Language Acquisition and the Form of the Grammar

David Lebeaux

LANGUA SITION AND THE FORM OF THE GRAMMAR

DAVID LEBEAUX
NEC Research Institute

JOHN BENJAMINS PUBLISHING COMPANY
PHILADELPHIA/AMSTERDAM

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Lebeaux, David.
Language acquisition and the form of the grammar / David Lebeaux
p. cm.
Includes bibliographical references and index.
1. Language acquisition. 2. Generative grammar. I. Title.
P118.L38995 2000
401'.93--dc21 00-039775
ISBN 90 272 2565 6 (Eur.) / 1 55619 858 2 (US)

John Benjamins Publishing Co. • P.O.Box 75577 • 1070 AN Amsterdam • The Netherlands
John Benjamins North America • P.O.Box 27519 • Philadelphia PA 19118-0519 • USA

There are two ways of painting two trees together. Draw a large tree and add a small one; this is called *fu lao* (carrying the old on the back). Draw a small tree and add a large one; this is called *hsieh yu* (leading the young by the hand). Old trees should show a grave dignity and an air of compassion. Young trees should appear modest and retiring. They should stand together gazing at each other.

Mai-mai Sze
The Way of Chinese Painting

Acknowledgments

This book had its origins as a linguistics thesis at the University of Massachusetts. First of all, I would like to thank my committee: Tom Roeper, for scores of hours of talk, for encouragement, and for his unflagging conviction of the importance of work in language acquisition; Edwin Williams, for the example of his work; Lyn Frazier, for an acute and creative reading; and Chuck Clifton, for a psychologist's view. More generally, I would like to thank the faculty and students of the University of Massachusetts, for making it a place where creative thinking is valued. The concerns and orientation of this book are very much molded by the training that I received there.

Further back, I would like to thank the people who got me interested in all of this in the first place: Steve Pinker, Jorge Hankamer, Jane Grimshaw, Annie Zaenen, Merrill Garrett and Susan Carey. I would also like to thank Noam Chomsky for encouragement throughout the years.

Since the writing of the thesis, I have had the encouragement and advice of many fine colleagues. I would especially like to thank Susan Powers, Alan Munn, Cristina Schmitt, Juan Uriagereka, Anne Vainikka, Ann Farmer, and Ana-Teresa Perez-Leroux. I am also indebted to Sandiway Fong, as well as Bob Krovetz, Christiane Fellbaum, Kiyoshi Yamabana, Piroska Csuri, and the NEC Research Institute for a remarkable environment in which to pursue the research further.

I would also like to thank Mamoru Saito, Hajime Hoji, Peggy Speas, Juergen Weissenborn, Clare Voss, Keiko Muromatsu, Eloise Jelinek, Emmon Bach, Jan Koster, and Ray Jackendoff.

Finally, I would like to thank my parents, Charles and Lillian Lebeaux, my sister, Debbie Lebeaux, and my sons, Mark and Theo. Most of all, I would like to thank my wife Pam, without whom this book would have been done badly, if at all. This book is dedicated to her, with love.

Preface

What is the best way to structure a grammar? This is the question that I started out with in the writing of my thesis in 1988. I believe that the thesis had a marked effect in its answering of this question, particularly in the creation of the Minimalist Program by Chomsky (1993) a few years later.

I attempted real answers to the question of how to structure a grammar, and the answers were these:

(i) In acquisition, the grammar is arranged along the lines of *subgrammars*. These grammars are arranged so that the child passes from one to the next, and each succeeding grammar contains the last. I shall make this clearer below.

(ii) In addition, in acquisition, the child proceeds to construct his/her grammar from *derivational endpoints* (Chapter 5). From the derivational endpoints, the child proceeds to construct the entire grammar. This may be forward or backward, depending on what the derivational endpoint is. If the derivation endpoint, or anchorpoint, is DS, then the construction is forward; if the derivational endpoint or anchorpoint is S-structure or the surface, then the construction proceeds backwards.

The above two proposals were the main proposals made about the acquisition sequence. There were many proposals made about the syntax. Of these, the main architectural proposals were the following.

(iii) The acquisition sequence and the syntax — in particular, the syntactic derivation — are not to be considered in isolation from each other, but rather are tightly yoked. The acquisition sequence can be seen as the result of derivational steps or subsequences (as can be seen in Chapter 2, 3, and 4). This means that the acquisition sequence gives unique purchase onto the derivation itself, including the adult derivation.

(iv) Phrase structure is not given as is, nor is derived top-down, but rather is composed (Speas 1990). This phrase structure composition (Lebeaux 1988), is not strictly bottom up, as in Chomsky's (1995) Merge, but rather involves

(a) the intermingling or units, (b) is grammatically licensed, and not simply geometrical (bottom-up) in character (in a way which will become clearer below), and (c) involves, among other transformations, the transformation Project-α (Chapter 4).

(v) Two specific composition operations (and the beginnings of a third) are proposed. Adjoin-α (Chapter 3) is proposed, adding adjuncts to the basic nuclear clause structure (Conjoin-α is also suggested in that chapter). In further work, this is quite similar to the Adjunction operation of Joshi and Kroch, and the Tree Adjoining Grammars (Joshi 1985; Joshi and Kroch 1985; Frank 1992), though the proposals are independent and the proposals are not exactly the same. The second new composition operation is Project-α (Chapter 4), which is an absolutely new operation in the field. It projects open class structure into a closed class frame, and constitutes the single most radical syntactic proposal of this book.

(vi) Finally, composition operations, and the variance in the grammar as a whole, are linked to the closed class set — elements like *the*, *a*, *to*, *of*, etc. In particular, each composition operation requires the satisfaction of a closed class element; as well as a closed class element being implicated in each parameter.

These constitute some of the major proposals that are made in the course of this thesis. In this preface I would like to both lay out these proposals in more detail, and compare them with some of the other proposals that have been made since the publication of this thesis in 1988. While this thesis played a major role in the coming of the Minimalist Program (Chomsky 1993, 1995), the ideas of the thesis warrant a renewed look by researchers in the field, for they have provocative implications for the treatment of language acquisition and the composition of phrase structure.

Let us start to outline the differences of this thesis with respect to later proposals, not with respect to language acquisition, but with respect to syntax. In particular, let us start with parts (iv) and (v) above: that the phrase marker is composed from smaller units.

A similar proposal is made with Chomsky's (1995) Merge. However, here, unlike Merge:

(1) The composition is not simply bottom-up, but involves the possible intermingling of units.

(2) The composition is syntactically triggered in that all phrase structure composition involves the satisfaction of closed class elements

(Chapters 3 and 4), and is not simply the geometric putting together of two units, as in Merge, and

(3) The composition consists of two operations among others (these are the only two that are developed in this thesis), Adjoin-α and Project-α.

With respect to the idea that all composition operations are syntactically triggered by features, let us take the operation Adjoin-α. This takes two structures and adjoins the second into the first.

(1)

s1:	the man met the woman	Adjoin-α →	the man met the woman
s2:	who loved him		who loved him

This shows the intermingling of units, as the second is intermeshed with the first. However, I argue here (Chapter 4), that it also shows the satisfaction of closed class elements, in an interesting way. Let us call the *wh*-element of the relative clause, *who* here, the relative clause linker.

It is a proposal of this thesis that the adjunction operation itself involves the satisfaction of the relative clause linker (*who*), by the relative clause head (*the woman*), and it is this relation, which is the relation of Agreement, which composes the phrase marker. The relative clause linker is part of the closed class set. This relative clause linker is satisfied in the course of Agreement, thus the composition operation is put into a 1-to-1 relation with the satisfaction of a closed class head. (This proposal, so far as I know, is brand new in the literature).

(2) Agree Relative head/relativizer ↔ Adjoin-α

This goes along with the proposal (Chapter 4), which was taken up in the Minimalist literature (Chomsky 1992, 1995), that movement involves the satisfaction of closed class features. The proposal here, however, is that composition, as well as movement, involves the satisfaction of a closed class feature (in particular, Agreement). In the position here, taken up in the Minimalist literature, the movement of an element to the subject position is put into a 1-to-1 correspondence with agreement (Chapter 4 again).

(3) Agree Subject/Predicate ↔ Move NP (Chapter 4)

The proposal here is thus more thoroughgoing than that in the minimalist literature, in that *both* the composition operation, and the movement operation are triggered by Agreement, and the satisfaction of closed class features. In the minimalist literature, it is simply movement which is triggered by the satisfaction

of closed class elements (features); phrase structure composition is done simply geometrically (bottom-up). Here, both are done through the satisfaction of Agreement. This is shown below.

(4)

	Minimalism	Lebeaux (1988)
Movement	syntactic (satisfaction of features)	syntactic (satisfaction of features)
Phrase Structure Composition	asyntactic (geometric)	syntactic (satisfaction of features)

This proposal (Lebeaux 1988) links the entire grammar to the closed class set — both the movement operations and the composition operations are linked to this set.

The set of composition operations discussed in this thesis is not intended to be exhaustive, merely representative. Along with Adjoin-α which Chomsky-adjoins elements into the representation (Chapter 3), let us take the second, yet more radical phrase structure composition operation, Project-α. This is not equivalent to Speas' (1990) Project-α, but rather projects an open class structure into a closed class frame. The open class structure also represents pure thematic structure, and the closed class structure, pure Case structure.

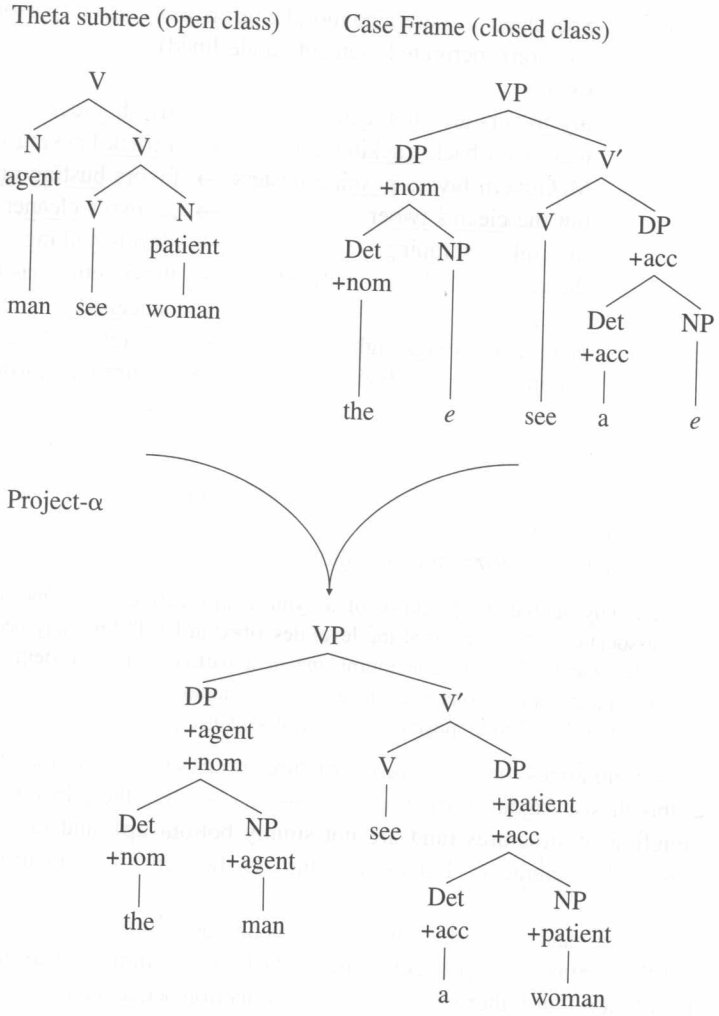
This operation, for a simple partial sentence, looks like (5) (see Lebeaux 1988, 1991, 1997, 1998 for further extensive discussion).

The operation projects the open class elements into the closed class (Case) frame. It also projects up the Case information from Determiner to DP, and unifies the theta information, from the theta subtree, into the Case Frame, so that it appears on the DP node.

The Project-α operation was motivated in part by the postulation of a subgrammar in acquisition (Chapters 2, 3, and 4), in part by the remarkable speech error data of Garrett (Chapter 4, Garrett 1975), and in part by idioms (Chapter 4). This operation is discussed at much greater length in further developments by myself (Lebeaux 1991, 1997, 1998).

I will discuss in more detail about the subgrammar underpinnings of the Project-α approach later in this preface. For now, I would simply like to point to the remarkable speech error data collected by Merrill Garrett (1975, 1980), the MIT corpus, which anchors this approach.

(5)



Garrett and Shattuck-Hufnagel collected a sample of 3400 speech errors. Of these, by far the most interesting class is the so-called "morpheme-stranding" errors. These are absolutely remarkable in that they show the insertion of open class elements into a closed class frame. Thus, empirically, the apparent "importance" of open class and closed class items is reversed — rather than open class items being paramount, closed class items are paramount, and guide the derivation. Open class elements are put into slots provided by closed class elements, in Garrett's remarkable work. A small sample of Garrett's set is shown below.

(6) Speech errors (stranded morpheme errors), Garrett (personal communication) (permuted elements underlined)

Error		Target
my frozers are shoulden	→	my shoulders are frozen
that just a back trucking out	→	a truck backing out
McGovern favors pushing busters	→	favors busting pushers
but the clean's twoer	→	… two's cleaner …
his sink is shipping	→	ship is sinking
the cancel has been practiced	→	the practice has been cancelled
she's got her sets sight	→	… sights set …
a puncture tiring device	→	… tire puncturing device …

As can be seen, these errors can only arise at a level where open class elements are inserted into a closed class frame. The insertion does not take place correctly — a speech error — so that the open class elements end up in permuted slots (e.g. a *puncture tiring device*).

Garrett summarizes this as follows:

> … why should the presence of a syntactically active bound morpheme be associated with an error at the level described in [(6)]? Precisely because the attachment of a syntactic morpheme to a particular lexical stem reflects a mapping from a "functional" level [i.e. "grammatical functional", i.e. my theta subtree, D.L.] to a "positional" level of sentence planning …

This summarizes the two phrase structure composition operations that I propose in this thesis: Adjoin-α and Project-α. As can be seen, these involve (1) the intermingling of structures (and are not simply bottom up), and (2) satisfaction of closed class elements. Let us now turn to the general acquisition side of the problem.

It was said above that this thesis was unique in that the acquisition sequence and the syntax — in particular, the syntactic derivation — were not considered in isolation, but rather in tandem. The acquisition sequence can be viewed as the output of derivational processes. Therefore, to the extent to which the derivation is partial, the corresponding stage of the acquisition sequence can be seen as a subgrammar of the full grammar. The yoking of the acquisition sequence and the syntax is therefore the following:

(7) ACQUISITION subgrammar approach
SYNTAX phrase structure composition from smaller units

The subgrammar approach means that children literally have a smaller grammar than the adult. The grammar increases over time by adding new structures (e.g. relative clauses, conjunctions), and by adding new primitives of the representational vocabulary, as in the change from pure theta composed speech, to theta and Case composed speech.

The addition of new structures — e.g. relative clauses and conjunctions — may be thought of as follows. A complex sentence like that in (8) may be thought of as a triple: the two units, and the operation composing them (8b).

(8) a. The man saw the woman who loved him.
b. (the man saw the woman (rooted), who loved him, Adjoin-α)

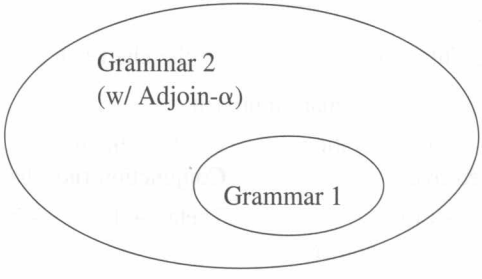
Therefore a subgrammar, if it is lacking the operation joining the units may be thought of as simply taking one of the units — let us say the rooted one — and letting go of the other unit (plus letting go of the operation itself). This is possible and necessary because it is the operation itself which joins the units: if the operation is not present, one or the other of the units must be chosen. The subgrammar behind (8a), but lacking the Adjoin-α operation, will therefore generate the structure in (9) (assuming that it is the rooted structure which is chosen).

(9) The man saw the woman.

This is what is wanted.

Note that the subgrammar approach (in acquisition), and the phrase structure composition approach (in syntax itself) are in perfect parity. The phrase structure composition approach gives the actual operation dividing the subgrammar from the supergrammar. That is, with respect to this operation (Adjoin-α), the grammars are arranged in two circles: Grammar 1 containing the grammar itself, but without Adjoin-α, and Grammar 2 containing the grammar including Adjoin-α.

(10)



The above is a case of adding a new operation.

The case of adding another representational primitive is yet more interesting.

Let us assume that the initial grammar is a pure representation of theta relations. At a later stage, Case comes in. This hypothesis is of the "layering of vocabulary": one type of representational vocabulary comes in, and does not displace, but rather is added to, another.

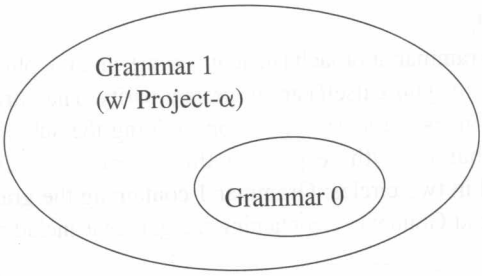
(11) theta → theta + Case
Stage I **Stage II**

The natural lines along which this representational addition takes place is precisely given by the operation Project-α. The derivation may again be thought of as a triple: the two composing structures, one a pure representation of theta relations, and one a pure representation of Case, and the operation composing them.

(12) ((man (see woman)), (the __ (see (a __))), Project-α)
the "sees" in theta tree and Case frame each contain partial information which is unified in the Project-α operation.

The subgrammar is one of the two representational units: in this case, the unit (man(see woman)). That is a sort of theta representation or telegraphic speech. The sequence from Grammar 0 to Grammar 1 is therefore given by the addition of Project-α.

(13)



The full pattern of stage-like growth is shown in the chart below:

(14) ACQUISITION: Subgrammar Approach

Add construction operations to simplified tree	Relative clauses, Conjunction (not discussed here)
Add primitives to representational vocabulary	Theta → Theta + Case

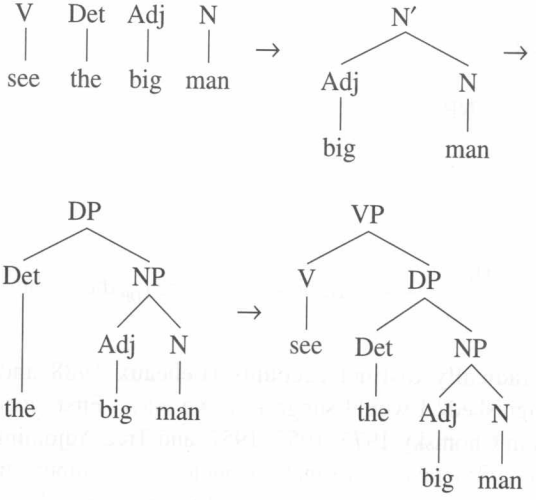
As can be seen, the acquisition sequence and the syntax — syntactic derivation — are tightly yoked.

Another way of putting the arguments above is in terms of distinguishing

accounts. I wish to distinguish the phrase structure operations here from Merge; and the acquisition subgrammar approach here from the alternative, which is the Full Tree, or Full Competence, Approach (the full tree approach holds that the child does not start out with a substructure, but rather has the full tree, at all stages of development.) Let us see how the accounts are distinguished, in turn.

Let us start with Chomsky's Merge. According to Merge, the (adult) phrase structure tree, as in Montague (1974), is built up bottom-up, taking individual units and joining them together, and so on. The chief property of Merge is that it is *strictly* bottom-up. Thus, for example, in a right-branching structure like "see the big man", Merge would first take *big* and *man* and Merge them together, then add *the* to *big man*, and then add *see* to the resultant.

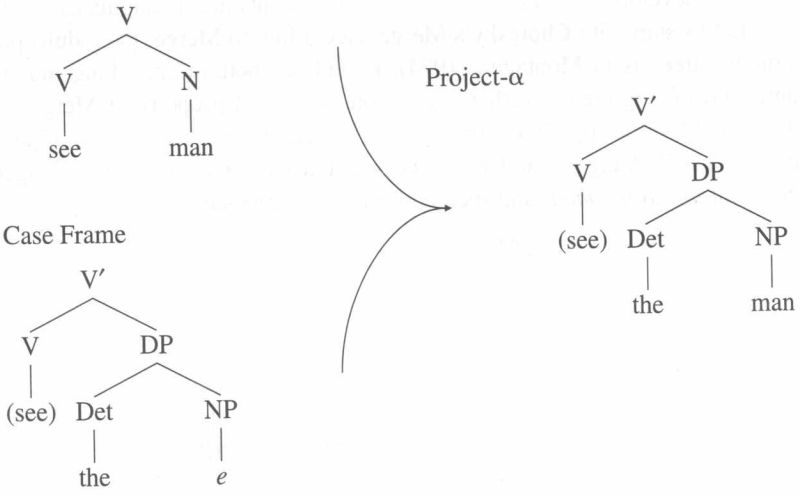
(15) Application of Merge:



The proposal assayed in this thesis (Lebeaux 1988) would, however, have a radically different derivation. It would take the basic structure as being the basic government relation: (*see man*). This is the primitive unit (unlike with Merge). To this, the *the* and the *big* may be added, by separate transformations, Project-α and Adjoin-α, respectively.

(16)

a. Project-α



b. Adjoin-α

($_{V'}$ see ($_{DP}$ the man))
($_{ADJ}$ big)
→ Adjoin-α → ($_{V'}$ see ($_{DP}$ the ($_{NP}$ big man)))

How can these radically distinct accounts (Lebeaux 1988 and Merge) be empirically distinguished? I would suggest in two ways. First, conceptually the proposal here (as in Chomsky 1975–1955, 1957, and Tree Adjoining Grammars, Kroch and Joshi 1985) takes information nuclei as its input structures, not arbitrary pieces of string. For example, for the structure "The man saw the photograph that was taken by Stieglitz", the representation here would take the two clausal nuclear structures, shown in (17) below, and adjoin them. This is not true for Merge which does not deal in nuclear units.

(17) s1: the man saw the photograph
s2: that was by Stieglitz
→ Adjoin-α → the man saw the photograph that was by Stieglitz

Even more interesting nuclear units are implicated in the transformation Project-α, where the full sentence is decomposed into a nuclear unit which is the theta subtree, and the Case Frame.

(18)

The structure in (18), *the man saw the woman*, is composed of a basic nuclear unit, (man (see woman)), which is telegraphic speech (as argued for in Chapter 2). No such nuclear unit exists in the Merge derivation of "the man saw the woman": that is, in the Merge derivation, (man (see woman)) does not exist as a substructure of ((the man) (saw (the woman)).

This is the conceptual argument for preferring the composition operation here over Merge. In addition, there are two simplicity arguments, of which I will give just one here.

The simplicity argument has to do with a set of structures that children produce which are called *replacement sequences* (Braine 1976). In these sequences, the child is trying to reach (output) some structure which is somewhat too difficult for him/her. To make it, therefore, he or she first outputs a substructure, and then the whole structure. Examples are given below: the first line is the first outputted structure, and the second line is the second outputted structure, as the child attempts to reach the target (which is the second line).

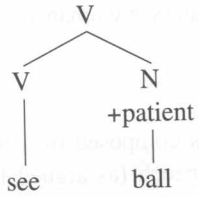
(19) see ball (first output)
see big ball (second output and target)

(20) see ball (first output)
see the ball (second output and target)

What is striking about these replacement sequences is that the child does not simply first output random substrings of the final target, but rather that the first output is an organized part of the second. Thus in both (19) and (20), what the child has done is first isolate out the basic government relation, (see ball), and then added to it: with "big" and "the", respectively.

The particular simplifications chosen are precisely what we would expect with the substructure approach outlined here, and crucially not with Merge. With the substructure approach outlined here (Chapter 2, 4), what the child (or adult) first has in the derivation is precisely the structure (see ball), shown in example (21).

(21)



To this structure is then added other elements, by Project-α or Adjoin-α. Thus, crucially, the first structure in (19) and (20) actually exists as a literal substructure of the final form — line 2 — and thus could help the child in deriving the final form. It literally goes into the derivation.

By contrast, with Merge, the first line in (19) and (20) *never underlies the second line*. It is easy to see why. Merge is simply bottom-up — it extends the phrase marker. Therefore, the phrase structure composition derivation underlying (20) line 2, is simply the following (Merge derivation).

(22) Merge derivation underlying (20) line 2
($_N$ ball)
($_{DP}$ ($_D$ the) ($_N$ ball))
(see ($_{DP}$ ($_D$ the) ($_N$ ball)))

However, this derivation crucially does not have the first line of (20) — (see (ball)) — as a subcomponent. That is, (see (ball)) does not go into the making of (see (the ball)), in the Merge derivation, but it does in the substructure derivation.

But this is a strong argument against Merge. For the first line of the outputted sequence of (20), (see ball), is presumably helping the child in reaching the ultimate target (see (the ball)). But this is impossible with Merge, for the first line in (20) does not go into the making of the second line, *according to the Merge derivation*.

That is, Merge cannot explain why (see ball) would help the child get to the target (see (the ball)), since (see ball) is not part of the derivation of (see (the ball)), in the Merge derivation. It is part of the sub-derivation in the substructure approach outlined here, because of the operation Project-α.

The above (see Chapters 2, 3, and 4) differentiates the sort of phrase structure composition operations found here from Merge. This is in the domain of syntax — though I have used language acquisition argumentation. In the domain of language acquisition proper, the proposal of this thesis — the hypothesis of substructures — must be contrasted with the alternative, which holds that the child is outputting the full tree, even when the child is potentially just in the one word stage: this may be called the Full Tree Hypothesis. These