78, 79, 82, 83, 85, 87, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 100, 101, 106, 108, 121, 179
head movement constraint, 42
Heraclitus, 134
Hero of Alexandria, 136
Hero's principle, 136, 137
heuristic algorithms, 131
higher apes, 30
Hinzen, W., 72, 104, 110, 112, 113, 115, 117, 175, 177
history of science, 140, 147, 165, 168
Hofstadter, D. R., 90
homo sapiens, 83, 87
Hornstein, N., 6, 17, 19, 34, 36, 53, 54, 66, 124, 126, 178, 181
Huarte, J., 28, 29
Hulst, H. van der, 90
human cognitive faculties, 24
human culture, 103
human evolution, 80, 86
human genome, 22, 151
human infants, 74
  preverbal, 183
human intelligence, 28, 176

human language, 25, 41, 45, 47, 50, 54, 55, 89, 90, 102
  and animal communication, 85, 87
  and language as such, 181
  and UG, 101
  as a natural object, 144
  complexity of, 48
  computational system of, 38, 79
  distinctive feature of, 29
  distinguishing aspects of, 80
  domain of, 106
  expressive power of, 29, 86
  naturalistic approach to, 144
  nature of, 70, 184
  special properties of, 64
  syntax of, 108
human linguistic capacities, 25
human mind, 24, 33, 144, 176
human nature, 44, 45, 144
human prerogative, 28, 83, 86
human thought, 113
Hume, D., 164
Husserl, E., 145
Huygens, C., 139
hypothetical statements, 157

idealization, 146
idealism, 165
identity theory, 154, 155, 156, 157, 169
  and logical behaviourism, 156
  token physicalism, 154, 159, 160
  token physicalism vs. type physicalism, 154
  type physicalism, 154, 155, 156, 157, 158, 160, 180
Imai, M., 183
immediate successor, 103
impenetrability, 167
imperfection, 62, 63, 65, 70, 71, 72
  and perfection, 65
  apparent, 73
  notion of, 71, 73
  of Move, 39
  source of, 40
  strategy, 52, 61, 62, 66, 69, 70, 77, 179
implausibility, 161
inclusiveness condition, 37, 39
inconsistency, 51, 100, 105, 179
indirect object, 121
indispensability, 65
individuated vs. unindividuated entities, 183
induction, 23
inertia, 167
inertiatic state, 74
inference rules, 7, 9
infinite number, 131
infinite series, 103
inflection, 64
information processing systems, 156
Ingram, J. C. L., 174
inherent property of nature, 137
initial state, 32, 42, 146
initial strings, 7
innate knowledge, 74, 75
innate structure, 23, 24, 31
innateness, 23, 31, 44, 178
inorganic world, 126, 140
input data, 18
input-output relations, 35
integers, 103, 104
integral equation, 129, 130
intelligent design, 119, 120
intentional realism, 163
interdisciplinary cooperation, 85
interface conditions, 33, 34, 35, 40, 43, 51, 53, 54, 56, 59, 69, 76, 77, 79, 80, 81, 82, 93, 107, 109, 118, 121, 179, 183, 184
interface constraints, 183
interface requirements, 41, 63, 111
interface-based explanation, 109, 110, 111, 113, 114, 115, 117, 118, 121, 133, 179, 183
interfaces, 27, 33, 37, 38, 41, 53, 56, 85, 86, 87, 92, 93, 97, 98, 111, 113, 114, 116, 124, 141, 184
interface-story, 117
internal justification, 19
Internal Merge, 40
interpretability condition, 37
Interpretable features, 37
intonation pattern, 15
isomorphism, 46

Jackendoff, R., 22, 28, 30, 32, 44, 45, 78, 90, 91, 92, 93, 96, 99, 100, 101, 120
Jacob, F., 49
Jenkins, L., 48
Johansson, S., 110, 151
Johnson, D. E., 44
Just-so stories, 85, 116

Kant, I., 164, 167, 168
Kepler, J., 135
kinetic energy, 127, 129, 132
Kinsella, A., 28, 78, 82, 83, 89, 90, 92, 93, 94, 95, 96, 97, 99
Kirby, S., 83
Koyré, A., 166, 167

La Mettrie, J. O., 144
label, 39
Lagrange, J. L., 138, 139
Lamarckian story, 116
Lanczos, C., 139
Langendoen, T., 53, 54
language acquisition device, 35
language and mind, 144, 145
language as such, 55, 65, 66, 67, 68, 69
language design, 35, 62, 72, 119, 151
language evolution, 41, 45, 75, 82, 85, 86, 110, 116, 151, 179
  problem of, 46
language faculty, 25, 26, 29, 53, 54, 106, 112
  and innate structure, 23
  and legibility conditions, 112
  and neighbouring systems, 110
  and recursion, 28
  and the lexicon, 37
  and usability, 26
  as a computational device, 35
  classic view of the, 23
  complexity of the, 151
  computational account of the, 174
  conception of the, 22
  evolutionary development of the, 50
  genetic component of the, 3, 70, 76, 101
  nature of the, 3, 28, 33, 66, 70, 77
  principled explanation of the, 75, 121
  specificity of the, 27
  structure of the, 27, 56
language properties, 56, 57, 59, 80, 96, 110, 117, 119
language-like system, 38, 71
Lappin, S., 44, 133
Larson, R., 83
Lasnik, H., 122
last resort, 40, 41, 122
law of reflection, 136
Law, A., 90
law-like statements, 57
laws of nature, 34, 48, 122, 140
laws of physics, 2, 4, 34, 48, 49, 59, 109, 121, 124, 125, 126, 128, 140, 147, 151, 152, 181
learning capacities, 90
least action, 127, 133, 134, 138, 139
  as a pitiable axiom, 139
  notion of, 127
  principle of, 127
least distance, principle of, 136
least effort, 41
least time, 133, 134, 137, 138
Lees, R. B., 9, 10
legal slots, 148
legibility conditions, 1, 27, 36, 37, 56, 68, 71, 72, 111, 112, 113, 115, 119, 123, 124, 125, 184
  and Agree, 67
  and conceptual necessity, 60, 67, 68
  and Merge, 65
  and minimalist explanation, 111, 121
  and symbolic systems, 64
  and the principle of economy, 122
  as a source of perfection, 63
  as an aspect of the SMT, 75
  as an empirical issue, 113
  as interface conditions, 33
  elimination of, 117
legibility considerations, 125
legibility constraints, 117
Leibnizian optimism, 35
Lenneberg, E. H., 1
lesions, 105
levels of representation, 36, 72
Levine, R. D., 126
Levinson, S. C., 90
Lewontin, R. C., 116
lexical items, 35, 36, 37, 38, 89, 103
lexicography, 66
lexicon, 36, 37, 38, 53, 68, 96, 100, 102, 103
LF-interface, 38
linguistic analysis, 2, 14, 62, 71, 184
linguistic computation, 134
linguistic domain, 183
linguistic experience, 14, 15, 16, 23, 24, 33, 46, 75
linguistic expressions, 54, 81, 89, 103, 104, 107
linguistic intuition, 10, 11, 13, 15
linguistic knowledge, 14, 23, 50
linguistic phenomena, 1, 2, 4, 31, 75, 182
linguistic sign, 90

linguistic structure, 10, 11, 12, 13, 15, 16, 170, 172
  complexity of, 23, 24
linguistic theorizing, 50
linguistic theory, 4, 6, 7, 9, 10, 11, 13, 16, 21, 89
linguistics, 1, 145, 156, 184
  and cognitive psychology, 28
  and neurobiology, 175
  and physics, 4, 121, 126, 141, 145
  and Plato's problem, 14
  and the natural sciences, 34
  and unification, 146
  Chomskyan, 1, 2, 6, 10, 19, 21, 31, 152, 178
  recursion in, 89
  scientific merit of, 2
  scientific status of, 174
  the Galilean style in, 146
living cells, 157
locomotion, 66
logic, 7, 9
logical behaviourism, 154, 155, 156, 157
logical connection, 156
logical form, 36, 39
logical systems, 7, 8
look-ahead, 119, 131, 133
loose correlations, 125

macrostate, 131
Mahajan, N., 183
major recantation, 30
Mandelbrot, B. B., 90
Mandelstam, S., 132, 134, 136, 138, 139
Marantz, A., 41, 42
markedness, 17
Markson, L., 91
Marr, D., 174
Marr's three levels of analysis, 174
  computational and algorithmic levels, 174
  implementation level, 174
Martin, D. M., 34
Marty, G., 4, 144, 162, 165, 169
mass, 134, 137, 145
Matching, 97
material world, 166
materialism, 153, 154, 161, 165, *see also* logical behaviorism and identity theory
materiality, 166
mathematical calculations, 104
mathematical capacity, 102
mathematical computational ability, 104
mathematical form, 147
mathematical logic, 89
mathematical spaces, 104
mathematics, 102, 104
  development of, 103
  foundations of, 102
Maupertuis, P.-L., 127, 137, 138, 139
Maupertuis's principle, 133
maximum disorder, 132
Maxwell, J. C., 165
Mayr, E., 83
mechanical philosophy, 165, 168
mechanics, 127, 128, 133, 138, 165
medieval scholasticism, 138
mental aspects, 34, 144
mental causation, 153, 154, 155, 156
mental causes vs. mental effects, 155
mental computation, 26, 152
mental events vs. brain events, 156
mental organ, 1, 25
mental phenomena, 144, 162, 176
mental representation, 152
mental states, 121, 152, 153, 154, 155, 156, 157, 158, 164, 169, 174
mental structures, 24, 26, 27, 141
mental substance, 153
mental token vs. brain token, 154
mental type, 156, 171
  vs. brain type, 154
mental-to-mental causation, 155
Merge, 3, 64, 70, 73, 77, 80, 82, 93, 97, 98, 105, 109
  and Agree, 40
  and cognitive systems, 80
  and discrete infinity, 89, 103, 104
  and movement, 39
  and recursion, 3, 28, 81, 82, 93, 94, 97, 98, 99, 101
  and the genetic instruction, 107, 108
  and the SMT, 98
  and UG, 102
  and unbounded Merge, 107
  as a (virtual) conceptual necessity, 40, 53, 65, 66, 67, 70
  as a recursive operation, 75, 82
  as a unique property of language, 102, 106
  as an aspect of perfection, 65
  conceptual necessity of, 67
  language specificity of, 105
  ontological status of, 4, 38, 77, 78, 179
  value of, 39

Merge-only hypothesis, 3, 82, 83, 93, 99, 102, 106, 107, 108, 179
metaphors, 133, 150
metaphysical foundation, 135
metaphysical ideals, 134
metaphysical import, 144
metaphysical naturalism, 161
metaphysical significance, 127
methodological consequences, 146
methodological considerations, 3, 19, 21, 146
methodological dualism, 144, 145
methodological naturalism, 144, 146, 162
  vs. substantive naturalism, 146
microstates, 131
mind/brain, 54, 64, 68
mind-body problem, 153, 163, 164
minimal principles, 133, 134, 137, 138, 139
minimal rationality, 147
minimal search, 170
minimalism, 1, 2, 5, 6, 19, 20, 21, 22, 30, 32, 35, 36, 39, 43, 47, 50, 53, 55, 69, 93, 104, 109, 113, 118, 126, 141, 147, 151, 169, 176, 178, 180, 181, 182, 184, 195
  and GB framework, 36–7
  and language specificity, 25, 27
  and physics, 4, 20, 109, 133, 134, 141, 180
  and the argument from poverty of stimulus, 23
  and the conception of language, 34
  and the notion of economy, 41–3
  and the SMT, 1, 2, 20, 51, 55, 178
  and virtual conceptual necessity, 51, 52
  as a research program, 2, 46
  emergence of, 6, 18, 23, 32, 44, 46, 48
  explanandum of, 32, 33
  explanantia of, 33–4, 36, 41
  major topic of, 32
  revised version of, 5, 180
  shift to, 3, 6, 19, 22, 27, 31, 43–9, 110
  tenets of, 75
minimalist explanation, 4, 33, 34, 41, 42, 74, 109, 112, 115, 117, 119, 120, 121, 125
minimalist literature, 51, 67, 72, 114, 121, 126, 141, 180
minimalist program. *See* minimalism
minimization, 137
minimum principles, 127, 135, 136, 137, 139, 141, 180
miracle, 76, 107, 178
Mirowski, P., 138
mobility, 167
mode of explanation, 115, 124, 167, 168, 180
modern cognitive psychology, 164, 176
modularity, 24, 25, 26, 27, 44
molecules, 132
monograph, 126
morphology, 63, 64, 96
motion, 127, 129, 130, 132, 136
Move(ment), 38, 39, 41
  as a (virtual) conceptual necessity, 40
  as a composite operation, 40
  as a virtual conceptual necessity, 73
  as an imperfection, 39, 73
  ontology of, 39
MRA, 158, 159, 160, 161, 164, 168, 171, 172, 173, 174, 175, 176, 180, 195
multi-dimensional space, 104
mutation, 46, 48
mystical aura, 127
mystical ideas, 137
mystical implications, 128, 141
mystical origins, 139
mysticism, 131, 134, 137

Narita, H., 114, 117
narrow syntax, 87, 95, 97
nativist view, 26
natural class, 59
natural kind predicates, 160
natural languages, 31, 32, 40, 62, 63, 64, 70
natural number system, 103
natural numbers, 102, 103, 104
natural object, 35, 54, 124, 144
natural phenomena, 20, 134, 166
natural philosophy, 135, 165
natural sciences, 126, 141, 145, 163, 175
  and linguistics, 34
natural selection, 45, 46, 48, 83, 84, 85, 86, 91, 94, 110, 115, 119, 151
natural set, 59
natural world, 1, 34, 54, 55, 135, 136, 138, 140, 144, 162, 180
natural biological order, 103
naturalization of mind, 163
naturalness, 19
nature vs. nurture dichotomy, 150
navigation, 87, 88
neat mind vs. tidy brain, 176
neo-Darwinian explanation, 120
Neuenschwander, D. E., 129
neural connections, 148

neural networks, 48
neural optimization, 150
neural structure, 48, 150, 170, 172
neural systems, 149, 172
neural wire connections, 170
neural wiring, 152
neurobiology, 175
neurological findings, 157
neurology, 4, 168
neuron arbors, 151
neuronal plasticity, 156
neurophysiological events, 156
neuroscience, 48, 50, 141, 144, 146, 147, 156, 171, 178
neutrality, 23
Newmeyer, F. J., 18
Newton, I., 135, 144, 161, 162, 163, 164, 165, 166, 167, 168
  definition of "body," 167
  Principia, 135
Newtonian mechanics, 129
non-genetic factors, 32
non-genetic nativism, 22, 23, 30, 31, 178
non-genomic nativism, 150, 151, 180
non-human animals, 87, 91, 92, 113, 183
non-human communicative systems, 29, 47, 106
non-human primates, 90
non-human species, 24, 76
non-linguistic domains, 25, 106
nonnaturalistic, 175
non-optimal computation, 173
non-optimization, 173
non-reducibility, 152
no-tampering condition, 170
number quantification, 87, 88
numeration, 38

O'Keefe, T., 166
object of inquiry, 10, 34, 146
observable world, 134
Occam's razor, 20
one-to-one correlation, 171
ontological commitment, 152
ontological distinction, 154, 183
ontological dualism, 165
ontological kind, 104
ontological thesis, 34
optical law, 136
optics, 127, 128, 133
optimal brain structure, 150
optimal computation, 41, 60, 68, 71, 82, 121, 133, 161, 174, 179, 180
  and Agree, 67
  and conceptual necessity, 67
  and convergence, 122
  and Factor III, 117
  and interface conditions, 59
  and minimalist explanation, 112
  and optimal-wiring, 147
  and physical law, 4
  and physical principles, 133, 141, 143
  and the MRA, 171, 172
  and the notion of "look-ahead," 119
  and the philosophy of mind, 143
  and true explanation, 171
  and VCN, 53
  as a consequence of a physical law, 125
  as a mode of explanation, 124
  as a primitive, 176
  as a source of minimalist explanation, 109
  as a source of perfection, 63
  as a tool of minimalist explanation, 34, 36
  as an aspect of the SMT, 75, 109
  aspects of, 152, 170
  deduction of, 4
  explanatory power of, 124
  explanatory role of, 121
  explanatory status of, 4, 123, 171
  general principles of, 56
  notion of, 78, 124, 181
  recourse to, 123
  two views of, 125
optimal design, 34, 121, 147
optimal neural network, 150
optimal solution, 20, 43, 51, 56, 79, 80, 93
optimal types, 74
optimality, 33, 34, 36, 41, 118, 125, 147, 170, 179
optimal-wiring, 48, 147
optimization, 35, 140, 149, 151, 152, 172
organical structure, 144, 162
organized matter, 162
Ott, D., 126

paleontological data, 84
Pandora's box, 126
Pangloss, Dr, 85, 115
panglossianism, 115
parametric settings, 17, 18, 32
parametric variation, 90
Peano, G., 89

Penrose, R., 90
perfect design, 61, 62, 63, 69, 71
perfection, 62, 63, 65, 66, 69, 71, 134
performance, 12, 26, 33, 104, 107, 109, 114
performance systems, 26, 33, 114, 116, 118
Perruchet, P., 90
phase theory, 124
phenotype, 74
phenotypic development, 74
philosophical doctrine, 4, 143, 153, 158, 164, 168, 180
philosophy of mind, 5, 143, 152, 153, 180
phonetic and semantic features, 37
phonological form, 36
phonological rules, 98
phonological system, 95, 97
phonology, 26, 68, 90, 93, 95, 96, 97, 98, 99, 100
phrasal categories, 37
phrase structure grammars, 7, 8, 9, 90
phrase structure rules, 7
phrase-marker, 128
physical basis, 126, 141, 152, 158
physical causation, 153, 166, 167
physical composition, 158, 160
physical constraints, 49
physical explanation, 150
physical factors, 74
physical law, 22, 23, 48, 50, 122, 125, 131, 147, 151, 180
  notion of, 4
physical laws, 49, 50, 122, 124, 125, 143, 151, 152, 161, 169, 172
physical neatness, 177
physical optimization, 172
  vs. physical non-optimization, 173
physical phenomena, 129
physical predicates, 160
physical principles, 50, 109, 125, 133, 134, 141, 143, 180
physical quantities, 134
physical sciences, 126, 140
physical substance, 153
physical theory, 139
physical world, 139, 153, 154
physics, 128, 133, 146, 152, 160
  and linguistics, 4, 121, 126
  and minimalism, 7, 20, 141, 145
  and mysticism, 131, 137
  and the special sciences, 159
  as the most fundamental level of explanation, 122
  branch of, 132
  Cartesian approach to, 144
  considerations of economy in, 127
  cutting edge of, 175
  development of, 140
  economy principles in, 133
  generalizations of, 159
  history of, 136
  ideally completed, 57, 160, 161
  kind of, 4
  language of, 169
  minimal principles in, 134, 135, 139
  model of, 180
  modern, 109, 135, 138, 140
  of the brain, 48, 141
  of the nineteenth century, 139
  of the pre-nineteenth century, 118
  principle of least action in, 127, 134
  theorems of, 127
Piaget, J., 25, 42, 47, 49
Piagetian view, 26
Piattelli-Palmarini, M., 2, 25, 26, 34, 53, 89, 107, 126, 182
pick-and-choose attitude, 124
Pinker, S., 22, 30, 44, 45, 78, 84, 85, 90, 91, 92, 93, 96, 99, 100, 110, 115, 120
Plato's problem, 7, 11, 13, 14, 17, 18, 19, 20, 43
Poisson, S., 139
politics, 59
Popper, K., 21, 164, 166, 181, 182
Popperian empiricism, 181
Post, E., 7, 49
Postal, P. M., 65, 67
potential energy, 127, 129
poverty-of-stimulus argument, 14, 23, 24, 30, 35
predicates, 159, 160, 173
pre-existing communication systems, 86
prelinguistic human cognition, 183
Premack, D., 25
pre-minimalism, 23
  defining features of, 3, 22
pre-Newton era, 161
Priestley, J., 144
primary linguistic data, 16, 17, 35
primary qualities of matter, 167
primary stress, 74

principled explanation, 1, 4, 59, 75, 109, 110, 113, 119, 121, 161, 169, 180
principles and parameters, 16, 17, 18, 19
principles of language, 2, 4, 109, 121, 125, 126, 128, 140, 147, 181
prior knowledge, 129
probability amplitude, 138
probe-goal relations, 106
problem of justification, 10
problem of language acquisition, 14, 19, 35, 46, 74
procedural thesis, 34
Procrastinate, 41, 131
pro-drop parameter, 18
production rules, 7, 8
Progovac, L., 82, 93
projection principle, 36
proof, 7, 9, 49, 99, 138
proof theory, 9, 49
properties of language, 2, 25, 26, 27, 66, 70, 76, 101, 115, 147
propositional attitudes, 163
propositional content, 9
propositional structures, 84
propositionality, 113
psychological generalizations, 161
psychological kinds vs. physical kinds, 158
psychological law, 160
psychological states, 160
psychology, 13, 28, 45, 58, 59, 156, 157, 159, 171, 176, 180
Pullum, G. K., 87, 90
Putnam, H., 158, 159
Pythagoras, 147
Pythagorean legacy, 134
Pythagorean truth, 135
Pythagoreans, 134

quantifier/variable structure, 114
quantum electrodynamics, 134, 138

Raj, M. G., 138
Ramón y Cajal, S. R., 156
random mutations, 47, 48, 49, 50
rational inquiry, 144, 162, 163
rational numbers, 104
real number system, 103
real refutation, 182
realm of nature, 106
recursion, 46, 78, 80, 81, 89, 92, 93, 95, 97, 98, 100, 101, 102, 108, 121, 179
  and discrete infinity, 87
  and FLN, 86, 98, 99
  and minimalism, 78, 82, 93, 94
  and non-human domains, 87
  and phonology, 97
  and the SMT, 93–4
  as a unique property, 28, 83, 89, 90, 95, 96, 97, 100
  computational mechanisms of, 87
  definition of, 86
  notion of, 3, 28, 82, 83, 89, 94, 99
  options for, 99
  phonological, 100
  range of, 97
recursion-only hypothesis, 3, 52, 77, 78, 81, 82, 83, 88, 89, 90, 92, 94, 97, 98, 99, 102, 106, 108
recursive device, 81, 97, 98, 108
recursive enumeration, 46
recursive generative capacity, 44
recursive mapping, 97
recursive mechanisms, 87, 89, 98
recursive patterns, 90
recursive power, 45, 93
recursive property, 29, 30, 100
recursiveness, 104
reduction, 36, 46, 103, 112, 147, 160, 171
  and the SMT, 99
reductionism, 154, 159, 160, 171
refraction, 137
rewriting rules, 7, 53, 94
Rey, A., 90
rhetorical question, 146, 182
Ringo, J., 48
Rizzi, L., 42
rule writing system, 16
rules of formation, 7, 8
rules of transformation, 7
Russell, B., 164

Sakel, J., 90
saltational evolution, 46
saltational view, 45, 47
saltationism, 46, 47, 48
Samuels, B. D., 82, 83, 92, 93, 95, 96, 97, 99
save wire, principle of, 49
saving wire, 148
skepticism, 25, 26, 105, 161, 163, 180, 181

Scheer, T., 82, 90
scholastic doctrine, 135
Scholz, B. C., 87, 90
Schreuder, M., 90
scientific attitude, 181
scientific feasibility, 174
scientific psychology, 163
scope and discourse properties, 64
Scott-Phillips, T. C., 83
seal of science, 174
second law of thermodynamics, 131
Seely, T. D., 126
segmentation of words, 74
semantic system, 95, 97
semantics, 68, 96, 97, 98, 100, 112, 117
Semenza, C., 104
sensorimotor apparatus, 69
sensory inputs, 156
sensory/motor connections, 149
sensory-motor system, 79, 80, 85
set of properties, 44, 102
  vs. set of laws, 173
Sherwood, B. A., 131
Shortest Move, 41, 42
Siegal, M. L., 74
silicon chips, 159
silicon-based Martians, 158
simplicity, 16, 18, 19, 20, 21, 26, 53, 54, 93, 128, 129, 134, 135, 136, 137, 181
single-dimensional space, 104
SL mechanisms, 74
Smith, J. M., 48
Smith, N., 53, 54, 67, 68, 73, 90
SMT, 3, 4, 20, 26, 43, 50, 51, 52, 61, 69, 71, 72, 73, 77, 80, 82, 93, 98, 99, 101, 102, 106, 112, 117, 125, 140, 143, 171
  a weaker version of, 57
  and *ceteris paribus* clauses, 57–60
  and interface requirements, 63
  and Merge, 77
  and the imperfection strategy, 61, 62, 72
  and UG, 76
  as a strict generalization, 56
  as an empirical thesis, 56, 182
  as an explanatory thesis, 56
  content of, 178, 181
  definition of, 1, 20, 51, 55, 78, 94
  departure from the, 61
  falsification of, 62, 181, 183
  nature of, 179
  three formulations of, 52
  two aspects of, 75, 179
  validity of, 63
social relationships, 87, 88
social-historical order, 103
software, 159
sound and meaning, 20, 26, 27, 35, 37, 60, 68, 79
space and matter, 166
special sciences, 57, 58, 59, 60, 159
species-specific capacity, 55
species-specificity, 28, 31
specificity of language, 27, 30, 106
specified subject condition, 42
specifier, 39, 42
specious explanations, 114, 184
speech, 14, 26, 33, 54, 74, 87, 91, 92
Spell-Out, 38, 41, 131
spoken language, 83
S-structure, 36, 62
standard theory, 32
Stapert, E., 90
Stationary Action, law of, 128
statistical learning, 74
stimuli and responses, 154
structural constraints, 73
structural dependency, 104
structure of language, 18, 19, 31, 120, 151, 152
structured images, 105
structured thoughts, 105
subjacency, 42
substance dualism, 153
sugar molecules, 131
super-engineer, 119
superiority, 42
superiority condition, 42
supply and demand, law of, 57
symbolic systems, 40, 41, 63, 64, 65, 70
symmetric relationship, 160
symmetry, 134
syntactic derivations, 132, 173
syntactic features, 37
syntactic mathematical impairments, 104
syntactic modules, 36
syntactic objects, 38, 41, 89
syntactic representation, 36
syntactic system, 117, 119
syntactic theory, 7, 49, 124
syntax, 18, 26, 90, 92, 93, 97, 98, 100, 108, 112, 115, 117, 119

syntax (cont.)
as the skeleton of thought, 112
radical approach to, 112
syntax-phonology interface, 68
syntax-semantics interface, 68
systematicity, 113

tautological explanation, 114
tautological form, 115
tautology, 100, 109, 117, 179
teleological character, 118, 121, 133, 137
teleological considerations, 119
teleological explanation, 118, 119, 131
teleology, 109, 118, 119, 133, 137, 179
natural vs. artificial, 119
temporary convenience, 4, 165, 168
Thales, 134
theological foundation, 138
theological interpretation, 139
theoremhood, 7
theory construction, 21, 145
theory of computation, 156
theory of language, 34, 50
thermodynamic equilibrium, 132
thermodynamics, 124, 132
theta theory, 36
third factor, 50, 74, 75, 76, 117, 151
third factor conditions, 151
Thompson, D. W., 147
thought system, 105, 112, 120
three factors framework, 33, 52, 73, 74, 76, 77, 78, 79
token physicalism. *See* identity theory
Tomalin, M., 89
total connection costs, 148
transformational rules, 32
transitivity, law of, 171
tree-diagram, 38
tree-like structures, 151
true explanation, 2, 4, 50, 170, 171, 180
Turing, A., 89
Tuscany, grand duke of, 135
type identity, 170, 171
type physicalism. *See* identity theory
typology, 31

UG, 3, 6, 16, 18, 24, 25, 30–4, 35, 42, 46, 65, 73, 75, 76, 78, 79, 99, 102, 122, 159, 178, 179
and FLN, 80–1, 101
as a mental program, 158
as a software, 159
complexity of, 22, 30, 46, 55, 56, 69
content of, 77
notion of, 66
size of, 46
ultimate explanation, 20, 166
unbiological, 174, 175
unbounded Merge, 101, 106
unboundedness, 107
unification in science, 2, 4
prospect of, 144, 146, 147, 161, 164, 168, 180
uniformity, 26, 31, 181
uninterpretable features, 37, 40, 64, 114
uniquely human capacities, 88
uniqueness, 44, 96, 100, 102, 106
unity of science, 169
universe, 43, 125, 127, 128, 134, 145
Uriagereka, J., 4, 34, 42, 53, 112, 117, 124, 125, 126, 131, 132, 133, 134, 140, 141, 175, 177
usability, 111, 112, 124, 125

vacuity, 57, 58
vague analogies, 126, 141
variational principles, 132, 138, 139
Varley, R., 104
vector, 138
vector addition, 138
velocity, 134, 137
Vergnaud, J. R., 4, 34, 114, 115, 126, 134, 140, 141
verificationism, 181
virtual conceptual necessity, 3, 52, 60, 70, 71, 73, 179, 181
interpretations of in the literature, 53–5
vision system, 105
VLSI microcircuit, 148, 149
vocabulary of genetics, 59
void, 166, 167

Waddington, C. H., 74
Walton, M. T., 135
Walton, P. J., 135
wave nature of light, 138
Webelhuth, G., 32
Weinberg, S., 145
well-designed brain, 176

Whitehead, A. N., 5
wh-words, 42
Williams, E., 17
wire length, 148, 149, 150, 170, 171, 172, 173
word learning, 91
X-bar theory, 36, 37, 94
Yang, C., 74
Yourgrau, W., 132, 134, 136, 138, 139