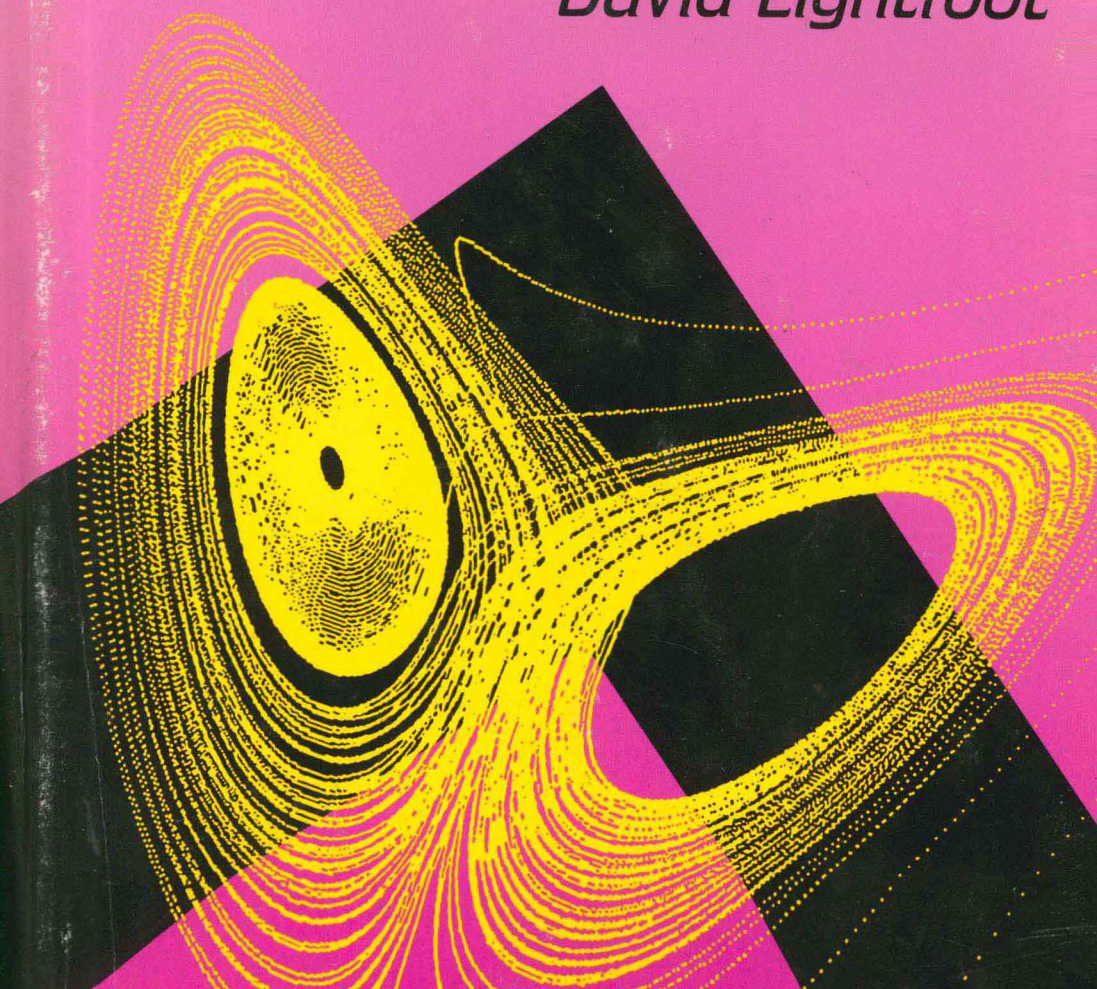


*How to Set Parameters:*

*Arguments  
from  
Language  
Change*

*David Lightfoot*



---

How to Set Parameters:  
Arguments from Language  
Change

David Lightfoot

A Bradford Book  
The MIT Press  
Cambridge, Massachusetts  
London, England

© 1991 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

The book was set in Times Roman by Asco Trade Typesetting Ltd. in Hong Kong and was printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Lightfoot, David.

How to set parameters: arguments from language change / David Lightfoot.  
p. cm.

“A Bradford book.”

Includes bibliographical references (p. ) and index.

ISBN 0-262-12153-0

1. Language acquisition. 2. Linguistic change. 3. Grammar, Comparative and general. I. Title.

P118.L46 1991

401'.93—dc20

90-13513

CIP

---

## How to Set Parameters

---

## Preface

For more than ten years generative grammarians have construed language acquisition as a matter of setting parameters—that is, of fixing option points defined in Universal Grammar. They have busily constructed “parametric differences” between various languages, while drum rolls and trumpet fanfares have heralded a major conceptual shift from the earlier “evaluation” of grammars. There has indeed been such a shift, which has drawn linguists closer to researchers in other fields, and this makes it all the more surprising that so little attention has been paid to what it takes to set these parameters. Sometimes this lack of attention undermines the claims being made, as when an alleged parametric difference is based entirely on data unavailable to children, on negative data, or on data about subtleties of quantifier scope. My most general goal here is to begin to correct this omission by making some claims about the child’s triggering experience and about how parameters are set.

I shall argue first that the triggering experience consists only of robust elements which are structurally simple, and that parameter setting is not sensitive to embedded material. My most precise arguments are based on data from diachronic changes, and claim that grammars would not have been reanalyzed in the way that they were if complex structures influenced the form of the emerging grammars. In fact, the nature of certain changes makes it possible to define the structural limits to the triggering experience rather exactly, and to define some parameters in clearer fashion. My second major claim is that morphology plays an important role in setting parameters which have widespread syntactic effects. I shall examine some consequences of the loss of the rich Old English case system and of the breakdown of the verb classes.

Languages’ histories are typically punctuated by occasional large-scale changes, corresponding to parametric shifts. French, for example, lacks

some Romance characteristics and is sometimes said to have a more Germanic flavor, having set some parameters in the Germanic fashion. Examining the particular clusters of new phenomena that arose during these large-scale changes casts light on the exact nature of the parameters that were set differently and on the consequences; often one finds that a new parameter setting leads to further associated changes. Particularly illuminating, I claim, are cases of obsolescent structures: if one aims to understand language change partly in terms of the way languages are acquired by young children, obsolescence must be treated as a by-product of some new parameter setting, defined at an appropriately abstract level and set by positive primary data. Children don't stop saying things that they hear from their older models simply because they are seeking some stylistic effect or imitating new models. This observation undermines some lexicalist models of language change, which have been fashionable with historical linguists who have defined their interests too narrowly.

In general, I shall seek "ahistorical" explanations for language change, invoking no real theory of change and no "diachronic universals." In parallel fashion, I shall avoid a historicist approach to Universal Grammar, shunning principles that are motivated solely by some observed historical tendency (Lightfoot 1987). But I aim to show how the demands of language acquisition can shape the way in which languages change when certain changes occur in the triggering experience. That is not to say, of course, that all language change is to be explained in this fashion; that cannot be true if various languages change in different ways. But some very precise claims about the triggering experience will explain some old historical puzzles.

One of my claims is that the triggering experience consists not of raw data but of partially analyzed structures. Whenever one postulates non-trivial analyses, there are likely to be some theory-internal claims and a certain degree of technicality. That is true here. Moreover, in making claims about language acquisition based on evidence from language change and in using current work from syntactic theory, I am dogged by a concern for the compartmentalization of academic life and haunted by a fear that my argument will be buried in everybody's backlog of good intentions, forever piled unobtrusively in dark, unvisited corners. What can be done? Well, I have written for a heterogeneous audience, trying not to presuppose years of concern with the history of English, with the binding theory, or with acquisitional concerns, keeping my Maryland graduate students in mind. My readers must do me the favor of treating this as a specific discussion

of the triggering experience, and not as a comprehensive survey of language change or language acquisition.

Several readers and audiences have reminded me many times of the wide audience I'm trying to reach, forcing me to spell out an unstated assumption, to clarify a piece of jargon, and, it must be confessed, to correct an error here and there. I am grateful to the many people who have discussed the historical material in my *Principles of Diachronic Syntax*, to the "peerage" who took some of the central ideas on language acquisition seriously enough to write commentaries on a 1989 article in *Behavioral and Brain Sciences*, and to the audiences who have endured lectures on these topics in the United States, in Europe, and in Brazil. Science is an inherently cooperative enterprise, and I have benefited enormously from all these people. I hope I have not mistaken forbearance for enthusiasm, but I now feel ready to go into print and to face another group of reviewers; no doubt they will induce me to write more.

A fellowship from the American Council of Learned Societies and an award from the Graduate Research Board at the University of Maryland gave me a reprieve from chairing a young and vigorous department, enabling me to work where the telephone didn't ring. Research grant BNS-8812408 from the National Science Foundation supported the work reported in chapter 3. Kathi Faulkingham delicately transformed messy material into scripts neat enough for others to read, and resisted the temptation to press the "delete file" command in desperation. And now some old friends have helped in the final stages: I am greatly indebted to Peter Coopmans, Norbert Hornstein, and Anthony Warner, who read the whole manuscript from quite different angles and made enough constructive suggestions to earn a dinner at Sergio's. There they will no doubt convince me that I should have followed more of their advice.

Also welcome at that dinner will be anybody who publishes an unequivocally generous review.

---

# Contents

Preface ix

---

## Chapter 1

**Parameters and Triggers** 1

---

1.1

---

A Selective Theory of Language Acquisition 1

---

1.2

---

Arguments from the Poverty of the Stimulus 3

---

1.3

---

Negative Data 10

---

1.4

---

Not Every Experience Is a Trigger 13

---

1.5

---

Conclusion 19

---

## Chapter 2

**Simple Triggers** 22

---

2.1

---

Degree-0 Learnability 22

---

2.2

---

Bounding Nodes 24

---

2.3

---

Dutch Government 28

---

2.4

---

Chinese AGR 32

---

2.5

---

Italian Again 36



		2.6	
		Conclusion	38
<b>Chapter 3</b>		3.1	
<b>Loss of Object-Verb Order</b>	42	An Empirical Argument for Degree-0 Learnability?	42
		3.2	
		Acquiring Object-Verb Order	44
		3.3	
		Old English	56
		3.4	
		Reanalysis in English and Its Consequences	63
		3.5	
		Comparisons	72
		3.6	
		Conclusion	75
<b>Chapter 4</b>		4.1	
<b>Infinitives</b>	78	New Accusative Subjects and Passives	78
		4.2	
		Infinitives in Brazilian Portuguese	96
		4.3	
		Conclusion	103
<b>Chapter 5</b>		5.1	
<b>The English Case System</b>	104	An Inadequate Account of NP Movement	105
		5.2	
		A Theory of Abstract Case	111
		5.3	
		The History of NP Movement	115
		5.4	
		Triggering a New Abstract Case System	120

**Chapter 6**

**Obsolescence and Lexicalism** 126

6.1

Impersonal Verbs 128

6.2

Auxiliary Verbs 141

6.3

Conclusion 154

**Chapter 7**

**Chaos, Catastrophes, and  
Creoles** 157

7.1

Chaos and the Gradualness of  
Change 157

7.2

Catastrophes 166

7.3

Creoles 174

7.4

Conclusion 182

Notes 185

Bibliography 199

Index 211

#### 1.1 A Selective Theory of Language Acquisition

Linguists have traditionally maintained that language is not acquired by children only on the basis of experience; rather, children must themselves contribute something, if only an appropriate “disposition to learn.” As careful studies were conducted, “discovery procedures” and “analogical principles” of increasing complexity were postulated as part of the child’s contribution to the process of acquiring a language.

Over the last thirty years generative grammarians have been developing a selective theory of language acquisition. We have sought to ascertain what information must be available to children, independent of any experience with language, in order for the eventual mature linguistic capacities to emerge on exposure to some typical triggering experience. Cutting some corners, we have assumed that this unlearned information is genetically encoded in some fashion, and we have adopted (1) as our explanatory model.

- (1) a. trigger (genotype → phenotype)
- b. primary linguistic data (Universal Grammar → grammar)

The goal is to specify relevant aspects of a child’s genotype such that a particular mature state will emerge when a child is exposed to a certain triggering experience, depending on whether the child is raised in, say, a Japanese or a Navaho linguistic environment. (1b) reflects the usual terminology: “Universal Grammar” contains those aspects of the genotype that are directly relevant for language growth, and a “grammar” is taken to be a phenotypic property, a part of a person’s mental makeup that characterizes his or her mature linguistic capacity. The *primary* linguistic data are those data to which children are exposed and which actually

determine or “trigger” some aspect of their grammars, having some long-term effect.

The theory is “selective” in the same sense that current theories of immunology and vision are selective and not “instructive.” Under an instructive theory, an outside signal imparts its character to the system that receives it, instructing what is essentially a plastic and modifiable nervous system; under a selective theory, a stimulus may change a system that is already highly structured by identifying and amplifying some component of already available circuitry. Put differently: a selective theory holds that an organism experiences the surrounding environment and selects relevant stimuli according to criteria that are already present internally. Jerne (1967) depicts antibody formation as a selective process whereby the antigen selects and amplifies specific antibodies, which already exist. Similarly, Hubel and Wiesel showed that particular neurons were preset to react only to specific visual stimuli (for example, to a horizontal line); exposure to a horizontal line entails a radical increase in the number of horizontal-line receptors, and a horizontal line can be said to elicit and select specific responses within the organism. Changeux (1980, 1983) argues along similar lines for a theory of “selective stabilization of synapses” whereby “the genetic program directs the proper interaction between main categories of neurons.” “However,” he continues, “during development within a given category, several contacts with the same specificity may form,” and other elements, which are not selected, may atrophy (1980, p. 193). Thus, to learn is to amplify certain connections and to eliminate other possibilities (see also Mehler 1974 and Edelman 1987). Jerne (1967) argues as follows: “Looking back into the history of biology, it appears that wherever a phenomenon resembles learning, an instructive theory was first proposed to account for the underlying mechanisms. In every case, this was later replaced by a selective theory.” For more discussion, see Piattelli-Palmarini 1986 and Jerne 1985.

Under current formulations of grammatical theory, the linguistic genotype, Universal Grammar, consists of principles and parameters that are set by some linguistic environment, just as certain receptors are “set” on exposure to a horizontal line. So the environment may be said to “select” particular values for the parameters of Universal Grammar. Universal Grammar must be able to support the acquisition of any human grammar, given an appropriate triggering experience. Of course, Universal Grammar need not be seen as homogeneous; it may emerge piecemeal, parts of it being available only at certain stages of a child’s development. Grammars are not only attainable under normal childhood conditions; they are also

usable for such purposes as speech production and comprehension, they are vulnerable to the kinds of aphasias that are actually found, and one expects that they will provide part of the basis for understanding the developmental stages that children go through. There is no shortage of empirical constraints on hypotheses about the elements of (1).

## 1.2 Arguments from the Poverty of the Stimulus

The “logical problem of language acquisition” has provided much of the empirical refinement of (1). Apparent poverty-of-stimulus problems have led grammarians to postulate particular principles and parameters at the level of Universal Grammar. The stimulus or triggering experience that children have appears to be too poor to determine all aspects of the mature capacities that they typically attain. It is too poor in three distinct ways: (a) The child’s experience is finite, but the capacity eventually attained ranges over an infinite domain and therefore must incorporate some recursive property not demanded by experience. (b) The experience consists partly of degenerate data which have no effect on the emerging capacity (see section 1.4). (c) Most important, it fails to provide the data needed to induce many principles and generalizations manifested by the mature capacity. Of these three, (a) and (b) have been discussed much more frequently than (c), although (c) is by far the most significant factor and provides a means for elaborating theories of Universal Grammar, as I shall now illustrate.

Any argument from the poverty of the stimulus makes crucial assumptions about the nature of the triggering experience. One simple argument, which has been used frequently, concerns the relationship between a statement (e.g., *the book on the shelf is dull*) and a corresponding question (*is the book on the shelf dull?*). There is an operation that places *is* at the front in this particular example, but how is this operation to be stated? It might be stated in structure-dependent fashion, locating *is* to the right of a subject noun phrase and moving it over that noun phrase (here, *the book on the shelf*). Alternatively, the operation might be structure-independent, making no reference to structural notions such as *noun phrase* and being sensitive only to the sequence of words; such an operation might simply identify the first *is* and move it to the front. Both proposals are adequate for the simple cases, but a slightly more complex case necessitates the first option. In (2a) the structure-independent operation would move the first *is* to the front and yield the nonoccurring (2b) (where *e* indicates an empty position); the structure-dependent operation would identify *the book which*

*is on the shelf* as a noun phrase and would move *is* from the position immediately to its right, yielding—correctly—(2c).

- (2) a. the book which is on the shelf is dull.  
 b. \**is* the book which *e* on the shelf is dull?  
 c. *is* <sub>NP</sub>[the book which is on the shelf] *e* dull?

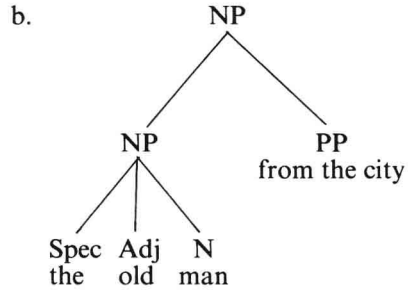
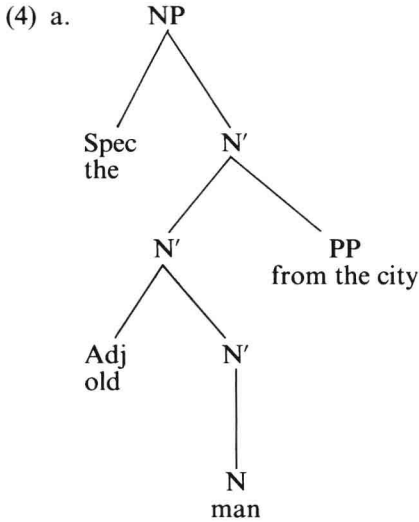
The crucial evidence for the structure-dependent formulation and for the nonavailability of structure-independent operations is the nonoccurrence of (2b). However, children are not systematically informed that certain forms do not exist, or that they are “ungrammatical,” and so the crucial evidence—the nonoccurrence of forms like (2b)—is not contained in normal childhood experience. Nonetheless, despite the lack of an environmental stimulus for structure dependence, and despite the simplicity of the structure-independent formulation, children invariably use the structure-dependent operation when first uttering questions of this kind; questions like (2b) simply do not occur and are not among the “errors” made by children. There seems to be no learning in this context. Moreover, whenever grammars have movement operations they are structure-dependent. If this is invariant and not deducible directly from childhood experience, it is reasonable to suppose that structure dependence is part of what the mind brings to the analysis of experience, not something hypothesized on the basis of evidence. One arrives at this conclusion by noting how the stimulus fails to determine certain aspects of mature grammars, and any such argument clearly depends on assumptions about the triggering experience—here, that the nonoccurrence of (2b) is not part of that experience. Furthermore, if the arguments to be offered in this book are correct, the relative clause in (2c) is also not part of the triggering experience.

Arguments from the poverty of the stimulus reveal not only gross properties of Universal Grammar, such as structure dependence, but also more fine-grained aspects of actual structures. To illustrate, I shall briefly rehearse an argument of Baker (1978), discussed by Hornstein and Lightfoot (1981) and then further by Lightfoot (1982). It has long been generally agreed that linguistic expressions are made up of subunits and have an internal hierarchical structure. It is also generally agreed that a grammar (in the sense defined) is not just a list of expressions but is a finite algebraic system that can “generate” an infinite range of expressions. One might imagine, in that case, that English noun phrases have the structure of either (3a) or (3b); proposals have been made along both lines. If a noun may project to a phrasal category in the manner defined by the phrase-structure rules of (3a), a phrase like *the old man from the city* will have the internal

structure of (4a). If the projections are along the lines of the rules in (3b) the structure will be (4b) with quite different subunits. In (4a) *the old man* does not form a single phrasal unit, but in (4b) it does. The crucial difference is that the rules of (3a) refer to  $N'$ , an element intermediate between the head noun (N) and the maximal phrasal projection (NP) of that noun.

- (3) a.  $NP \rightarrow \text{Specifier } N'$   
 $N' \rightarrow (\text{Adj}) [N \text{ or } N'] PP$

- b.  $NP \rightarrow NP PP$   
 $NP \rightarrow \text{Specifier (Adj) } N$



Now, it can be shown that any noun phrase that occurs in English, and thus any noun phrase that an English-speaking child is likely to hear, can be generated by both sets of projection types. However, linguists believe that something along the lines of (3a) must be correct, or at least preferred to (3b), because (3b) is consistent with certain phenomena that do *not* occur in English, unlike (3a). (3b) has no  $N'$  node, and therefore provides no straightforward way to distinguish between (5a) and (5b) and no ready means to capture the ambiguity of (6a), which may have the meaning of (6b) or (6c). The details of the analysis need not concern us here.<sup>1</sup>

- (5) a. \*the student of physics is older than the one of chemistry  
 b. the student from NY is older than the one from LA
- (6) a. he wants an old suit but he already has the only one I own  
 b. he wants an old suit but he already has the only suit I own  
 c. he wants an old suit but he already has the only old suit I own

What *is* relevant here is the following problem: It is reasonable to suppose that children might be exposed to any noun phrase that may occur in English, but it is not the case that they are systematically informed that sentences like (5a) are not uttered by most speakers and that (6a) has two possible meanings. In fact, “negative data” (i.e., information that certain data do not exist) are generally not available to children, and perception of ambiguity is a sophisticated skill that develops late and not uniformly; most ambiguities pass unnoticed, and people take the most appropriate of the available meanings. To be sure, children come to know these things, and this knowledge is part of the output of the language-acquisition process; but it is not part of the input, not part of the “evidence” for the emerging system, and thus not part of the triggering experience. Consequently, although linguists are able to choose hypothesis (3a) over (3b) on the basis of phenomena such as (5) and (6), children have no analogous basis for such a choice if such data are not available to them. It is in this sense that the stimulus is too impoverished to fully determine the emergent analysis. In that case children must arrive at (3a) on some other, presumably nonexperiential basis. As an account of this basis, linguists have postulated genotypical information that phrasal categories have the structure of (7). By (7a) any noun phrase (NP) consists of a Specifier (e.g. an article) and a subphrase N' in some order to be determined by the child's particular linguistic experience, the “trigger” of (1a). Similarly, a verb phrase (VP) consists of a Specifier and a V' in some order, and likewise the other phrasal categories. By (7b) the N' consists of a nucleus (N or N') and perhaps some satellite material in some order. (The curly brackets indicate an unordered set, and YP covers any phrasal category or a clause.<sup>2</sup>)

- (7) a.  $XP \rightarrow \{\text{Specifier}, X'\}$   
 b.  $X' \rightarrow \{X \text{ or } X', (YP)\}$

- (8) a. the house  
 b. students of linguistics, belief that Susan left

Under (7), the linear order of constituents constitutes a parameter that is set on exposure to some trigger. The English-speaking child hears phrases like (8a) and, after some development, analyzes them as consisting of two words, one of a closed class (*the*) and the other of an open class (*house*); in the light of this and in the light of the parameter in (7a), the child adopts the first rule of (3a). Likewise, exposure to phrases like (8b) suffices to set the parameter in (7b), such that the second rule of (3a) emerges.<sup>3</sup> Because of the parameters in (7), rules like those of (3b) are never available to children and therefore do not have to be “unlearned” in any sense. Al-



though no “evidence” for the existence of a phrasal category N' seems to be available in a child's experience, it is provided by the genotype and therefore it occurs in mature grammars. (I shall consider an alternative account later.)

Consider, for a moment, the development that must take place before these parameters can be set. Children acquire the sounds of their languages and come to use *men* as a word and a noun with the meaning roughly of the plural of ‘man’. This is a nontrivial process, and many people have examined how it happens. Having established that *men* is a noun, children later acquire the constituent structure of *men from the city*, if I am right, by setting the parameters in (7) and projecting to NP accordingly via N', yielding

NP[Spec<sub>N'</sub>[N'<sub>N</sub>[N[men]]] PP[from the city]].

Lebeaux (1988) discusses this aspect of language acquisition very interestingly. In setting these particular parameters, children operate with partially formed representations that include N[men], P[from], Spec[the], and N[city]. They are operating not with “raw data” or mere words but with partially analyzed structures.

*Men from the city* and similar expressions occur in the child's environment with an appropriate frequency, and, given a partially formed grammar whereby *men* and *city* are classified as nouns, a child can assign a projection conforming to (7). Contrast this with some ill-formed expressions that a child might encounter for various reasons: *from the city men*, uttered by a guest who speaks a head-final language, is not sufficiently frequent to have any effect; a mixed form such as *men uit de stad*, uttered by a Dutch house guest, could not be analyzed by means of a partially formed grammar in which *uit* is not classified. Of course, if the Dutch guest stayed long enough and spoke often enough about city people, the child might come to classify *uit*, *de*, and *stad* as P, Spec, and N respectively, with the relevant meanings, might interpret the string correctly as a phrasal satellite of *men*, and might even use such phrases; in that case, one would witness an instance of code switching, a frequent phenomenon among children exposed to more than one language. If there is a great deal of this kind of mixture in the trigger experience, and if a pidgin is involved, the child will attain some form of creole. This extreme case reflects the quite general capacity to operate with a heterogeneous grammar.

There is much more to be said about NP structure and about its consequences. I have sketched the argument briefly here in order to demonstrate that a poverty-of-stimulus argument is based on assumptions about