GRAMMATICAL THEORY AND LANGUAGE ACQUISITION

Lydia White

A dissertation submitted to the Graduate Faculty, McGill University, in partial fulfillment of the requirements for the degree of Ph.D. in Linguistics

Department of Linguistics
McGill University
October 1979

ABSTRACT

GRAMMATICAL THEORY AND LANGUAGE ACQUISITION

Lydia White

This thesis examines the interaction of a particular theory of grammar, the extended standard version of generative grammar, with language acquisition. It discusses the kind of explanation of acquisition that a restrictive theory of grammar can offer and the kinds of prediction that it can be expected to make, as well as ways in which consideration of the facts of acquisition must shape proposals for the theory. Acquisitional data are considered as part of the data base to which the theory of grammar is responsible and the use of such data to argue for particular grammars or the form of grammars in general is discussed. A number of issues, such as the psychological reality of child grammars, the question of their optimality, and the types of change possible in child language, are re-examined. It is suggested that a lack of coherence in these areas has caused serious confusion in the past.

RESUME

GRAMMATICAL THEORY AND LANGUAGE ACQUISITION

LA THEORIE DE LA GRAMMAIRE ET L'ACQUISITION DU LANGAGE

Lydia White

Cette thèse examine l'interaction entre une théorie particulière de la grammaire, la version standard étendue de la grammaire générative, et l'acquisition du langage. On y discute le genre d'explication qu'une théorie restrictive de la grammaire peut fournir pour l'acquisition, de même que le genre de prédictions qu'elle peut être susceptible de faire. On y discute également comment l'étude des faits d'acquisition doit être en mesure de formuler des propositions pour la théorie. Les données d'acquisition sont considerées comme faisant partie des données dont la theorie de la grammaire doit rendre compte. L'utilisation de telles données pour appuyer les grammaires particulières ou la forme des grammaires en général y est discutée. Certaines questions sont reconsidérées, telles que la realité psychologique des grammaires des enfants, leur optimalité, et le type de changement possible dans le langage des enfants. Il y est suggéré qu'un manque de cohésion dans ces domaines a été autrefois la cause de beaucoup de confusion.

PREFACE

While there has long been an interest in the relevance of linguistic theory to acquisition, and vice versa, this has involved very little critical consideration of the nature of the interaction between these two areas. This thesis will examine exactly what explanation of and predictions for acquisition can be expected from a restrictive theory of grammar (namely the extended standard version of generative grammar) and what falls outside the scope of such a theory, as well as looking at ways in which acquisition data should and should not be used in support of proposals for particular grammars or for the theory of grammar.

Many assumptions are implicit in the literature which lead to contradictory proposals. I shall attempt to make these assumptions explicit and to show which are justified and which are not. Where some consideration has been given in the past to the relationship of grammatical theory to acquisition, the apparent failure of certain predictions of the theory of grammar, a failure due to a misunderstanding of what can be expected of linguistic theory in this area, has led either to a total rejection of any direct interaction at all, or to the uncritical separation of concepts such as psychological reality and optimality. This thesis will

investigate the consequences of such moves and show that they are not only mistaken but that, in the latter case, while the intention is to maintain a direct link between grammatical theory and acquisition, the effect is quite the opposite, making it impossible to determine what acquisition data are data of. Such issues appear not to have been thoroughly or critically discussed in the literature and I hope that by attempting some clarification here the way will be open for a fruitful association of grammatical theory and acquisition.

I should particularly like to thank my thesis supervisor, David Lightfoot, for his encouragement and suggestions, and also the following people for comments and discussion on various parts of this thesis or on earlier papers closely related to the issues raised here: N. Domingue, N. Hornstein, M-L. Kean, E. Matthei, I. Mazurkewich, M. Paradis, G. Piggott, R. Tucker.

Research for this thesis was supported by a McGill University McConnell Fellowship and a Government of Québec Ministry of Education Bursary, for which I am grateful.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

1.1. The problem of acquisition

The theory of grammar proposed by Chomsky (1965 and subsequently) seeks to explain how children can master their native language. The problem to be accounted for has often been described: on being exposed to a finite amount of data from any particular language, children become able to understand and produce a potentially infinite set of sentences of that language, in a short space of time. Two characteristics of the primary data, the data to which children are exposed, are crucial for an understanding of the problem, since they show that such data are "deficient" for any account of language learning that would rely solely on induction. On the one hand, children may be exposed to performance errors in the form of false starts, slips of the tongue, hesitations, and so on, and the language learner has no means of knowing in advance what proportion of such errors to expect or when or where they might occur. They are, therefore, a potential source of difficulty (though perhaps the problems caused by this kind of degenerate data have been overestimated, as will be discussed in Chapter 3).

On the other hand, and more significantly, the input data is deficient in that it lacks direct evidence as to ungrammaticality, ambiguity or paraphrase relations. Information of this kind forms part of the non-primary data available to the linguist in constructing grammars. It is not available to the child and he nevertheless constructs a grammar that includes knowledge of such relationships. A grammar of competence must include knowledge of ungrammaticality, ambiguity and paraphrase relations, as children's and adult's speech behaviour shows. That is, children do not construct a grammar solely for the primary data but one which encompasses non-primary data as well, a fact which cannot be explained if the primary data is the only source of their knowledge.[1]

How, then, do children acquire a particular grammar in such circumstances? According to the theory of grammar presupposed in this thesis, children do not have to rely on the input data alone in order to acquire a language; rather, they have available certain a priori principles which guide grammar construction in particular ways. That is, the principles of grammar, or universal grammar, are innate; they are part of the biological endowment that the child brings to the acquisition task. The content of universal grammar is an empirical question, constantly under review. But the acceptance of some a priori principles, whatever they may be, answers the logical problem of acquisition, in that it offers an explanation of how acquisition is possible at all, given the shortcomings of the primary data.

But while the acceptance of certain principles may explain how acquisition is possible, many problems remain. Of particular concern is the "projection problem" (Peters, 1972; Baker, 1979a). Even assuming certain innately specified principles, the theory of grammar has not been sufficiently restrictive to allow one to predict a particular grammar given a particular set of data. The ideal interaction of data and grammar has been represented as follows (Chomsky, 1967; McNeill, 1970):

1.1. data ------> Language acquisition device (L.A.D.) ---> grammar

That is:

> "It is logically possible that the data might be sufficiently rich and the class of potential grammars sufficiently limited so that no more than a single permitted grammar will be compatible with the available data at the moment of successful acquisition."
>
> (Chomsky, 1965, p. 36)

However, this ideal is far from being attained. Given a set of primary data, it is often possible to construct several grammars which are compatible with that data. Thus, the problem for the child is sometimes seen as having to select one out of several possible grammars:

1.2. data ------> L.A.D. ------> grammar 1
grammar 2
grammar n

Here one needs to be quite clear on the difference between

the task facing the child and that facing the linguist. It is certainly a problem for the linguist if linguistic theory cannot limit the number of grammars consistent with the data but this does not mean that the child actually compares alternatives. One should be wary of imputing the shortcomings of the linguist to the child; the child may well have the means to construct one grammar on the basis of universal principles and primary data, although we have not yet discovered the means to do so.

The choice between competing grammars has traditionally been made by an evaluation metric, often conceived of in terms of simplicity, which has proved easier to define intuitively than formally. A more fruitful way of choosing between grammars has been to remove some of the choice altogether, by adopting a restrictive theory of grammar, such that universal grammar, the a priori principles imputed to the L.A.D., will not permit the range of grammars which is currently possible. Recently, interest has been concentrated on ways to restrict the theory (see Chapter 2), thus bringing us closer to the ideal of (1.1). The aim of linguistic theory, then, is to discover the principles that the child already has which makes such an interaction between the data and the L.A.D. possible.

In many cases, apparently competing grammars for the same data will turn out to be non-equivalent. For example, in Chapters 3 and 4, grammars will be considered which are

equally compatible with input data but where one grammar fails to account for knowledge of ungrammaticality, ambiguity or paraphrase relations, which must be embodied in a grammar of competence. The number of acceptable grammars is thus further limited if one takes into account exactly what a grammar must be a grammar of. The more one can restrict the number of grammars available in various ways, the less it will be necessary to appeal to a simplicity metric to choose between them.[2]

Thus, trying to answer the problem of acquisition involves discovering principles or restrictions that the child must be assumed to have innately. A failure to separate the linguist's task from the child's can make the problem seem even more complicated than it is. The projection problem is the linguist's problem, not the child's. We must assume that he already has the answer to it, that grammatical principles severely constrain the possibilities open to him.

1.2. Some misconceptions

Since the theory of grammar aims to explain the fact of acquisition by means of restrictive a priori principles, it has sometimes been assumed that the theory is itself a theory of acquisition (e.g. Derwing, 1973). But though it may be the aim of the theory of grammar to characterize the content of the "language acquisition device", it is in no sense a theory of how acquisition proceeds in "real time", it does not

provide a performance model for acquisition (or for any other aspect of speech behaviour).

Nevertheless, despite injunctions to the contrary (e.g. Chomsky, 1967), there was a tendency in the early psycholinguistic literature to assume that the theory of transformational generative grammar (T.G.G.) would offer not just a model of competence but some kind of performance model as well. The publications of Chomsky (1957, 1965), and the works which these stimulated, led to a search for insights that linguistic theory could offer to the study of acquisition. Initially, this took the form of trying to determine the child's underlying grammar and using T.G.G. models to describe his output.[3] In many cases, it was assumed that the division between the phrase-structure and transformational components of the grammar would be mirrored by a similar division in real-time acquisition; that is, that it would be reasonable to expect the child to proceed first by "speaking underlying structure directly" (McNeill, 1970). Klima and Bellugi (1966), McNeill (1966, 1970), Brown, Cazden and Bellugi (1969) all make similar assumptions: that the child first speaks base-strings, or something close to them, and only later develops transformational rules. As McNeill (1970) puts it "language acquisition consists of learning transformations".[4] But there is nothing whatsoever in the theory of grammar to support these proposals. While the theory distinguishes between the base and the transformational components of a grammar, it

makes no claims about the order of their acquisition and there is no reason, as far as the theory of grammar is concerned, why phrase-structure rules and transformations should not be learned together. This was recognized by Bloom (1970), Bowerman (1973), Roeper (1973) who all propose analyses that involve transformations concurrent with phrase-structure rules in early speech. Whether or not their particular proposals are correct is not at issue here; what is important is that they recognize that the theory of grammar does not dictate any precedence in the learning of phrase-structure rules.

This early concentration on transformations in developmental psycholinguistics continued with proposals for the derivational theory of complexity (D.T.C.), where, again, the grammar is taken directly as a performance model, the proposal being that there is a one-to-one correspondence between transformational rules and psychological operations (see Chapter 5). Not surprisingly, with T.G.G. interpreted directly as a performance model, or, failing that, with no very clear picture of how performance factors should interact with a grammar of competence, there was some confusion as to what predictions could be expected from such a theory and when predictions were made, they did not seem to be borne out. With the apparent failure to find any close links between the predictions of the theory of grammar and the facts of acquisition, there was a move to deny a role to the theory of grammar in the study of acquisition and to replace it by emphasis on general

cognitive and inductive principles, perceptual and learning strategies and the like.[5]

1.3. The responsibility of the theory of grammar to acquisition.

This move away from the theory of grammar as a source of interest is at least partly based on misconceptions as to what grammatical theory can be expected to say about acquisition. Since the theory of grammar is not a performance model, it alone cannot be expected to make predictions about acquisition in real time, though it will interact with theories of maturation, cognitive development, parsing, etc., to provide proposals about the course of language development. For example, in Chapter 8, I shall examine markedness proposals and see how they affect the course of acquisition, on the assumption that markedness is a form of complexity and that complexity generally causes problems in acquisition (and not just in the field of language). Such proposals will only be as good as the theories of the other related domains. A theory of grammar cannot be responsible for unsatisfactory proposals which result from a poor theory of processing, for instance, as was the case with the D.T.C., where naive assumptions about processing were partly responsible for the failure of this theory (see Chapter 5).

As an answer to the logical problem of acquisition, the child's grammar construction is assumed to be constrained by a priori principles and so the class of grammars constructed

by children falls within the class of possible grammars allowed by the theory of grammar. Therefore, the theory of grammar will place bounds on grammar construction by the child, such that every grammar is a possible grammar, every change a possible change, though other, non-grammatical, factors may dictate that the child proceeds in a fashion which is in fact even more limited than the theory of grammar would in principle allow. The theory of grammar, then, will yield certain predictions about the kinds of grammar the child constructs, for example about the class of errors that may be expected in the acquisition process, or the kinds of changes that may be found between grammars of different stages of acquisition. Some of the predictions that can be expected directly from the theory will be discussed in Chapters 3 and 7, as well as ways in which further restrictions may be placed on the child's grammar construction, due to the intervention of non-grammatical factors, such as limitations on memory capacity and data-handling or processing abilities.

Not only can linguistic theory offer insights into acquisition but acquisition data itself can be of relevance to linguistic theory. That is, acquisition data are part of the data base to which the theory of grammar is responsible. They rank with other sources of data - adult synchronic data about grammaticality, ambiguity, and paraphrase, diachronic data, data from language pathology - as an area in which to test the hypotheses of the theory of grammar and to construct

particular grammars. Furthermore, given the deficiencies that exist in the primary data available to the child, a consideration of the exact nature of such data - the lack of negative evidence, of information on ambiguity and paraphrase - can suggest further restrictions on the theory of grammar. The study of "learnability" suggests how important the data are in determining what form of grammar construction should be permissible by the theory of grammar (Chapter 4).

Even amongst those who adopt a restricted framework there are constant misconceptions about the relationship between the theory of grammar and acquisition, for example about what it means to say that a child's grammar is psychologically real and optimal. Such matters will be discussed in Chapters 5 and 6, where I shall take a critical look at some attempts to deny the psychological validity of the theory of grammar.

This thesis, then, is essentially a critical examination of some current assumptions about acquisition and linguistic theory, and an attempt to provide some conceptual "tidying up", by showing just how much linguistic theory can be expected to say about acquisition and the importance of acquisition data for the theory of grammar. It will not attempt to provide a theory of acquisition or give arguments for how to embed a theory of grammar within a theory of performance.

FOOTNOTES TO CHAPTER 1

1. It is, of course, denied by some that the child has to construct a grammar at all, e.g. Fodor, Bever and Garrett, (1974). See Chapter 5 for discussion.

2. For the kind of evaluation of grammars implied by the concept of markedness, that is for a ranking in terms of "accessibility", see Chapter 8.

3. See Bloom (1975), Menyuk (1977) for a review of the literature.

4. Similar misconceptions prevail in pidgin/creole studies, where pidgins are felt to reveal universal phrase-structure rules while creoles develop transformations (Kay and Sankoff, 1974). The assumption that a grammar showing no transformations is somehow closer to universal grammar than one with such rules is not justified by the theory of grammar, either for acquisition or pidgin/creole studies and stems from an over-emphasis on certain substantive universals. See also Chapter 3.

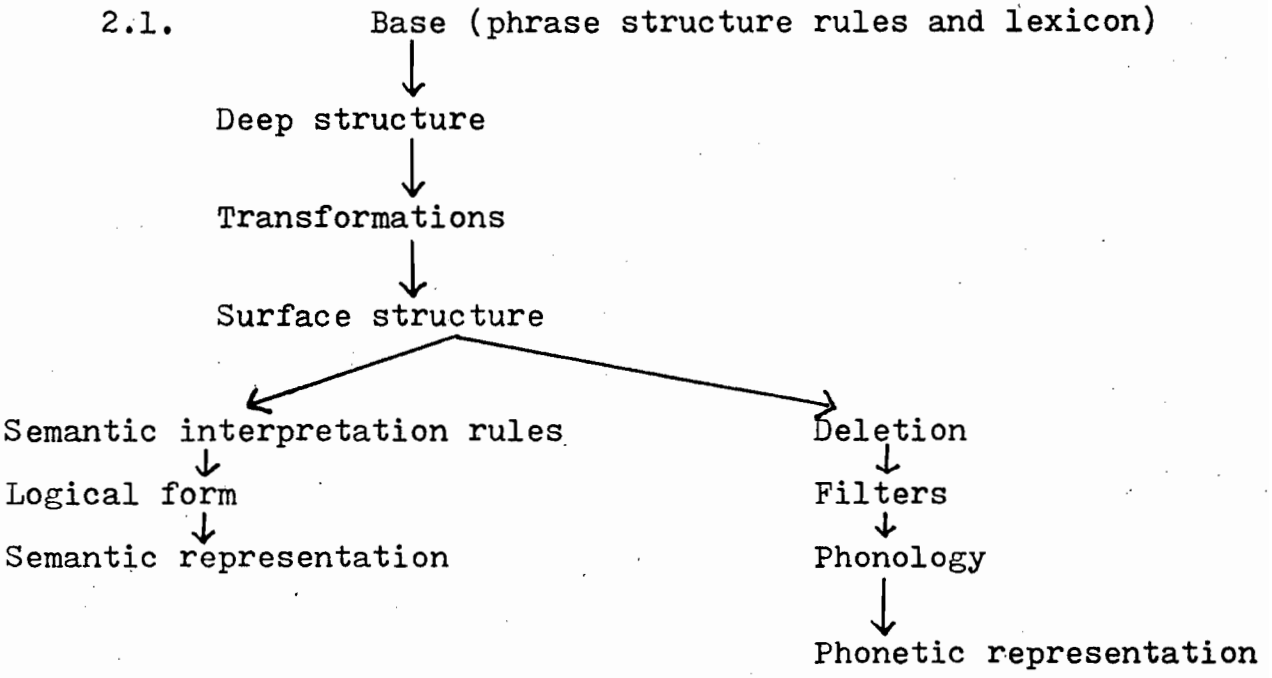
5. See Sheldon (1978) for a summary of this position.

Chapter 2

A THEORY OF GRAMMAR

The restrictive theory of grammar which will be presupposed throughout this thesis is the Extended Standard Theory (E.S.T.) of Chomsky (1970), together with more recent developments (Chomsky, 1973, 1977, 1978; Chomsky and Lasnik, 1977). Any particular grammar must conform to the following form, which is specified by the theory of grammar:

2.1.



In this way, sound (phonetic representation) can be correlated with meaning (semantic representation) over an infinite domain.

I shall briefly discuss each component of this model[1], as

well as showing how it differs from earlier versions of the theory current when many psycholinguists were investigating child language in terms of T.G.G.

In the present model, the phrase structure (P.S.) rules produce initial phrase markers, or deep structures, which can then be acted upon by the transformational component. The P.S. rules are limited by $\bar{X}$ conventions (Jackendoff, 1977), which constitute an attempt to limit the form of P.S. rules for all languages. In any particular language, a P.S. rule should be of the form:

2.2. $\bar{\bar{X}} \longrightarrow \bar{X} \quad (\text{Spec } \bar{X})$

or the reverse:

2.3. $\bar{\bar{X}} \longrightarrow (\text{Spec } \bar{X}) \; \bar{X}$

(where X is a major category such as noun, verb, etc.). A language should have all its specifiers (Spec) following the major categories or preceding them, but not both. For example, in a grammar where auxiliary (aux) and determiner (det) are specifiers of verb (V) and noun (N), that grammar should contain V aux, N det or aux V, det N but not aux V and N det or V aux and det N. (In fact, such combinations do occur and they are considered "marked".) Thus, phrase structure rules are language-specific but must preferably fall within the framework of the $\bar{X}$ proposals.

The base component of the grammar also contains the lexicon. Originally (Chomsky, 1965), the lexicon was morpheme-based, with transformational rules, such as the nominalization

transformation, resulting in word-formation. In other words, both derivational and inflectional morphology were conceived of in transformational terms. More recently, following proposals in Chomsky (1970), it has been suggested that the lexicon is word-based (Halle, 1973; Aronoff, 1976) and that morphological rules are a form of redundancy rule connecting related words and operating within the lexicon (Jackendoff, 1975a). Thus, words like "decide" and "decision" will both occur as entries and a redundancy rule will express the relationship between them. While the above proposals agree that derivational morphology operates in this way, it is not clear whether inflected forms also occur as full entries (Jackendoff, 1975a) or whether they are produced by inflectional rules operating outside the lexicon (Aronoff, 1976). As well as expressing the relationships between lexical items that are morphologically connected, lexical redundancy rules also express relationships between various syntactic constructions that a given item can occur in (see discussion of dative-movement, p.15-16). Criteria for distinguishing between such lexical rules and transformational rules are discussed by Wasow (1977).

The transformational component of the grammar assumed a large part of the descriptive burden in early versions of T.G.G. There were many transformational rules and these were very detailed. Specific items could be mentioned in rules, inserted, deleted or moved. Transformational rules were used

to express the relationship between sentences which differed in surface structure but were similar in meaning. For example, active and passive sentences would arise from a common deep structure and the passive transformation would carry out two operations: moving the subject into an agent by-phrase (which was created by the original rule (Chomsky, 1957) but subsequently base-generated with an empty slot to accommodate the deep subject (Chomsky, 1965)) and moving the object into the vacated subject position. The by-phrase could subsequently be deleted. Similarly, the relationship between the following sentences:

2.4. a. John gave a book to Fred
b. John gave Fred a book

was captured by giving them the same deep structure, similar to (2.4.a), with a transformation of dative-movement to account for (2.4.b). The phrase "to Fred" would be moved and the preposition subsequently deleted. Another pair of transformationally related sentences would be the following:

2.5. a. the man who is tall
b. the tall man

where the prenominal adjective in (2.5.b) was considered to be derived from the relative clause in (2.5.a) (Smith, 1961).

Many such pairs of sentences are no longer considered to be transformationally related, or if they are, the transformations are rather different from those originally proposed. Thus, amongst subsequent proposals for passive have been that the agent should be base-generated in its by-phrase, so that

only object pre-posing takes place (Hornstein, 1975), that passive is a lexical rule not a transformational rule at all (Bresnan, 1978), that some passives are transformational and some lexical (Wasow, 1977). The two different kinds of dative construction are now considered to be base-generated (Oehrle, 1976; Baker, 1978, 1979a) and the relationship between them is captured by lexical redundancy rules. Adjectives in pre-nominal position are base-generated (Baker, 1975). Such structures were the subject of a great deal of investigation by psycholinguists, who, not surprisingly considering the time when the experimentation was done, assumed a transformational analysis. The change of focus away from the traditional kind of transformation has certain consequences for work on acquisition, as we shall see (Chapters 4 and 5).

In E.S.T. there has been a move away from a transformational component containing detailed and specific rules, towards rules which can be written in terms of limited formal properties, thus reducing the expressive power of transformations. Details of a rule's application, such as contextual restrictions, are no longer mentioned in the rule itself. Instead, in the "core" grammar of English, proposed by Chomsky (1977, 1978), Chomsky and Lasnik (1977), transformational rules are limited to one or two, namely "move NP" and "move WH" (Chomsky, 1977) or just "move α" (Chomsky, 1978), where α is a category in the $\bar{X}$ system.

Such rules, because of their simplicity, over-generate

greatly. Instead of building restrictions into individual transformations, the operation of these rules is limited by conditions on transformations (Chomsky, 1973), conditions which are universal and apply to all unmarked, core rules. Thus, any rule may operate freely but an ungrammatical derivation is prevented by these conditions. More recently (Chomsky, 1978), the conditions have been seen as conditions on representations in logical form, rather than on rule functioning.

Some of the proposed conditions limit the functioning of transformational rules, while others limit the form of grammars. Some affect interpretive rules as well as transformational ones, for example the Specified Subject Condition (S.S.C.), and the Tensed S Condition or Propositional Island Constraint (P.I.C.), to which we shall return when considering the interpretive rules. The concern of E.S.T. is to find grammatical principles which apply as generally as possible, and which can be thought of as innate, rather than in the description of individual rules, which was a major part of early T.G.G.

An example of a condition limiting the functioning of transformational rules is provided by the Subjacency Condition (Chomsky, 1973). This condition states that:

2.6. No rule can involve X, Y, if Y is not subjacent to X.[2]

"Subjacent" means that there is at most one cyclic category

between X and Y i.e. that X and Y are in the same or adjacent cycles.[3] The structure (2.7.b) is prevented from being derived from (2.7.a.) by subjacency:

2.7.a. COMP he believes $[_{NP}$ the claim $[_{\bar{S}}$ COMP John saw who$]]$

b. *who does he believe the claim John saw

Wh-movement applies from COMP to COMP. COMP means "complementizer" and is a node which occurs at the beginning of clauses, the English complementizers being "that", "for" and "Wh" (Bresnan, 1972). However, in (2.7.a) there is no COMP node on the NP cycle. As the two COMP nodes are not on adjacent cycles, Wh-movement is prevented, explaining the ungrammaticality of (2.7.b) without any need to put a specific restriction on the rule of Wh-movement itself.

A condition on the form of rules is the Structure-Preserving principle of Emonds (1976). This principle states that non-root transformations, those that can apply in embedded clauses, must be structure-preserving. This means that an element cannot be moved unless a node of the relevant category is present in deep structure in the position to which the element will be moved. In other words, it is generable in that position by the phrase-structure rules. Root transformations, on the other hand, which apply only in main clauses,[4] do not have to be structure-preserving. Subject-auxiliary inversion, for example, is a rule which applies only to root sentences and is not structure-preserving, since there is no P.S. rule for English which generates an auxiliary node

in sentence initial position. The passive transformation, on the other hand, is structure-preserving, since the object moves into subject position, an NP slot generated by the P.S. rules.

Of the two transformational rules proposed by Chomsky (1977) as part of "core" grammar, Wh-movement covers a number of cases which were previously dealt with by different transformations. Wh-movement now accounts for structures which explicitly involve Wh elements, such as direct and indirect questions and relative clauses, as well as cases where no Wh element is apparent at surface, such as comparative-deletion, topicalization, clefting and tough-movement.

Transformations which involve the movement of NPs or Wh leave "traces" (Chomsky, 1973). A trace "t" is left in the original position of the NP or Wh element after it has been moved. For example, in the derivation of passives, the object NP will leave a trace when it moves into subject position. Thus (2.8.a) will become (2.8.b):

2.8.a. np was examined John by the doctor
b. John was examined t by the doctor

(assuming a base-generated agent by-phrase). Traces are interpreted by the rules of anaphora (see below). For such interpretation to take place, a trace must be "properly bound" by the NP that has moved: the NP must be to the left and higher in the tree. Any post-posing rule leaves an ill-bound trace which cannot be interpreted. For example, if the

derivation of passive includes moving the subject into a base-generated by-phrase with an empty NP slot, a trace will result which is not properly bound:

2.9.a. The doctor examined John by np
b. t examined John by the doctor

However, such traces can be covered by lexical material; in this case the subsequent movement of "John" into subject position will cover the ill-bound trace. Where such traces are not covered, they cannot be interpreted and the derivation will be rejected as ill-formed. Trace theory allows the conditions on rules to apply to cases which could otherwise not be excluded by them (Chomsky, 1973) and also allows semantic interpretation to be done at the level of surface structure.

The output of the transformational component is surface structure, a term which differs from its earlier usage, since surface structure now contains traces and has not yet undergone the effects of deletions or filters. Two different sets of things happen at surface and they are independent of one another. Proceeding down the left-hand side of figure (2.1), we have the semantic interpretation rules. In early T.G.G., semantic interpretation was considered to take place at deep structure (Katz and Postal, 1964; Chomsky, 1965). Later, Chomsky (1970) and Jackendoff (1972) argued that at least some semantic interpretation must be done at surface and it now appears, given trace theory, that all semantic inter-

pretation can be done there (Jackendoff, 1975b; Chomsky, 1977; White, 1977).

Among the interpretive rules that operate at this level are the rules of construal, including rules of anaphora, which relate anaphors, including traces, to antecedents, and the rule of disjoint reference, which establishes when items cannot be anaphoric. The interpretive rules themselves map into logical form which "has basic properties of some variant of predicate calculus" (Chomsky, 1978) and is interpreted in a similar way to it (Chomsky and Lasnik, 1977).

The rules of construal are part of the core grammar and are bound by certain of the conditions which also govern the application of transformational rules, in particular S.S.C. and P.I.C. The S.S.C. states:

2.10. No rule can involve X, Y in the structure

$$\ldots\ldots X \ldots\ldots \left[_{\alpha} \ldots\ldots Z \ldots\ldots WYV \ldots\ldots \right] \ldots\ldots X \ldots\ldots$$

where Z is the specified subject of WYV and α is a cyclic node

"Specified subject" means that Z is either lexically filled or it is a PRO (empty subject) controlled by an element other than X. Tensed-S or P.I.C. states:

2.11. No rule can involve X, Y in the structure

$$\ldots\ldots X \ldots\ldots \left[_{\alpha} \ldots\ldots Y \ldots\ldots \right] \ldots\ldots X \ldots\ldots$$

where α is a tensed sentence.

To see how the conditions affect the interpretive rules, consider the following sentences, which are ungrammatical:

2.12.a. *the men saw [$_{NP}$John's pictures of each other]

b. *the candidates expected [$_{\bar{S}}$ that each other would win]

In (2.12.a), the interpretation of "each other" with "the men" is blocked by the S.S.C., since "John's" is a specified subject, while in (2.12.b) "each other" cannot be interpreted with "the candidates" because "each other" is inside a tensed sentence and P.I.C. intervenes.

The transformational rules of NP movement and Wh-movement and the interpretive rules of construal, the core grammar, are unmarked (Chomsky and Lasnik, 1977). That is, they do not violate the universal conditions on rules. However, not all conditions are absolute and it is possible to formulate rules in such a way that they can by-pass the conditions, so that they are not subject to them. Unmarked rules are written in terms of variables; they are maximally general. To "immunize" a rule against the conditions, one must specify the constants involved in the conditioning environment of that rule. Thus, a rule of the form:

2.13. .....X..... [$\bar{N}$P....Y...].....

would not be subject to conditions on rules because of the constant, NP, intervening between X and Y, which a condition referring only to conditioning environments expressed in terms of variables cannot take account of. But the use of constants to express a rule involves a cost, in that such rules cannot be generalised; they are exceptional or "marked".

Chomsky (1977) gives an example of a marked rule using

material from Kayne (1975). Kayne argues for a general, unmarked rule of L-tous in French, which moves quantifiers to the left and obeys the conditions on rules:

2.14. W V Q Z ----> W Q V Z

Here, V (verb) and Q (quantifier) would be the X and Y involved in the rule and all other material can be expressed in terms of the variables W, Z. This rule accounts for cases of quantifier movement, such as (2.15.b) from (2.15.a):

2.15.a. Les garçons sont partis tous à la guerre
b. Les garçons sont tous partis à la guerre

It cannot, however, account for the following sentences, which are acceptable to some speakers:

2.16.a. Il faut toutes /qu'elles s'en aillent/
b. Il faut tous /qu'on se tire/

The interpretive rule needed here must construe the quantifier with a pronoun which falls within a tensed sentence, apparently violating P.I.C. To avoid this, the rule must be expressed in terms of constants, so that it is no longer subject to the condition:

2.17. vbl V* Q que α Pro vbl

(Chomsky's rule (15))

where Q may be construed with the pronoun (Pro).[5] Rule (2.17) is more highly marked than (2.14). As we shall see in Chapter 8, claims about markedness in grammars may be considered as claims about accessibility; in acquisitional terms, this means that the more marked grammars are harder to acquire.

To return to figure (2.1), on the right-hand side of the

diagram, the first thing to take place after surface structure is deletion. Deletions in early versions of the theory were transformations. They are no longer considered to be such and, as can be seen from the diagram, semantic interpretation is done off surface before deletion takes place. What can be deleted is fairly restricted (Chomsky, 1978) and deletions are subject to a recoverability condition (Chomsky, 1965).

After deletion come the surface filters (Chomsky and Lasnik, 1977), which determine well-formedness. These filters do away with the need for statements of rule ordering, obligatoriness and contextual dependencies in the transformational component. Transformations are optional and unordered and the effects of obligatoriness and ordering are achieved by the filters. For example, Chomsky and Lasnik's rule of free deletion in COMP allows the derivation of the following ungrammatical sentence:

2.18. *the man met you is my friend

where the subject relative pronoun, in the COMP slot, has been deleted. This is ungrammatical in Modern English, although it was acceptable in Middle English and early Modern English. One could propose a special limitation on the COMP deletion rule which would exclude deletion in just such cases but this would complicate the rule and cause it to lose generality. Instead, Chomsky and Lasnik propose the following filter:

2.19. $*[_{\overline{NP}}$NP tense VP$]$

(Chomsky and Lasnik's filter (20))

which excludes the relevant cases. The filter removes the need to deal with the exception at the transformational level, preserving the unmarked nature of the core rules (of which COMP-deletion is one).

Similarly, their derivation of relative clauses allows the generation of structures like:

2.20. *the man who that I saw

which, again, was grammatical in Middle English but is not so now. Instead of stating this as an obligatory occasion for the COMP deletion rule, they propose the following filter:

2.21. $*[\overline{\ }_{COMP}$ wh-phrase + complementizer$]$

(Chomsky and Lasnik's filter (18))

This rules out the undesirable sequence. Such filters may be language specific, since they apply at some stages of English but not others, or they may be conceived of as the unmarked case of universal filters, operating in the absence of positive evidence to the contrary (see Chapters 4 and 8).

Surface structure as now understood in E.S.T. is more abstract than it was in earlier proposals. It contains traces and PROs (empty but controlled NPs) (Chomsky, 1978) and deletions have not yet taken effect. This, and other factors which have been discussed, such as the limited number of transformational rules and the limitations on their operations, has helped to bring deep and surface structure closer together and the requirement of recoverability of deletions will further limit abstractness. In addition to these limitations,

further restrictions on abstractness have been proposed for phonology (Kiparsky, 1968a) and syntax (Lightfoot, 1979). As a result, it is not possible to conceive of syntactic deep structure as being universal, the so-called "universal base" hypothesis, since, on such a view, deep and surface structure must inevitably be very different and a powerful transformational component will be required to produce the surface structures of particular languages off the universal base. One of the consequences of the restrictive theory adopted here is to rule out such an analysis, an analysis which is implied in some early psycholinguistic work (e.g. McNeill, 1966, 1970).

It can also be seen that the interest of E.S.T. is in the pursuit of universals concerning the form of rules, such as $\overline{X}$ limitations, and their functioning, such as S.S.C. and P.I.C. Less weight is assigned to the substantive universals. However, much early psycholinguistic work was concerned with substantive universals, such as word classes, and little has been done to study the claims for acquisition that arise out of a consideration of the more abstract formal universals discussed here. It is the relationship between these formal universals and acquisition data that I shall be concerned with.

One final note of explanation is necessary here: when in this thesis I speak of universal grammar or universals, I mean the principles that have been outlined in this chapter, namely principles of the theory of grammar. However, the term "universal" is also used to describe certain frequently

occurring tendencies in different languages (Greenberg, 1966). This kind of universal can hardly contribute towards an explanation of the acquisition problem, cannot be a candidate for membership in the class of a priori principles with which the child is endowed, since one cannot impute to the child knowledge of all languages and the tendencies to be found in them (though, of course, some of the typological universals may be consequences of principles of the theory of grammar; for example, $\overline{X}$ conventions will predict relative order of major categories and their specifiers, which typological universals also describe).

E.S.T., then, as outlined in this chapter and described in the sources cited, is the restrictive theory of grammar which will be presupposed throughout this thesis. I shall go on to examine the kinds of prediction about acquisition that can be expected from such a theory, the ways in which it constrains the child's grammar construction, as well as ways in which acquisition data should and should not be used in support of such a theory.

FOOTNOTES TO CHAPTER 2

1. I shall not discuss the phonological component at all, since this thesis is not concerned with the acquisition of phonology. I presuppose the framework of Chomsky and Halle (1968).

2. "A transformational rule 'involves' X and Y when it moves a phrase from position X to position Y and a rule of construal 'involves' X and Y when it assigns Y the feature $[\pm$ anaphoric to $i]$, where X has the index i (or conversely, in both cases)" (Chomsky, 1977, p. 75).

3. Cyclicity is itself a condition on rule function (Fillmore, 1963) which states that all rules apply first to the lowest cyclic node in a tree, then to the next lowest, and so on. The cyclic nodes for English are $\overline{S}$ and NP (Chomsky, 1973) and possibly also S (Chomsky, 1977). Certain rules, however, are not considered to be cyclic but are pre- or post-cyclic; for example, affix-hopping is post-cyclic, applying after all the other cyclic rules.

4. There are exceptions to this. See, for example, Hooper and Thompson (1973).

5. "Where V* is a certain class of verbs including 'falloir', 'vouloir', Q is a quantifier and α is either null or is a 'sufficiently short' NP" (Chomsky, 1977, p. 77). Note that this rule is not just arbitrarily created to solve the problem of the "violation" of conditions by forms such as

(2.16.a, b); a separate rule is motivated anyway, as they require an interpretive rule, not a movement rule like (2.15.a, b).

CHAPTER 3

PRINCIPLES OF GRAMMAR AND ACQUISITION

3.1. The explanatory power of a priori principles.

The theory of grammar outlined in the previous chapter addresses itself towards explaining the logical problem of acquisition, how it is that despite an environmental stimulus deficient in certain respects, including performance errors and lacking information about ungrammaticality, ambiguity or paraphrase relations, the child learns his mother tongue in a short space of time. Explanation is achieved if one thinks of:

> "'Universal grammar' (UG) as the system of principles, conditions and rules that are elements or properties of all human languages not merely by accident but by necessity."
>
> (Chomsky, 1975, p. 29)

In other words, universal grammar is part of the child's biological endowment, so that the child's grammar construction falls within the limits laid down by those principles.

A way in which such a priori principles can overcome the shortcomings of the primary data, the data that constitute the input to the child, can be illustrated with an example from Baker (1978). Baker shows that the input to the child learning English would be compatible with two grammars

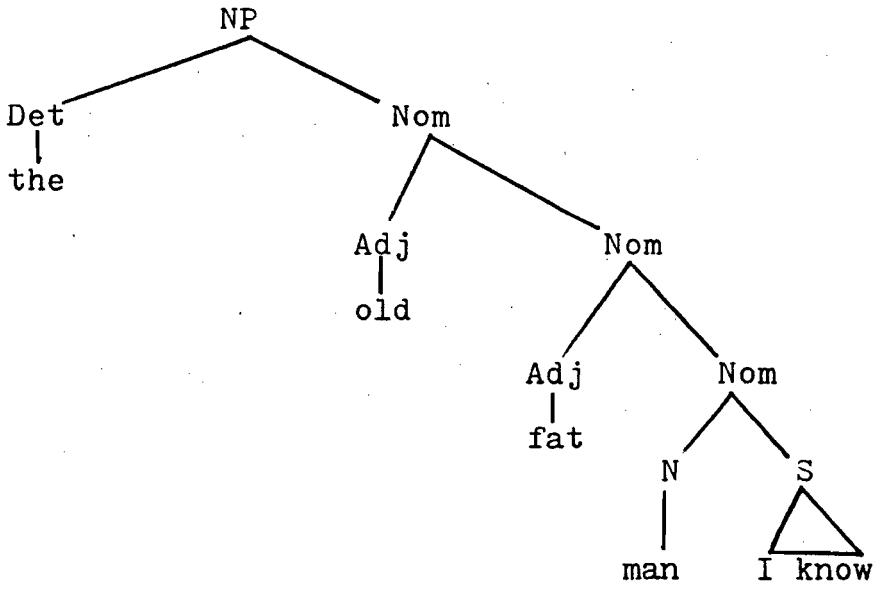
which account for the rule of One-substitution but that the non-primary data indicate that these grammars are not, in fact, equivalent and that only one correctly describes the adult's competence. Only by assuming certain restrictive principles of grammar can the learning of that grammar be explained.

Baker argues that the facts of One-substitution in English are better explained if complex NPs are analysed into a determiner plus a nominal, the latter itself a constituent consisting of the noun and its modifiers (the Det-Nom analysis), than by more traditional accounts of an NP followed by an S (or PP), where the noun and its modifiers do not make up a single constituent. That is, an NP such as:

3.1. the old fat man I know

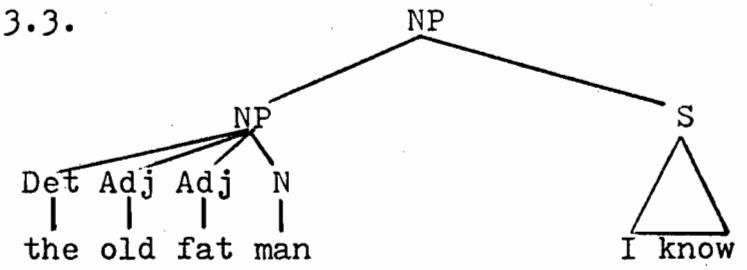
is represented as follows on the Det-Nom account:

3.2.



The NP-S analysis, on the other hand, would look like this:

3.3.



On the Det-Nom analysis, all phrases for which One-substitution can occur are also constituents under a nominal node, whereas, on the NP-S account, the word 'one' replaces groups of words which are not constituents. For example, in:

3.4. The old fat man from Cleveland was more helpful than the young thin one

"one" substitutes for "man from Cleveland", which is itself a nominal on the Det-Nom account but which is not a constituent on the alternative proposal.

Furthermore, many sentences containing a One-substitution are ambiguous. A rule written in terms of nominals can account for this fact, since there can be more than one nominal in any given tree, but a rule depending on the NP-S analysis, written in terms of nouns instead of nominals, cannot. In:

3.5. I would rather tell you another funny story but I've already told you the only one I know

"one" can stand either for "story" or "funny story" but the NP-S account predicts only the former interpretation, since "funny story" is not a constituent on that analysis, whereas it is a nominal on the Det-Nom account.

Finally, the NP-S proposal makes the incorrect prediction that "one" can be substituted for the noun "student" in (3.6) to give:

3.6. *The student of chemistry was more thoroughly prepared than the one of physics.

However, "the one of physics" is ungrammatical. The Det-Nom account does not predict that "one" can replace "student" here, since the lowest nominal in this case is "student of chemistry" and not "student".

It follows from these arguments that the NP-S account is not descriptively adequate for adult English. However, as Baker points out, from the data available to the child learning English, it is impossible to tell when One-substitution produces ambiguities and when it produces ungrammatical results. The child is not told explicitly that (3.5) is ambiguous or (3.6) ungrammatical and yet he becomes aware of this. If the child hears sentences where "one" is used and has some indication that "one" must be interpreted as referring to a previous expression and has been exposed to data showing the kinds of constituent that can occur within NPs, then, on the basis of such experience alone, the NP-S account and the Det-Nom account look equally good. The primary data are not informative enough to single out the analysis which correctly captures the child's and adult's competence, their knowledge of where and when "one" may substitute for another expression. How does the child acquire the Det-Nom grammar, which his knowledge of ambiguity and ungrammaticalness suggests that he must?

If a number of restrictive assumptions are made about universal grammar, then the acquisition of the Det-Nom analysis can be explained by ruling out in advance the NP-S grammar.

Baker suggests a number of restrictions on phrase-structure rules, similar to $\bar{X}$ proposals, such that:

3.7.a. NP is a sequence of Det and Nom, the order to be specified in the grammar of each language.

b. Nom is a sequence of Nom and S, Nom and Adj, Nom and PP, etc., with the order to be specified in the grammar of each language.

c. Nom is a sequence of N and various NPs, PPs, etc., with the order to be specified in each language.

(from Baker's (20a, b, c), p. 418)

These proposals serve to exclude the NP-S analysis. To exclude a One-substitution rule that replaces N instead of Nom, he also proposes that:

3.8. In any rule that replaces one of two identical categories, only phrase structure categories higher than the lexical categories N, V, Adj, etc. are eligible to figure in the rule as the specification of the replaced category.

(his (21), p. 418)

Such proposals, whilst here meeting a particular need to explain the acquisition of the Det-Nom grammar, are also justified on independent grounds. For instance, the $\bar{X}$ proposals of Chomsky (1970), Jackendoff (1977) (see Chapter 2) are similar but made for different reasons. Thus, one set of proposals is likely to be able to account for a number of different facts.

The restrictive principles proposed by Baker explain the possibility of acquisition in this case; although the primary data alone are not sufficiently rich to determine the grammar uniquely, these data interacting with the proposed a priori

principles are. Given the input data alone, the two grammars appear equivalent but the child's and adult's competence shows that only one can be correct. Since the data cannot distinguish between them, one must look beyond the data for an explanation.

3.2. Children's grammars as "possible" grammars.

By accepting the principles of the theory of grammar as part of the biological specification of the child, one is including the class of child grammars within the class of possible grammars, as defined by E.S.T., so that data from child language is data about possible grammars, along with data from other areas. If the child's grammar construction is constrained by the theory of grammar, he ought not to hypothesize grammars that the theory does not permit. This is a working hypothesis which is testable on acquisition data; if the theory provides an upward boundary on child grammar construction, children ought not to produce forms which go beyond that boundary. Thus, the theory of grammar makes predictions about the kinds of "errors" that can be expected to occur in child language.

If, in constructing a grammar, children were to try out all logical possibilities, their language should show forms which, in fact, never seem to occur. For example, Chomsky (1975) argues that children never try out structure-independent hypotheses in grammar construction. Given declarative sentences such as:

3.9.a. The man is tall
b. The man who is tall is in the room

a child forms questions of the following kinds:

3.10.a. Is the man tall?
b. Is the man who is tall in the room?

He never produces the following form:

3.11. *Is the man who tall is in the room

Were the child simply looking for logical possibilities on the basis of the input data, it would be perfectly consistent for him to try out the structure-independent alternative of preposing the first occurrence of "is". Indeed, this would be expected were he operating on inductive principles alone. Instead, it appears that children never try out structure-independent hypotheses but are always constrained to work within the framework of structure-dependency. Since there is nothing in the input data alone to suggest such a principle, one must look beyond the data for an explanation of this restriction on grammar construction. As a principle of universal grammar, structure-dependency explains the child's ability to ignore structure-independent hypotheses in grammar construction and it is confirmed by the fact that child language shows no evidence of structure-independent hypotheses.

There are cases where acquisition data seem to show that the child does try out hypotheses which the theory of grammar would exclude. For instance, certain data provide an apparent counter-example to the restrictions on phrase-structure rules discussed by Baker (1978) and specifically to (3.7.a). Children are reported to produce forms like:

3.12.a. a this truck
b. a my pencil (Brown and Bellugi, 1964)
c. dats a your car
d. dats your a car (McNeill, 1966)

This looks as if they consider "Det Det Nom" to be a possible phrase-structure rule, which is ruled out by Baker's proposals. If knowledge of such proposals is innate, children might be expected to try different orders for determiners and nominals but not to produce forms like those in (3.12). Before taking these forms as evidence against a universal constraint, one must explore other possibilities. If the children are classifying part of the determiner system as adjectives, for instance, the problem would not arise, since the adjective would be a constituent under the nominal. If empirical support [1] for such alternative explanations proves lacking, then one has reason to question the proposed universal constraint.

Another apparent counter-example comes from Matthei (1978). Structure-dependency, as discussed above, is an absolute condition on rules; no rule may violate it. In the case of some of the conditions discussed in Chapter 2, such as the S.S.C. and P.I.C., rules can be written in such a way that they by-pass these conditions. However, any unmarked rule must be bound by them. If such conditions are biologically specified, no violations of them should occur in the acquisition of unmarked rules.[2] Matthei offers evidence which is, at first sight, very damaging, since he finds that, in the acquisition of the reciprocal "each other", children appear to disregard both S.S.C. and P.I.C. In sentences such as:

3.13.a. The chickens want the pigs to tickle each other
b. The chickens said that the pigs tickled each other

his subjects (17 children aged 4.2 to 6.6) chose the matrix clause subject as the referent for "each other", in 64.4% of their total responses. That is, the children acted out the sentences by making the pigs tickle the chickens, and vice versa. In other words, the interpretive rule was relating "each other" to a word over a specified subject or outside a tensed sentence. As "each other"-interpretation is an unmarked rule, part of the core grammar of English, such violations should not occur.

Matthei remarks that, contrary to his expectations, this result suggests that the conditions on "each other"-interpretation are learned. However, in order to count against the innateness of S.S.C. and P.I.C., the evidence must be shown to be relevant. In this case, it may turn out to be evidence precisely of the innateness of these conditions. Matthei notes that the children may be treating "each other" as an ordinary pronoun. If this observation is pursued, it is apparent that no violation of the conditions occurs. That is, suppose that in learning "each other " the child has to realize that it is reciprocal and anaphoric (referring to an antecedent) and bound (controlled by that antecedent). In that case, it is possible that the child reaches a stage where he knows that "each other" is reciprocal and anaphoric but does not realise that it is bound. In other words, he might treat "each other" like "them" in (3.14.a, b):

3.14.a. The chickens want the pigs to tickle them
b. The chickens said that the pigs tickled them

In this case, the rule of disjoint reference (Chomsky, 1973) applies to exclude a reading where "the pigs" and "them" co-refer. P.I.C. and S.S.C., far from being violated, apply to prevent the rule of disjoint reference from applying to "the chickens" and "them", so that these can be interpreted as coreferential. This explains the case of the pigs being made to tickle the chickens. The fact that the children also made the chickens tickle the pigs suggests they have some understanding of the concept of a reciprocal.

Thus these cases do not necessarily provide evidence of the child's constructing grammars beyond the bounds of U.G. but can be explained by differences in detail between the child's grammar and the adult's as far as adjectives and reciprocal pronouns are concerned. However, Fodor, Bever and Garrett (1974) claim that grammars which provide a correct description of child speech may nevertheless be ones which the theory of grammar would rule out, ones which would be "impossible". If this were true, the possibilities for the child's grammar construction would seem to be unlimited and the learning task would be impossible. In fact, their claim is based on a misconception over what the theory says about the form of grammars.

Fodor, Bever and Garrett justify their claim by taking Braine's (1963a) characterization of two-word speech in terms of pivot-open (P-O) grammars as a correct description of that stage. They then argue that such grammars do not observe universals found in adult grammars, in particular that pivot-open grammars have no transformational component.

There are two major issues here. The first is whether pivot-open grammars even provide a correct description of two-word speech at all. Bloom (1970), Bowerman (1973), Brown (1973), amongst others, suggest that such grammars are inadequate for the language that they attempt to describe. If they cannot even describe child language correctly, then the question of whether they are or are not possible is superfluous.

Supposing, however, for the sake of the argument that P-O grammars provide an adequate representation of the child's competence, one cannot show that they are impossible grammars simply because they lack a transformational component. This would be to assume that all universally available features of language must be available to the child from the beginning and that they must occur in all languages. I shall shortly argue that many universal features of language are triggered by the data; therefore, if

the child does not perceive the language in terms of transformations, then there will be nothing to activate the transformational component. But the fact that children's early speech may not include transformations does not mean that they are incapable of producing them as soon as their perception of the data merits it. As Roeper (1978) points out:

> "We know that children learn to identify verbs before they identify auxiliaries. Therefore it is not surprising that the power of transformational analysis comes into use in German at an earlier age than in English since the verb-second (or verb final) rule in German involves the main verb while in English the subject-auxiliary inversion rule involves the auxiliary. Thus the sequence of hypotheses that a child tests and the power they entail could be different for each language."
>
> (p. 10)

Roeper (1973) and Bowerman (1973) have argued, for German and Finnish respectively, that transformations are present in two-word speech in those languages. It is an accident of English, then, an accident of the data, that the transformational component has not been activated at a time when it is activated for children learning other languages. Indeed, there may be cases where the transformational component is never activated, for instance in languages with very rich inflectional morphology. The grammars of such languages are not impossible either, any more than child language describable by P-O grammars which happens, by chance rather than necessity, to have no transformations.

If one cannot expect necessarily to see a transformational component in operation, one might, instead, argue in terms of universals relevant to phrase-structure grammars, which is what P-O grammars are. The $\bar{X}$ conventions (Jackendoff, 1977) are constraints on phrase-structure rules (see Chapter 2). Fodor, Bever and Garrett observe that:

> "The order of the classes is relatively stable for a given child - that is, children who produce pivot + open do not produce open + pivot and vice versa."
>
> (p. 485)

If the pivot class is taken as equivalent to the specifier and the open class as $\bar{X}$, this is consistent with the demands of the $\bar{X}$ convention that all specifiers precede their major categories or follow them but not both. Thus, P-O grammars do appear to conform to phrase-structure universals for which they can be tested.[3]

It is not the case, then, that P-O grammars are both correct descriptions and impossible grammars. They lack a transformational component because the data they describe do not require to be accounted for in those terms and they do appear to abide by universals relevant to phrase-structure grammars. Of course, the question of whether they are even adequate descriptions of child language still remains, but this was not the issue in Fodor, Bever and Garrett's discussion of them.

3.3. The triggering effects of data.

The fact that certain principles of grammar are innate

says nothing about when they will be seen in action. I have already alluded in the previous discussion to the possibility that grammatical principles are only revealed by relevant data; if the child is unaware of auxiliary verbs, he or she will not propose a transformational analysis for yes/no questions but as soon as he becomes aware of auxiliaries, the transformational component can be activated (assuming that it has not yet been activated by other data).

According to Chomsky (1965):

> "Certain kinds of data and experience may be required in order to set the language-acquisition device into operation."
>
> (p. 33)

In that case, one cannot expect to see grammatical principles in operation without relevant data to trigger them. Thus, it is pointless to ask whether principles of grammar like S.S.C. or P.I.C. are present in the child from the beginning. To return to the example of structure-dependency, discussed in the previous section, at the earliest point when the child can ask:

3.10.a. Is the man tall?

he will be incapable of asking the more complicated question:

3.10.b. Is the man who is tall in the room?

This means that the non-occurrence of forms like:

3.11. *Is the man who tall is in the room

does not necessarily at this stage constitute evidence in favour of structure-dependency but merely results from the fact that the child cannot produce any complex structures at

all. However, this lack of evidence for structure-dependency in early speech is not crucial; what is important is that at the time that the child is producing complex sentences, he is never found to make errors like (3.11). That is, the constraint is triggered as soon as it is needed, when the relevant structures are being produced, without prior learning.

A similar point is made by Roeper (1978) who shows that as soon as children understand that determiners mark noun-phrases, they understand the control differences between NP and VP gerunds. In:

3.15. John likes singing songs

they realize that it is John who likes singing, whereas in:

3.16. John likes the singing of songs

they understand that the subject is unspecified. As Roeper remarks:

> "We cannot prove that the distinction is innate, but the assumption that it is innate has an important consequence. Under the innateness assumption, the child's task is not to infer the existence of the principle (a control difference) but simply to recognize where it applies. The child must learn to recognize that gerunds are nounphrases if they are marked by a <u>determiner</u> (the), <u>negative</u> (no), or <u>adjective</u>. If a gerund has a nominal marker, then it is automatically uncontrolled. Until, however, children learn to recognize these noun-markers, we would not expect them to exhibit knowledge of the nominal/verbal difference."
>
> (p. 15)

This brings out quite clearly the interaction of universal principles with the child's perception of the data at any particular time. It makes no sense for the principles to be activated before the data is seen in a relevant light. Roeper's

experimental results support his contention that the control difference between nominal and verbal gerunds is part of U.G. and available to the child as soon as the relevance of noun-markers is understood. That is, children who understand the determiner system do not interpret (3.16) as though it meant the same as (3.15).[4]

Thus, though the principles of universal grammar may be present in the language faculty all along, they will not necessarily all make their appearance at the same time. We are unable to see the effects of such principles until they are revealed by the data. When particular data occur or when the child comes to see the data in a particular way, he or she will be constrained by the principles of grammar which are relevant to that data.

This is very different from the position adopted by McNeill (1966, 1970) and others (e.g. Stampe, 1969; Traugott, 1973; Kay and Sankoff, 1974; Bickerton, 1977) that the child is somehow closer to universal grammar than the adult and that the language that emerges first reveals aspects of universal grammar, whereas later forms are more language-specific. This view stems from an over-concentration on substantive universals; it is difficult to see how formal universals such as structure-dependency or S.S.C. could possibly occur as part of the child's earliest language in the absence of the data which require such principles.

As well as being triggered by data, it is possible that

certain principles are more directly affected by maturation, just as other innate specifications, such as eye colour, are not realised immediately. For example, C. Chomsky (1969), in her investigation of pronominal reference in children, found that children under the age of 5.6 did not seem to be aware of the non-identity requirement for sentences such as:

3.17. He knew that Pluto was sad

They would interpret "he" and "Pluto" as coreferring, whereas children over 5.6 did not. This is inconsistent both with a claim that such principles have to be learned and with the kind of triggering effects of data discussed above. If the non-identity requirement were learned, one would expect considerable variation in the age at which it was acquired, as C. Chomsky found with the other structures that she investigated, rather than a fairly clear cut-off point at 5.6, as though some principle had suddenly become activated. Nor does the principle seem to be triggered by the needs of the data, since the children encountered sentences like (3.17) long before the age of 5.6 but misinterpreted them until this point. If we take as an initial hypothesis that all principles of universal grammar are present in the child from the beginning (though dormant until triggered by relevant data), this would appear to be falsified by cases such as the above where a principle seems to be activated maturationally. Thus, we must allow for both possibilities.

Given that a priori principles cannot be apparent unless one has data that is relevant to them, one must be wary

of using such principles in appropriate contexts. A misguided attempt to explain acquisitional data by means of structure-preservation (Emonds, 1976)[5] has been made by Parker (1977). In languages where base word-order is hard to determine, such as German, which has SVO order in main clauses and SOV in subordinate ones, Emonds's hypothesis means that the order of the subordinate clauses must be taken as the base order. In other words, SOV is basic and the SVO order of main clauses is achieved by a non-structure-preserving transformation. If SVO were basic, then the non-structure-preserving rule converting SVO to SOV would have to apply in non-root sentences, which is prohibited by the hypothesis.

What is the position for children learning German and trying to determine the base-order of that language? Parker (1977) maintains that structure-preservation predicts that children will learn SOV order right from the beginning. But in the early stages of acquisition, when children show no evidence of understanding or using embeddings and when few sentences containing subordinate clauses are directed at them (Drach, 1969; Pfuderer, 1969; Snow, 1972), the data that the child is most likely to draw on involve main clauses only, which in German are SVO. If the data are exclusively of this type, then the principle of structure-preservation is simply irrelevant, since non-root sentences are not yet in question.

In order, therefore, to explain the child's supposed ability to find the SOV order by using the structure-preserving principle, Parker proposes a special perceptual

strategy telling the child to pay attention to the order of constituents in embedded sentences, to bring non-root sentences into consideration. This is implausible if the child does not hear many embedded clauses in the first place and ignores them if he does (Solan and Roeper, 1978). If children do show SOV order before they are aware of embeddings, it must be due to something other than the operation of structure-preservation.

The actual evidence on word-order acquisition in German is conflicting. Park (1971, cited in Roeper (1973) and Parker (1977)) finds that children at the two-word stage show OV order 80% of the time. For three-word speech, Roeper (1973) found only 50% OV. Roeper suggests, following an observation of H.Sinclair-de-Zwart, that the two-word forms may be based on adult main clauses containing modals, where the order is S modal OV. As far as Roeper's findings are concerned, either order might be the base order with the other derived transformationally (or the order might be free) since, if one is dealing with root sentences only, the structure-preserving hypothesis cannot rule out the selection of SVO as basic, with a transformation resulting in SOV in 50% of the cases. Though such a transformation would not be structure-preserving, it would not have to be so if it operated only in root sentences. The evidence from German main clauses suggests that either order might be basic, depending on whether the child was concentrating on modal or

non-modal main clause types. There is a certain reluctance in the literature to accept that a child might start out with a base order that differs from the adult's. But there is no reason why the child should automatically arrive at the deep structure order of the adult language without trying out alternatives.[6] Children do appear to try out difference orders in acquisition (McNeill, 1966; Slobin, 1966a) and this is to be expected if child grammars are considered in their own right, as attempts to construct grammars for particular data and not as attempts to acquire the adult grammar as such.

However, Emonds's hypothesis does have specific predictions to make regarding later phases of acquisition (as Roeper (1973) also notes). When children reach a stage where they clearly attend to embedded clauses (where one does not have to rely on an implausible strategy to get this result), the data is now relevant for the triggering of the structure-preserving principle. Therefore, one would expect their subordinate clauses generally to show SOV order. If the child had postulated underlying SVO on the basis of evidence from main clauses, or if he had assumed a free word-order, he should restructure his grammar to produce underlying SOV, on the basis of evidence from subordinate clauses and the principle of structure-preservation. Of course, if the child has postulated SOV order all along, then no restructuring will be necessary. But if he does choose this order initially, it

cannot be argued that the choice was determined by structure-preservation, since the relevant triggering data is not present from the beginning.

An equivalent prediction for English that structure-preservation does make is that subject-auxiliary inversion, a non-structure-preserving rule, should not occur in subordinate clauses. However, cases have been reported, for example by Menyuk (1971):

3.18. I saw where is the hat

It may be that such indirect questions with subject-auxiliary inversion are not truly generated by the grammar but that the question word produces an automatic response of inversion because direct questions are so much more common in child language. Subject-auxiliary inversion never occurs in subordinate clauses after "whether" or "if", the equivalents of main clause yes/no questions. This suggests that it is not the grammar which generates the inversion in indirect questions, since it does not occur in all types of indirect questions, but rather that certain wh-words produce an idiom-like response. Alternatively, this may be a case of the "mixed indirect discourse" discussed by Emonds (1976), where certain indirect forms have characteristics of direct quotation and are considered root sentences.

Current grammatical theory offers certain hypotheses about the nature of universal grammar, about the a priori principles available to the child. These proposals are based

on investigations of different kinds of data; the hypotheses can be tested and they may or may not prove tenable. What the principles are is an empirical question, to be approached in various ways, using data from synchrony, diachrony, pathology, acquisition, etc. While it may be difficult to use acquisition data to make direct proposals on the nature of universal principles, due to the many other factors that are involved in the acquisition process, such data provide a useful testing ground for the proposals, via the occurrence or non-occurrence of certain forms, provided that the principles are considered with respect to relevant triggering experience.

Certain proposals for U.G. may turn out to be unsatisfactory, in that their apparent universality could be due to causes other than biological specification. For example, Olmsted (1966), Sampson (1978) claim that many phonological universals are due to acoustic or articulatory factors, so that one does not have to account for them in terms of a priori principles. Early work on syntax acquisition concentrated on substantive universals and was followed by suggestions that such universals, e.g. word classes, could in fact be established on the basis of distributional evidence alone (Braine, 1963b). But the fact that certain proposals may turn out to be mistaken, or to need amending, is no argument against the need for a priori principles as such, but only against particular unsatisfactory proposals. Though physiological factors or regularities in the corpus of data may

account for certain aspects of acquisition, there remain many things which cannot be explained in such ways and which require the assumption of some kind of a priori knowledge to constrain the child's grammar construction and explain his abilities. This is not to deny the crucial importance of the evidence available to the child, which will be considered in the following chapter, but the input data to the child cannot alone account for language learning, although it must shape our proposals for the theory of grammar.

FOOTNOTES TO CHAPTER 3

1. For example, one might apply tests of adjectival characteristics (c.f. Wasow, 1977) to the child's use of the possessives.

2. Unless there are principles of UG which are activated when the child is more mature. This will be discussed later in the chapter but I do not consider it in this case.

3. However, the identification of the pivot class with the specifier class may turn out to be impossible here, since pivots typically include words like "hi"; "here"; "see" (Braine, 1963a), which hardly fall into the adult definition of specifier. On the other hand, children may have a concept of "specifier" which includes a wider range of categories than the adult allows.

4. Roeper's experiment is concerned with wider issues than just this one. He isolates conflicting claims made by the theory of grammar and theories relying on perception strategies or pragmatic strategies. He finds that a perception strategy, namely that NVN sequences are SVO, is adopted by children who do not yet understand the significance of the determiner system for control. As soon as they understand significance of the determiners, they use the a priori principle of grammatical control. In other words, perceptual strategies are useful where a grammatical principle has not yet been triggered. No child made use of pragmatic strategies and there was no evidence for other perception strategies such as "first N is subject".

5. See Chapter 2.

6. Unless, of course, they use other cues deriving from principles of grammar. For example, $\overline{X}$ conventions might help the child to establish word-order by means of the order of specifiers and $\overline{X}$ within various categories. However, this in itself might create problems for learners of German. In SOV clauses (if the children are aware of them) the order of specifiers relative to major categories is inconsistent: noun specifiers precede nouns but verb specifiers follow verbs. This is a marked ordering and not likely to make the determining of base order easy. If children are attending only to main clauses, both noun and verb specifiers precede their categories, which is in accordance with $\overline{X}$ proposals and suggests that the children might be led into selecting SVO as basic until the activation of some other principle (structure-preservation) forces them to restructure.

CHAPTER 4

THE IMPORTANCE OF DATA

4.1. Underestimating the role of data.

In the previous chapter, it has been argued that certain deficiencies in the primary data, such as the lack of information about ambiguity, must lead to proposals for specific a priori principles, and the kinds of prediction that can be expected for acquisition data, particularly predictions about "errors", have been discussed. In this chapter, I shall consider a move to restrict the class of grammars which concentrates on the nature of the positive data available to the child, bringing in considerations of "learnability".

One effect of early proposals for the innateness of grammatical principles was the occasional neglect of consideration of the role of data. Indeed, in certain cases, universal grammar was apparently considered powerful enough to create structures in the absence of any relevant data at all. For example, Gruber (1967) proposes that the speech of the child he investigated should be analysed in terms of topic-comment rather than subject-predicate constructions. Noting that the child was not exposed to such constructions at all in his parents' speech,[1] he remarks:

> "The spontaneous creation of the topic-comment construction could then be attributable to the innate capacities of the child, to his innate 'knowledge' of language."
>
> (p. 39)

Similarly, Bickerton (1977) states that "there can be rules of language that are not derived from any linguistic input" and Traugott (1973) argues that children might, for no apparent reason, develop marked categories not present in the language they are learning by "reaching down" into innate hierarchies. Why they should want to do this is not clear. Such proposals, if they really mean what they seem to mean, attribute enormous power to the L.A.D.; it is difficult to see how one could restrict it in such circumstances. What would prevent the L.A.D. from "inventing" all universal categories which happened to be missing in any particular language that a child was learning? If universal grammar can really operate in the absence of data, as these suggestions imply, it is curious that this does not happen more often in acquisition. Such proposals stem in part from the position, outlined in Chapter 3, that the child somehow has closer access to U.G. than the adult and that his language will exhibit more universals than the adult's. But all grammars, child or adult, must be circumscribed by U.G. and in all cases grammar construction results from an interaction of a priori principles and data, and not from the former alone.

In fact, such claims indicate not that the principles of grammar are so powerful that data is unnecessary, but that

linguists are not being sufficiently precise in determining the kind of interaction that can be expected between grammatical principles and data, so that anything which occurs which cannot be explained is automatically attributed to innate principles, without being very clear what the principles would be in such circumstances. It is necessary to be quite specific about what such principles are and not to use innateness as a "catch-all" for apparently insoluble problems, since this degrades the value of the explanation that the theory of grammar can offer.

It must be remembered exactly what universal principles must do: they must be abstract enough to account for the learning of all languages yet rich enough to explain how any particular language is acquired. Data from a specific language are crucial for triggering principles relevant to that language. The child constructs a grammar for the language he is learning from the data of that language in conjunction with the principles of grammar. If there is no triggering data, it simply makes no sense to speak as if the innate principles could operate by themselves, creating language. This is reminiscent of the attempts of the Egyptian Pharoah, Psammetichus, to find out which was the oldest language of the world by bringing up two children in solitary confinement to see what language they would speak.

Closer examination of the cases where U.G. appears to be operating in the absence of input data reveals that this

is not the case at all. For example, Gruber's observation that topicalized structures appear spontaneously can be explained if Chomsky's (1977) proposal is taken into consideration, that the rule of Wh-movement covers a variety of cases previously thought of as separate rules, including topicalization. At the time that the topicalized forms were observed, the child was also using Wh-movement in Wh questions, which also occurred in the parental data. His use of topicalization could be considered to be an extension of the rule of Wh-movement into a domain where it is not used by the adult, but the evidence for it is, nevertheless there, in the form of Wh questions. (See, also, Chapter 8).

Similarly, when Bickerton (1975, 1977) claims that tense and aspect markers arise in creoles without any linguistic input, he dismisses as unimportant the fact that children learning creoles choose the same lexical items to function as the new markers. However, their use of the same words suggests that there is, indeed, relevant input. This can be seen clearly in the case discussed by Sankoff and Laberge (1974) of the development of "bai" from adverb to future marker in Tok Pisin. The adverb "bai" originally had relatively free word order but is gradually being restricted to a position immediately in front of the verb (or with, at most, a pronoun intervening between "bai" and the verb). Once the adverb is in close association with the verb, children learning Tok Pisin may interpret it as a future marker. U.G. may have

guided the child's interpretation of the data such that "bai" is seen as a tense marker rather than an adverb but cannot be said to have acted in the complete absence of data: the child is unlikely to interpret "bai" as a tense marker if it is not closely associated with the verb. The child's perception of the data often differs from the adult's (see Chapter 6), so that while there may appear to be no triggering data as far as the adult is concerned, there is in fact relevant input data for the child. Many cases of apparent lack of input data may be accounted for in this way.

To say that principles of grammar might guide the child's perception of existing data is quite different from claiming that U.G. can operate in the absence of input. For example, Slobin (1966a) suggests that children are predisposed to use word order to express grammatical relations so that even in the acquisition of an inflected language like Russian, they will hypothesize fixed word-order, where no such order is apparent in the data.[2] Even if a language as it appears to adults has a relatively free order, the child's perception of the data may be such that one of the possible orders stands out more than others; the child may have the ability to focus on that order. Grammatical principles may guide the child's perception to produce this effect. That is, the necessary information is present in the data, provided that the data is seen in a certain light.

Another tendency to rely too much on innate principles is seen in arguments for the universal base. In order to explain the child's ability to acquire deep structure on the

basis of surface structure information, it is sometimes proposed that the child has innate knowledge of deep structure and only has to learn transformations (McNeill, 1966, 1970). This implies that deep structure is identical for all languages.[3] However, while the universal base hypothesis would remove the need to learn deep structure, and thus answer the logical problem of acquisition in this case, it would also make the transformational component extremely complex. The transformations on such a model could hardly be very restricted, since the distance between deep and surface structure would be great, and, in consequence, the transformational rules would have to be very powerful. Therefore, many transformational accounts would be possible for the same surface data, raising the projection problem, the question of how to rule out some of the proposals to make such grammars learnable.[4]

4.2. Overestimating the role of data.

One reaction, an overreaction, to the approaches which overestimate the role of a priori principles or fail to constrain such principles sufficiently is to deny that they have any role whatsoever. For instance, Derwing (1973) says:

> "To invoke 'innateness' is useless as an explanatory device - one might just as well invoke 'revelation'."
>
> (p. 65)

While this may be true of some of the more extravagant claims made for universal grammar in the field of acquisition, it is not the case once a restricted framework is adopted in which

very specific proposals can be made which yield particular predictions and explanations of deficiencies in the data.

Derwing's alternative is to propose an "inductive requirement" on grammars, by which he means that:

> "Some reasonable provision must be made for the 'learnability' of a (linguist's) grammar by a child solely on the basis of the data which is available to him."
>
> (p. 49; italics in original)

The data, together with some very general inductive learning principles, must, he claims, be sufficient for acquisition to be possible.

However, this claim overlooks a number of facts. In the previous chapter, it was shown that the child does not try out all the logical alternatives in grammar construction that are available to him. A truly inductive approach would predict errors of the form of (3.11), repeated here as (4.1):

4.1. *Is the man who tall is in the room

That is, from the data and inductive principles alone, one would have no reason not to try out structure-independent hypotheses as well as structure-dependent ones.

Further, the primary data is often unsystematic as far as information about ambiguity, paraphrase relations, scope of quantifiers, anaphora, etc., are concerned. As we have seen, on the basis of primary data alone, it is not possible to guarantee construction of a grammar which accounts accurately for the child's competence in One-substitution (the Det-Nom grammar) over one which adequately characterizes the primary data

but which does not capture all aspects of the child's knowledge (the NP-S proposal). It is hard to see how inductive principles alone could resolve such problems for the acquirer of language. If such factors can be accounted for by means of unlearned principles and cannot be accounted for on the basis of learning alone, it is difficult to see in what sense invoking innateness does not constitute explanation.[5]

Finally, Derwing's proposals suffer from a failure to take into account the question of negative evidence. Recently, there has been much interest in the role of negative data in grammar construction and the implications for language learnability of the fact that the child appears to get no significant negative evidence in the course of acquisition (Gold, 1967; Wexler, Culicover and Hamburger, 1975; Baker, 1978, 1979a).

4.3. Negative evidence and learnability.

There are two ways in which negative evidence might be of importance to the child. It is often pointed out that children learning a language have to make do with "degenerate" data, that they hear speech full of false starts, hesitations, mistakes, etc., and that they have no means of telling which of the utterances they hear are a genuine reflection of adult competence and which are the result of performance errors. This is one sense in which children get no negative evidence: they do not get told if an utterance they hear is in fact grammatical or not.

However, it has been argued (e.g. Snow, 1972) that the language heard by many (middle class, North American) children is in fact highly structured, consisting of short, simple and grammatical sentences or phrases, that mothers avoid degenerate speech when addressing their children. The child also seems to have the ability to ignore data that he does not understand, and this could include degenerate data. Roeper (1978) suggests that the child is equipped to select data for himself (so that the mother's attempts to provide short grammatical sequences are actually redundant). If the child has this ability, then the fact that he does not have information identifying performance errors will not matter.[6]

Furthermore, Braine (1971) has presented adults with simulated languages to learn, both with and without degenerate data (ungrammatical strings) in the corpus. He found that his subjects had no difficulty in learning the language even with 7% of degenerate data in the corpus; they were not told that there would be any ungrammatical strings but they were capable of detecting "funny" sentences, ones which did not conform to rules which they could discover for the majority of the strings in the corpus.

Thus, when there is no negative evidence about the grammaticality of the <u>input</u> to the child, this may not constitute a serious block to acquisition. On the other hand, the question of negative evidence and the child's <u>output</u> is far

more crucial. The problem arises where the child might overgenerate a structure and could only find this out by being corrected. This raises problems for many transformational analyses, as well as for the inductive and analogical principles argued for by Derwing (1973).

That children do not receive significant negative evidence has been suggested by Brown and Hanlon (1970) who show that parental corrections usually concern the truth value of the child's utterance, not his syntactic structures. McNeill (1966) shows that where children are corrected they ignore the correction. McCawley (1976) points out that this finding is only indicative of the child's reaction to correction at a particular stage of development, that there might come a time when correction would be effective. This is plausible; when the child is beginning to focus on data in a particular way, a correction that he can relate to his perception of the data might be absorbed and used to help his grammar construction, whereas when the correction bears no relation to his perception of the data, the child will ignore it. This still means that corrections will be an unreliable source of negative data, since the parent is unlikely to know whether the child is in a fit state to understand them or not, i.e. unlikely to be able to choose an appropriate time to make a correction.

McCawley (1979) suggests that the current emphasis on the lack of negative evidence available to children is mis-

guided, in that sources of negative evidence other than parental corrections may be available. For instance, the adult may fail to understand what the child is trying to say and the failure to communicate would constitute a form of negative data. However, while this may sometimes be true, in many of the cases that will be discussed below, the overgeneralizations that might occur are unlikely to cause comprehension problems. Furthermore, many of the potential overgeneralizations do not actually occur at all, so that the question at issue is not what form of negative evidence is or is not available but how to account for the forms that do occur on the basis of positive data only.

In investigating the problem raised by the lack of negative evidence, Baker (1978, 1979a) discusses two kinds of exceptional statement which might occur in a grammar. One kind can be arrived at on the basis of positive data only, the other requires negative evidence. It would be desirable, therefore, to exclude the need for the latter kind of exceptional statement altogether from the grammar.

An example of the harmless kind of exceptional statement is furnished by "how come", which does not trigger subject-auxiliary inversion. Data provided by the adult show instances of "how come" used without inversion:

4.2. How come he's still here?

which indicate clearly that this form is not followed by subject-auxiliary inversion, unlike other Wh question words.

The child therefore, does not require negative evidence to learn this exceptional form. Of interest also, and only briefly touched on by Baker, is the question of how the child learns subject-auxiliary inversion in the first place, since question formation in early stages of child language does not involve inversion with any of the Wh-interrogatives. For instance:

4.3. Why he is still here?

will be grammatical for many children at a certain stage, though the adult data do not provide evidence of this form, and the child changes from using such forms to using subject-auxiliary inversion without negative evidence.[7]

This case is similar to the learning of "chose" as the past tense of "choose", another of Baker's "benign" exceptional statements. Baker (also Braine, 1971) proposes that more specific rules are ordered before more general ones, so that the child does not need to be told that "choosed" is ungrammatical; he will automatically select the less general form "chose". Again, this proposal does not explain why there is a stage at which the child considers "choosed" grammatical, or how, without negative evidence, he gets from that stage to the realization that it is ungrammatical.

In both the case of the child's grammar which does not contain subject-auxiliary inversion at all and the grammar which contains only regular inflections, the positive data itself can nevertheless tell the child that he is producing

the "wrong" forms. That is, as soon as he notices the mismatch between the adult data and his own rule, he has a form of correction available to him; he perceives the data in a new light, in a way which is consistent with the positive data. His original guesses turn out to be too restricted for the adult data, though correct for his current perceptions of that data.[8] Adjustments can be made to a grammar which is too restricted for the data, whereas it is much harder to constrain a grammar which is not restricted enough.

Grammars which are not restricted enough constitute the other kind of case that Baker considers, the case where exceptional statements in the grammar would be bothersome since they would be motivated only by negative evidence. Baker looks at various forms of dative construction, traditionally described by the dative-movement transformation. If dative-movement is a transformation, then, given data such as:

4.4.a. We sent the book to George
b. We sent George the book
c. We reported the accident to the police

the child would be justified in generating the following:

4.5. *We reported the police the accident

If he did produce such strings, he would have no means of knowing from the positive data alone that such forms might not materialize at some future time to confirm his hypothesis. He has overguessed and has no means of knowing this fact; he would have to be told that "report" does not allow dative-movement.

Baker reports that children do not, in fact, come up with forms such as (4.5), which suggests that the lack of negative evidence is not crucial because children do not actually construct grammars which result in such over-generalizations. In other words, children do not have a rule of dative-movement at all.

Instead, Baker favours an analysis of datives where both the V NP NP structures and the V NP PP ones are base-generated, verbs being subcategorized for either or both. The child will not assume that a verb takes one or other sub-categorization without positive evidence of it, so there is no learnability problem, no occasion for problematical over-generalizations.[9] Oehrle (1976) proposes a similar analysis of dative constructions, independently of the learnability issue.

Baker gives similar arguments against the rules of "to be" deletion and relative clause reduction/modifier shift. That is, the fact that sentences like the following do not occur in child language:

4.6.a. *John happens sleepy
b. *the awake man

suggests that forms such as:

4.7.a. John seems to be sleepy
b. John seems sleepy

4.8.a. the man who is fat
b. the fat man

must all be base-generated, rather than there being a trans-formational relationship between them, with (4.7.b) and (4.8.b)

derived from (4.7.a) and (4.8.a) respectively. If there were a transformational relationship, forms such as (4.6.a and b) would be predicted and it would require negative evidence to rule them out. As Baker (1979a) notes, the absence of such forms causes considerable problems for Derwing who seeks to explain acquisition solely on the basis of the data and analogical principles, since forms such as (4.5) and (4.6.a and b) would be predicted on analogical grounds. The fact that they do not occur is a problem for inductive approaches and if they did occur there would be the problem of the lack of negative evidence to rule them out.

Commenting on the difference between the two kinds of exceptional statement that might be necessary in a grammar, Baker says:

> "Any exceptional statements in a grammar are harmless if there is some basic data available to the child learning the language that indicates the necessity of making those exceptional statements. On the other hand, exceptional statements for which a child's basic data provide no evidence are troublesome, and any grammar that contains such exceptional statements is immediately suspect."
>
> (1978, p. 426)

In order to constrain grammar construction so that the theory of grammar will not admit analyses which would require exceptional statements based on negative evidence, in order, for example, to exclude a transformational analysis of dative-movement which would require a statement that "report" cannot allow dative-movement to occur, Baker (1979a) proposes his own

"inductive requirement" on grammar construction, a requirement that differs somewhat from Derwing's:

> "If a linguist proposes a grammar G for some language, then G meets the 'inductive requirement' if it is accompanied by a set of hypotheses about human cognitive capacities from which G can be deduced, given primary linguistic data (i.e. data of just the kinds that a child obtains from his environment)."
>
> (Baker, 1979a, his (59))

Thus, while the primary data are crucial, the grammar is not induced from the data alone but deduced from a combination of the principles of U.G. and the data. That is, the difference between Derwing's view of the data and Baker's is that the former considers it both necessary and sufficient for acquisition while the latter considers it only necessary.

The question of learnability of grammars is very much bound up with the projection problem. In the cases discussed above, the positive primary data allow two possible analyses, a transformational one versus one where forms would be base-generated. Only one of these grammars, in each case the non-transformational one, is learnable without negative evidence. The linguist has negative evidence available to him in the construction of grammars but the child does not. Baker is proposing to exclude in principle (by means of his inductive requirement) the kind of grammar which requires exceptional statements based on negative evidence. In this way, the child will never have to consider two possible grammars for the positive data but will automatically be restricted to one.

This proposal seems justified for the cases where errors of overgeneralization never occur. That is, the non-occurrence of forms such as (4.5) and (4.6.a and b), and many others discussed by Baker (1979a), is good evidence that the child does not try out transformational analyses in these cases at all. Nevertheless, there are problems with his proposal, in that certain errors which children make do suggest the need for exceptional statements which cannot be arrived at from the basic data. Consider structures involving datives and their interaction with pronouns. The following sentence is grammatical:

4.9. I gave it to my mummy

whereas, for most adults, (4.10) is ungrammatical:

4.10. *I gave my mummy it

Yet Fischer (1976) reports that forms like (4.10) are preferred over forms like (4.9) by children of about 5 years of age, whose parents consider (4.10) to be ungrammatical. On the analysis proposed by Baker, or by Oehrle (1976), verbs like "give" will be subcategorized to be followed by either NP NP or NP PP. For the NP NP case, the child has to learn that the second NP cannot be a pronoun.[10] Fischer proposes an output condition and Oehrle (1976) a surface filter to rule out forms such as (4.10). The question is, how would the child learn the ungrammaticality of (4.10) if he is not corrected i.e. how would he acquire such a filter? There is no positive evidence in the data that this is an error. Such forms simply

do not occur, so there is always the possibility that they might crop up at some future time in the adult language sample. Whereas in the other cases discussed by Baker, the deviant forms were never generated in the first place, this is not the case here.

Baker (1979a) suggests that the explanation may have something to do with cliticization. He notes that similar overgeneralizations involving pronouns occur for particle movement. There is an identical learnability problem for forms such as:

4.11. *I turned off it

Children at one stage produce such forms and later do not, without the benefit of negative evidence. Baker proposes that when children produce forms such as (4.10) and (4.11), they are unaware of the enclitic status of object pronouns. As soon as they become aware of it, on the basis of other positive evidence in the data, forms such as (4.10) and (4.11) are ruled out without the need for exceptional statements based on negative evidence. Some explanation of this kind may well be appropriate, though if it is really a case of cliticization both for particles and datives, one would expect the realization of cliticization to affect the two constructions at the same time. That is, as soon as the child becomes aware of the enclitic status of object pronouns, then all object pronouns should be produced in positions which take account of this fact. However, Fischer (1976) reports

that 5 year olds go through a stage where they prefer dative constructions with the pronoun separate from the verb and particle constructions with the pronoun next to the verb. For these children the following will both be grammatical:

4.12.a. I saw a pretty flower and I picked it up
b. I saw a pretty flower and I gave my mommy it

This suggests that the same explanation cannot account for both cases. Nevertheless, while this specific proposal may not be accurate, some other proposal of this type could explain the children's behaviour if their perception of the positive data changes so that they become aware of some factor which will cause them to drop the "ungrammatical" constructions without need for negative evidence.

A similar problem for learnability arises with relative pronoun omission. In the grammar of Modern English, the following sentences are permissible for adults:

4.13.a. I saw the man whom I know
b. I saw the man I know
c. I saw the man who knows me

On the other hand, (4.14) is ungrammatical:

4.14. *I saw the man knows me

If on the basis of the positive evidence of (4.13.a, b and c) the child guesses a grammar which allows omission of any relative pronoun whether it is subject or object of the relative clause, he will generate forms like (4.14). He will then require negative evidence to find out that such forms are not permissible. Such possibilities cannot be ruled out by means

of an inductive requirement that the child should assume relative pronoun deletion is possible only where he has positive evidence of such deletion, because such forms do occur in child language. For example, my son Thomas for several months between the ages of 3 and 4 systematically omitted subject relative pronouns, producing forms like (4.14) and also like:

4.15. The man knows me is tall

(i.e. it made no difference whether the relativized NP was subject or object of its clause).

Chomsky and Lasnik (1977) propose a language-specific filter to rule out forms such as (4.14) and (4.15):

4.16. $*[_{\overline{NP}}$ NP tense VP$]$

(their filter (20))

Again, the question arises as to how this filter is learned. It cannot be assumed to be straightforwardly part of universal grammar, since subject relative pronoun omission was permissible in the grammars of Middle English and Early Modern English. Chomsky and Lasnik (1977) touch on this problem:

> "Suppose, in contrast, that the filter (20) is not universal but is simply a principle of the language under analysis. Then it must be learned. Specific evidence must be presented to show the child that (20) holds. The evidence might be negative evidence, i.e. corrections by the speech community when (20) is violated. Or it might be that only positive evidence, i.e. observation of grammatical speech, suffices. It is difficult to take the former suggestion seriously. That is, it is difficult to believe that everyone who knows the facts described above has been explicitly corrected (or has observed corrections) for violation. Dismissing this possibility, we must assign to the language faculty some principle that leads to postulating (20) as the 'unmarked case' unless there is specific evidence to the contrary."
>
> (p. 437)

Baker (1979b) considers a similar proposal, whereby:

> "Part of the grammar of every language consists of a 'filter component'. The filters listed in this component are specified universally, but languages differ in whether specific filters are assigned the value 'plus' (indicating that the filter applies) or 'minus' (indicating that the filter does not apply). The value 'plus' for a filter would be the 'unmarked' value."
>
> (p. 5)

In both these proposals, the filter is assumed to operate unless specific positive evidence to the contrary exists. That is, for the child learning Middle English, there would be positive evidence, in the form of omitted subject relative pronouns, that the filter does not apply, that the "marked" value is in operation. Baker actually rejects this solution for many cases because so many different filters would be required in the grammar, and proposes, instead, a series of "permitting filters" which would be motivated only by positive evidence. However, this would raise the learnability problem once again because the permitting filter:

4.17. $\lceil_{COMP}\ e\rceil$

would be motivated by object relative clauses which occur without relative pronouns and could then be overgeneralized to subject relatives.

Even if (4.16) is assumed to be the unmarked case, there is a problem if the data from Thomas are at all representative.[11] That is, having no positive evidence that the marked case is in operation in Modern English he ought to assume the unmarked

case. It is true that he does eventually do so, but what triggers this? Why does the filter not operate as soon as he starts using relative clauses? It is possible that these filters are not triggered by data but come into operation for maturational reasons, in which case "errors" involving the marked case may be expected before they come into operation. Alternatively, it may be that while the structures he produces look like relative clauses, they are not true relatives but in fact involve a conjoined clause analysis (as suggested by Tavakolian, 1978). In that case, the filter might not operate because the structure involved is not seen as an NP containing the sequence: NP tense VP. As soon as the child changes from a conjoined clause analysis to a true relative clause, the filter will become relevant and will come into operation.

So far, two situations have been considered as regards the relationship between exceptional statements and positive data. The first kind of exceptional statement required in grammars causes no learnability problems because the exceptional nature of any particular form is obvious from the positive data alone, as was the case with "how come" or "chose", discussed above.

Bothersome exceptional statements, on the other hand, are those which are required to rule out the over-application of an optional transformational rule to situations where it may not apply. Thus, dative-movement is traditionally thought

of as an optional rule with certain verbs being marked as exceptions to this rule, in order to prevent ungrammatical forms like (4.5) from occurring. Information to motivate such exceptional statements is not available in the positive evidence alone and, therefore, negative evidence is required, an undesirable consequence since children do not seem to get corrected in the relevant respects. To avoid the need for such exceptional statements, Baker proposes his inductive requirement that no grammar should be acceptable which cannot be deduced from the positive data alone, interacting with U.G.

This has certain implications for the way grammars are written. For example, how should verbs which are exceptions to general rules be entered in the lexicon? It is customary to enter an exception in the lexicon as [- rule X]. If Baker's requirement of positive evidence only were extended to these cases, instead of marking a verb like "ascertain" as [- COMP deletion], to show that the complementizer cannot be deleted after this verb, one would have to mark all verbs that can take the rule, for example "say", as [+ COMP deletion], since one would never know whether verbs like "ascertain" might not at some time undergo the rule. Under Baker's (1979b) proposals this would be equivalent to subcategorising each verb for the complementizer configurations it could be followed by, just as it is subcategorized for other configurations.

The proposal to rule out bothersome exceptional statements by disallowing the grammars that give rise to them is supported

in those cases where the child never makes the kind of error which the optional transformational account would predict. Problems occur in those cases where the child does make the kind of overgeneralization that would seem to require negative evidence to suppress it. It is interesting that many of these involve proposals concerning filters in the grammar, which may be a reflection of the fact that filters themselves are often attempts to cope with bothersome exceptional statements. In some cases, as we have seen, the child's changing perception of the positive data will result in the dropping of the exceptional form without the need for any negative evidence. Where this is not the case, reliance is placed on calling the filter which would be difficult to learn "the unmarked case". This, however, is no real explanation but a restatement of the problem, the problem being that of an apparently language-specific filter which has to be learned although there is no negative evidence to show the need for such a filter. To insist that the "unmarked case" can function without positive evidence, to assume that a filter comes into operation unless there is specific evidence to the contrary, is a description of the same facts under a different name. This problem will be discussed further in Chapter 8, on Acquisition and Markedness, where I shall consider the kind of evidence that would show whether the filters constitute the unmarked case or not.

FOOTNOTES TO CHAPTER 4

1. Though it is always difficult for an experimenter to be sure of this, since subjects may use forms in his absence which they do not use when he is present and the child is also exposed to language from other people who may use the relevant forms.

2. But see Maratsos (1978) for discussion of cases where this proposal is not supported.

3. It is not clear that this really is McNeill's position since he does discuss the learning of base word order for individual languages. There has been a certain confusion in the literature due to a tendency to equate deep syntactic structure with semantic structure.

4. Wexler, Culicover and Hamburger (1975) discuss the learnability of transformational grammars off a universal base. They find such grammars unlearnable. However, with enrichment of the data by assuming that the child can infer the interpretation of a sentence from context and then proceed from the interpretation to deep structure, together with certain restrictions on transformational power, they were able to prove learnability.

5. It is true that it may prove to be the case that what are now considered to be a priori linguistic principles might be something rather different, a subset of other cognitive universals. But, even so, this does not lessen the explanatory power of innateness but merely means that we might one day

know more about such principles. In any case, the way things are at present, far more precision is to be achieved by pursuing specifically linguistic universals than by looking for more general principles. It is foolish to ignore the kind of information that is available just because the explanations that follow from it may turn out to be subsumed under some other kind of explanation, the nature of which is as yet very unclear.

6. Of course, it may be precisely the child's a priori knowledge which gives him the ability to ignore aspects of the input, although it would also seem to be due to general perceptual factors.

7. Baker (1979a) discusses this in slightly different terms. He considers it a problem that the child might be provided with data of the following kinds:

i. What will he do?

ii. I don't know what he will do.

If he noticed both these sentence types, with inversion in main clauses and no inversion in subordinate, and assumed that subject-auxiliary inversion was an optional rule, he could not get the obligatory status of subject-auxiliary inversion without negative evidence. However, it seems unlikely that the child pays much attention to subordinate clauses at the time that he is learning subject-auxiliary inversion, and if he does the facts will follow from structure-preservation (see discussion of the interaction of the principle of structure-preservation and data in Chapter 3).

8. Children's forms such as "choosed" are often referred to as overgeneralizations in the literature. But though they are overgeneralizations of a particular rule into areas where it would not apply in adult language, they are restricted in the sense that the child has assumed that one rule will be sufficient for the adult data, where in fact several may be required.

9. Fischer (1976) reports that children acquire the V NP PP structure before the V NP NP. This fact in itself says nothing in favour of either the dative-movement analysis or the lexical analysis, although she assumes the former. While it could suggest a base form being learned before a transformation, it is just as possible that the base-generated V NP NP structure is more difficult than a base-generated V NP PP and that the former is, therefore, acquired later.

10. Actually, that it cannot be an unstressed pronoun for the adult grammar. Fischer (1976) reports that the children treated stressed and unstressed pronouns alike.

11. Although there have now been a number of investigations on relative clause acquisition, particularly on comprehension (Sheldon, 1972; Prideaux, 1978; Tavakolian, 1978), there has been little discussion of subject relative pronoun deletion. In studies investigating comprehension rather than production, such omissions might not be noticed at all. However, Cook (1973) found that, in an imitation task, young children frequently omitted subject relative pronouns.

CHAPTER 5

THE PSYCHOLOGICAL REALITY OF GRAMMARS

5.1. Three views of psychological reality.

In the previous chapter, some of the implications of "learnability" have been discussed. That grammars must be learnable presupposes that it is grammars that are learned. There has been considerable disagreement as to whether grammars are or are not psychologically real, so this matter deserves further attention, since it is the source of much confusion in the literature.

By the term "psychological reality", it is understood that:

> "A person's knowledge of language can properly be represented as a system of rules of grammar, and that process models concerned with language use will incorporate such representations of linguistic competence."
>
> (Chomsky, 1976, p. 15)

Three different positions have been taken on this issue. The first, as outlined in the above quotation, is that the best linguistic description is psychologically real, that the grammars proposed by linguists can also be considered, in some sense, to be "mental" grammars. This does not mean that the grammar is to be considered as a processing model in itself but only that it will be incorporated into such a model. For

example, Bresnan's (1978) proposals can be realized by an Augmented Transition Network (ATN) parsing system (Wanner and Maratsos, 1978) but that system is not itself a grammar although it includes one.

The second position, held, for example, by Bever (1970), Fodor, Bever and Garrett (1974), Sheldon (1978), is that a person's knowledge of language cannot be represented in any "real" sense by the rules of a grammar and, therefore, that such a system will not be incorporated into a model of language use. In other words, the linguist's task is to provide a description of facts about language but not to impute any knowledge of the resulting system to the speaker/hearer/acquirer. On this view, then, there are no "mental" grammars; the processor operates with reference to heuristic strategies rather than a grammar. The third position is that speakers do have mental grammars but that these are not necessarily the same as the grammar which the linguist would consider optimal, a position which leads to a great deal of confusion, as we shall see.

5.2. The best linguistic description is psychologically real.

If linguistic theory is to provide a grammar of adult competence, it makes no sense to propose a grammar which is an adequate description of that competence and yet is not psychologically real. The best description is the one which offers the best account of the facts. To revert to the case of One-substitution, discussed in Chapter 3, the NP-S grammar is

adequate as a description of the primary data but inadequate if non-primary data are taken into consideration. Since the adult competence includes knowledge of these non-primary data, of ambiguity, paraphrase and ungrammaticality, the NP-S grammar cannot be psychologically real. The grammar which provides the best account of the primary and non-primary data is the Det-Nom grammar, and this is the one which must, by definition, be psychologically real. If some evidence emerges to show that the Det-Nom grammar is not psychologically real, then it will also not be the best description of the facts.

Psychological reality, then, is simply a matter of the correctness of a linguistic description. This is argued by Kiparsky (1968b) who shows that the "brace notation" must be the correct, and hence the psychologically real, way to represent significant linguistic generalizations. What appears to worry many people in cases such as Kiparsky's use of the term "psychological reality" is that his evidence concerns only the product of the grammar i.e. the language that is spoken, and intuitions about it, and says nothing about the process of producing that language. There seems to be a feeling that only the latter should be considered psychological, because that is what psychologists study. Thus, data from reaction times, recall experiments or language pathology are somehow felt to offer more insight into mental grammars than hypotheses about data consisting of sentences people utter or insights into the grammaticality of such

sentences. Bresnan (1978) distinguishes between the "grammatical characterization" problem and the "grammatical realization" problem. Kiparsky's proposals would fall into the former, experiments on reaction times into the latter. But while this distinction can be made, it does not affect psychological reality. Characterization and realization must be two sides of the same coin, both potentially having evidence to offer as to the nature of grammars and neither being more "real" than the other.

The theory of grammar then, will have implications for a model of language use and vice versa. The predictions that can be made will depend both on one's theory of grammar and on one's theory of language processing or production. One of the earliest theories of processing/production that was adopted with respect to T.G.G. was the Derivational Theory of Complexity (D.T.C.). As the two positions, mentioned in section 5.1., which do not accept the psychological reality of the linguist's grammar stem from the results of experiments done with the D.T.C., it is worth looking at this work in some detail.

The D.T.C. in its earliest form states that the complexity of a sentence is a function of the number of grammatical rules involved in its derivation. Thus, for example, a passive sentence involves more transformations than an active one and should, therefore, be more "complex", as revealed by longer reaction times for comprehension of passives, longer times to

recall them, later acquisition, etc.[1] Since it is not possible to tell if all transformations are equally complex, if passive is as complex as affix-hopping for example, a refined version of the D.T.C. was proposed by Brown and Hanlon (1970). They suggested that structures should only be compared which shared the same derivation except for one or two transformations; for example, a negative question should be more complex than a negative statement, the two differing in that the former has an additional question transformation.

At first, experiments seemed to support the D.T.C., both in its original version and in the version proposed by Brown and Hanlon. (See Fodor and Garrett, 1966, and Fodor, Bever and Garrett, 1974, for a review of the literature). But the original experiments were followed by a great many which did not confirm the D.T.C. For example, on the model of T.G.G. which these experiments were based on, namely that proposed in Chomsky (1957, 1965), truncated passives (those lacking an agent in a by-phrase) were considered to be derived from full passives which had undergone deletion of the agent by-phrase. As deletions were transformations in this model, this meant that short passives were derivationally more complex than long.[2] Yet the evidence shows that short passives are either acquired before long ones (Watt, 1970; Bever, Carroll and Hurtig, 1976) or at the same time (Maratsos and Abranovitch, 1975) but certainly not after them. Again, the derivation of pre-nominal adjectives from relative clauses (Smith, 1961) suggests that

pre-nominal adjectives must be complex but, once more, this is disconfirmed by the evidence, since pre-nominal adjectives occur in early two-word speech (Braine, 1963a), long before relative clauses.

5.3. Arguments against "mental" grammars.

Fodor, Bever and Garrett (1974) review the literature and conclude that the D.T.C. is not supported by the experimental evidence. There are a number of ways one can proceed from this conclusion, some logical, some rather extreme. That is, the failure of the experiments might indicate something at fault with the theory of grammar or something wrong with the processing model, or, indeed, with both. But it is precipitate to claim, as Fodor, Bever and Garrett do, that the results of the experiments indicate that there is no relationship between grammars and models of production or processing, without first investigating revisions of the theory of grammar or other processing models, either or both of which would still allow a close relationship between the grammar and models of language use. The position of Watt (1970), on the other hand, is to maintain that both the theory of grammar and the processing model implicit in these experiments are correct and, accordingly, he has to do a considerable amount of patching up, the chief effect of which is to separate the linguist's grammar from the "mental" grammar, or grammar of competence, an undesirable consequence if explanatory adequacy is to be achieved, but a division which, unfortunately, continues to exist.

Fodor, Bever and Garrett discuss in passing, but dismiss as relatively unimportant, the possibility that the theory of grammar might no longer propose the kinds of derivation on which the D.T.C. was tested. Yet this is exactly what has happened. For example, many proposals for passive now involve only one transformational rule, NP preposing, and the agent by-phrase is base-generated. If there is no agent by-phrase, as in truncates, there will be no need to base generate one. Even if truncates were considered to have an empty by-phrase in deep structure, deletions are no longer considered to be transformations anyway. Therefore, no distinction is predicted between full and truncated passives (though the former might take longer to interpret). Similarly, adjectives in pre-nominal position are no longer derived from relative clauses (Baker, 1975). Many other derivations which caused problems for the D.T.C. have been revised so that they can no longer be used as counter-arguments to the claim that transformational complexity involves production or parsing complexity. In other words, the experimental evidence could have been used to indicate problems with certain aspects of transformational grammar at that time. In fact, the evidence was not usually used in this way and the changes that were made in the derivations that happened to be causing difficulties for the D.T.C. were arrived at on the basis of other considerations which had nothing at all to do with the experimental results.

In any case, not all of the problems with the D.T.C. could have been removed by making revisions in the theory of grammar. It is, therefore, necessary to consider whether the problem lies with the D.T.C. itself as a theory of processing. Originally, the D.T.C. was proposed as an obviously over-simplified initial hypothesis for part of a performance model, though this was sometimes lost sight of. Arguments against the D.T.C. as a processing model are summarized by Fodor and Garrett (1966), Fodor, Bever and Garrett (1974), who show the problems that result as far as search space, memory limitations, etc., are concerned for analysis-by-analysis or analysis-by-synthesis models, which are implicit in the D.T.C. They point out that as a model of processing it fails to account for other forms of complexity, such as non-linguistic or perceptual complexity, and, of course, it only accounts for one kind of linguistic complexity, namely trans-formational, so that it is far too simplistic a model. They also argue that it fails to predict the unit of perception (which they consider to be the clause). But Frazier (1979) suggests that the D.T.C. in fact presupposes that the unit of perception is the entire sentence; a complete sentence would have to be available before analysis could begin, in order for the processor to assess what transformations might be involved, so that no analysis of the incoming signal could take place as it arrives. Frazier convincingly argues against analysis in terms of the entire sentence. In fact, the D.T.C.

does not seem inconsistent with the clausal unit proposed by Fodor, Bever and Garrett, since the clause will be the domain of many transformations, but it is inconsistent with the phrasal units of perception proposed by Frazier, since her units are arrived at by means of a combination of syntactic phrases and memory limitations and are not consistently clauses or phrases.

But the failure of a particular processing model, in this case the D.T.C., does not mean that there is no relationship between a grammar and models of language processing. All it indicates is that another, more satisfactory, model of processing must be sought. And it is perfectly possible that other processing models will retain a close relationship with the grammar, that they will incorporate a grammar in a fairly concrete way. Indeed, from the point of view of explanatory adequacy, only such processing models will be of interest. A number of recent proposals for parsing models do relate closely to various versions of E.S.T., for example Wanner and Maratsos (1978), J. D. Fodor (1978), Frazier (1979). I do not wish to discuss these models but merely to point out that the logical alternative to the failure of the D.T.C. is not to assume that the relationship between grammar and processing must be very abstract, or even non-existent, but to try out alternative models.

The two possibilities that have been considered so far have been that the experimental evidence might have indicated

that the theory of grammar required some revision, or that the theory of parsing needed alteration. It is indeed possible that both were in need of revision.[3] But the need to revise both one's theory of parsing and one's theory of grammar does not automatically lead to the conclusion that the relationship between grammar and processing must be very remote. Nevertheless, this is maintained by Fodor, Bever and Garrett (1974) who leap from observation of the failure of the D.T.C. to the position that a "grammar is only very abstractly related to a model of the adult sentence producer-perceiver". While they accept that recognition procedures must somehow "employ the kind of information that is represented by grammatical rules" they do not accept that this information is used in the form of a grammar. They are convinced only of the psychological reality of the structural descriptions proposed by the theory of grammar and not of the transformations or derivations. It is not clear that one can really make such a distinction, since presumably the structural description of a surface structure (which the parser has to deal with) must include, in some form or other, information about its own derivation, about the transformations that have led to that surface structure.

As a result of this rejection of a close relationship between grammar and performance, attention has turned to other areas which might relate more significantly to a model of language use. There has been particular interest in

heuristic strategies, perception strategies, learning strategies, etc, which might guide a language parser (e.g. Bever, 1970; Slobin, 1973; Fodor, Bever and Garrett, 1974). Certainly a theory of grammar cannot be expected to account for all aspects of parsing complexity, some of which will arise from properties of the parser being assumed, and certain facts about parsing may well be described by means of strategies rather than in terms of a grammar. However, it is by no means clear that perception strategies alone can really provide a more realistic model of language processing than models which explicitly incorporate a grammar. See Frazier (1979), for example, for a discussion of parsing problems that arise with some of Bever's (1970) perception strategies, and for ways to incorporate a grammar directly into a parsing model and Roeper (1978) for a theory of acquisition that specifically includes a grammar as part of the model.

Implicit in much of the use of terms like "perception strategy" is the assumption that perception strategies are psychologically real simply by virtue of the word "perception", by virtue of the fact that they do not rely on grammars. However it is not enough to dismiss the psychological reality of grammars and take for granted the status of such strategies. The psychological status of perception strategies is equally questionable. As Bever (1970) discusses, one must have a clear idea of what a perception strategy can be, how language specific strategies are acquired, and so on. As it is, perception strategies are often invoked to cover some fact, to

describe some inconvenient data rather than to explain it; the implicit assumption that any such strategy must automatically be psychologically real acts as an excuse for rather implausible proposals.

For example, as discussed in Chapter 3, Parker (1977) proposes a strategy that children pay attention to embedded clauses even in the very early stages of speech, since he wishes to argue that German children postulate SOV word-order in their earliest grammars, and SOV order in German occurs only in subordinate clauses, whereas main clause order is SVO. Such a strategy is rather unlikely since the evidence suggests that speech directed at children contains few embedded sentences (Pfuderer, 1969; Drach, 1969; Snow, 1972). This means that children would have actively to seek out embedded sentences in other speech samples, again rather unlikely, since children appear to ignore embedded sentences even when they do occur (Solan and Roeper, 1978). Thus, Parker's proposal for SOV order rests on an implausible perception strategy, so that if there is no support for that strategy, the proposal itself must be questioned or given more solid justification. In using perception strategies either to replace or complement the role of the grammar in language use, it is important that the proposed strategies can themselves be viewed in the context of a theory, so that one can decide whether such strategies contribute towards explanation or not. The problem of psychological reality does not disappear simply by removing grammars from consideration.

5.4. The separation of the linguist's grammar from the psychologically real one.

Another position which was adopted after the failure of the D.T.C. experiments to confirm a direct relationship between processing and transformational complexity was the view that speakers do have mental grammars but that these are distinct from the optimal grammar which a linguist might construct. This is the stand taken by Watt (1970) who proposes that for just those cases where the D.T.C. is not borne out, such as in the comprehension or acquisition of truncated passives, prenominal adjectives, etc., speakers have a mental grammar which will lead to derivations which have the complexity that the D.T.C. predicts, while the linguist's grammar retains the derivations which were not supported by the experimental results. He specifically considers and rejects a solution which would require the linguist's grammar itself to be revised, arguing, for instance, that to change the underlying form of the passive would be to lose the generalisations which it was originally intended to account for. (Nevertheless, the proposals which he considers and rejects for the linguist's grammar for passives have subsequently been adopted in E.S.T. for reasons independent of the psycholinguistic arguments).

Watt notes that young children produce truncated passives quite early and suggests that in the mental grammars of children these passives are analysed as adjectival and treated accordingly. The optimal grammar for the linguist is,

he assumes, to derive truncated passives from full passives by deletion, the full passives themselves being transformationally related to their active forms, the then standard derivation for all passives. He considers it implausible that children should alter the analysis of truncates in their mental grammars when they become aware of full passives and, hence, that their mental grammars differ from the optimal grammar. A similar argument is made by Maratsos (1978). Rather than making this distinction between mental and optimal grammars, the evidence from child language could have been used to suggest that the optimal grammar itself was wrong or, in fact, non-optimal. Indeed, recent proposals for the derivation of passives suggest that this is the case. Wasow (1977) distinguishes between adjectival and verbal passives. Only the latter are derived transformationally; the former are related to their active forms by lexical redundancy rules. Now, if the truncates that children produce early on are in fact adjectival passives, then there is no difference between their mental grammar which analyses them as adjectives and the optimal grammar which does the same thing, and which continues to analyse them adjectivally in all the grammars that the child may pass through.[4] Judging by the examples given by Watt and others (e.g. Bever, Carroll and Hurtig, 1975) of early truncates, they are indeed what Wasow would classify as lexical passives.

The distinction between lexical and verbal passives

now proposed may help to clarify some of the apparently contradictory results obtained in many early experiments on the acquisition of passives. As already noted, Watt and others suggest that truncated passives are present in children's grammars before full ones because they appear earlier in spontaneous speech. Maratsos and Abranovitch (1975), on the other hand, present evidence to show that children comprehend full and truncated passives at the same stage, arguing that this is evidence that full passives and truncates must be treated together in child grammars. The passive forms that they use seem to consist almost entirely of what Wasow would consider to be verbal, or transformational, passives.[5] Thus, they do show that there is no difference in the stage of acquisition of a full or a truncated verbal passive, a fact which, as they point out, is no problem for theories which propose a similar derivation for these forms, such as E.S.T. or Bresnan (1978), though it is a problem for the D.T.C. But since they did not investigate adjectival passives (which had not as yet been distinguished in the literature), their result says nothing about a difference between the acquisition of adjectival truncates, as opposed to verbal truncates and full passives, or about a dual analysis based on these lines.

This view of a difference between optimal grammars and mental grammars is still very prevalent, particularly in the context of language acquisition. For example, J. D. Fodor and

Smith (1978) argue that there are two situations in language learning which would lead to the failure to learn the grammar the linguist considers optimal. On the one hand, the child might be a "blinkered" learner and simply not think of the optimal solution. This stems from a confusion over what is meant by optimal. Fodor and Smith appear to think that the same grammar can, indeed must, be optimal for various dialects, for the child and for the adult. They then argue that the evidence suggests that the child does not acquire the optimal grammar, because it is impossible to determine what the optimal grammar actually is in certain circumstances. But a grammar can only be optimal with respect to a particular set of data. If the child fails to see an optimal grammar it is because it is not, in fact, optimal for his perception of the data (see Chapter 6 for further discussion).

Their second proposal is that the child might be a "lazy" learner; he might notice the optimal solution but fail to reanalyze his grammar in accordance with it because his own version works and he cannot be bothered to change it; there is a psychological cost in doing so.[6] This is analogous to Watt's proposal that it is implausible for the child to abandon his mental grammar where it differs from the optimal grammar. The same position is advanced by Maratsos (1978), who argues that evidence from child language on the acquisition of short passives, adjectives and VP complements suggests that these are not first acquired with their own base rules

and later restructured in terms of long passives, relative clauses, or subjectless S complements, because such reanalyses, whilst capturing apparent linguistic generalisations, involve the child in extensive restructuring, where his original grammar could still deal adequately with the data.

But if his original grammar can still deal adequately with the data, then perhaps it <u>is</u> the optimal grammar. Maratsos tries to exclude restructuring on psychological, or even logical grounds, whereas some of these reanalyses will simply be excluded by current theories of grammar.[7] Rather than excluding restructuring, which I shall argue in Chapter 7 does indeed take place in children's grammars, Watt, Maratsos, and Fodor and Smith could remove their pseudo-division between children's real grammars and linguists' optimal grammars by accepting that evidence from child language is evidence about optimal grammars, that restructuring is not necessary in these particular cases because the child's optimal grammar and the adult's coincide, because they view the data in the same way for these structures. This says nothing about restructuring for other cases. There will be many occasions when the child's optimal grammar at one stage is not the same as his optimal grammar at another stage; then restructuring will be necessary.

Indeed, Maratsos does want to use evidence from child language to argue for the psychological reality of the theory proposed by Bresnan (1978). He also wants to discredit the theory of grammar current in Chomsky (1965). By referring to

the latter as the linguist's optimal grammar, which it is no longer considered to be in that form (a fact he fails to point out), Maratsos is trying both to show that the optimal grammar (Chomsky's) cannot be real and that someone else's optimal grammar (Bresnan's) can be. This involves some rather loose use of the term "optimal". Both theories cannot be optimal; both cannot be psychologically real. If Bresnan can show that hers is the optimal grammar, then psychological reality follows automatically, as she herself recognizes, and there is no need for Maratsos to argue for the unreality of someone else's optimal grammar, because that grammar cannot also be optimal. The acquisition data can help to decide between theories only if the term "optimal" is reserved for the real, mental grammar.

The failure of certain psychological experiments to confirm the predictions of the D.T.C. could reasonably, assuming the psychological reality of grammars, have been considered to indicate something at fault with part of T.G.G. or with the D.T.C. as a theory of processing, or, indeed, with both. But it was over-reacting to that evidence to accept it as proof that no close relationship exists between a grammar and a model of processing. To reject the concept of the psychological reality of grammars on the basis of one, subsequently altered, form of the theory, rather than considering revisions of that theory or alternative parsing models, seems totally unjustified. On the other hand, it would seem to be under-

reacting to the evidence to try and ignore it altogether and propose, in order to account for the inconsistencies, a division between mental and optimal grammars, a division which prevents any useful kind of interaction between a theory of grammar and data from acquisition or information from language processing.

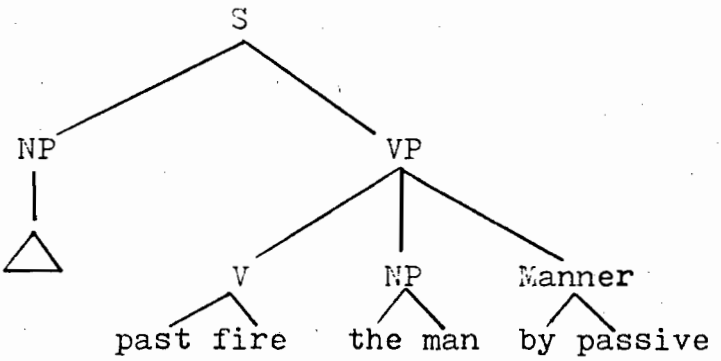
In conclusion, with the goal of explanatory adequacy in mind, we must assume that the grammar that the linguist proposes as optimal is the psychologically real one and, hence, that evidence from acquisition or processing can be relevant to the formulation of grammars. Although experimental evidence can give some insights into psychologically real grammars, given sufficiently well-formulated theories of parsing, this evidence has often been of somewhat limited use because it is not specific enough: evidence from processing experiments may be consistent with a number of theories of grammar (Maratsos and Abranovitch, 1975; Fodor, Bever and Garrett, 1974) or different processing models may be consistent with one theory of grammar (Dresher and Hornstein, 1976). Therefore, a fruitful approach to the determination of the psychologically real grammar is likely to continue to be the linguist's search for the optimal grammar.

FOOTNOTES TO CHAPTER 5

1. Many of the early experiments concerned claims about the comprehension of sentences: the more transformations involved in a sentence, the longer it would take to process. Although production does not necessarily involve a mirror-image of comprehension, it was also assumed that the D.T.C. would predict production complexity. The predictions were further extended to language acquisition: the more transformations in a derivation, the harder that structure would be to acquire and, therefore, the later it would emerge. I am discussing all the evidence together, even where the experiments do not specifically bear on acquisition.

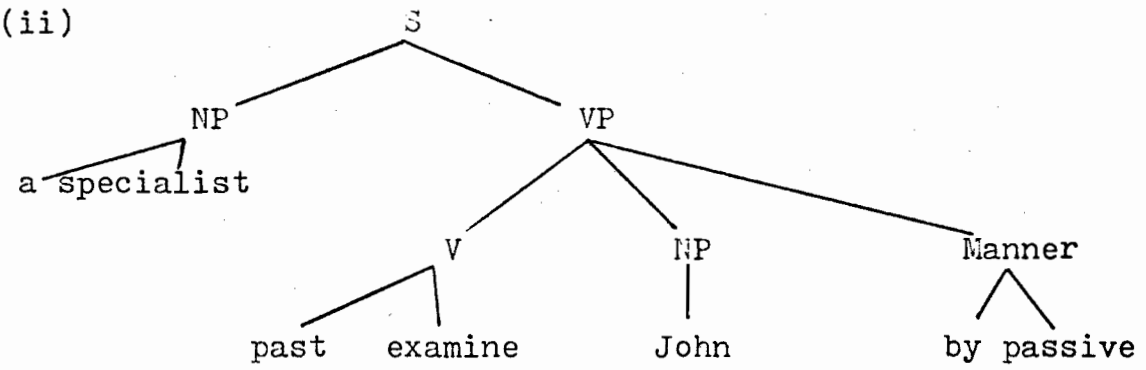
2. Actually, it is not clear that this prediction is correct even in the "Aspects" framework. In "Aspects", Chomsky proposes that the deep structure for a truncated passive is as follows:

(i)



The object moves into subject position and the unspecified agent by-phrase is deleted producing "the man was fired". Thus, two operations are involved. A full passive, on the other hand, has the following deep structure:

(ii)



Here, the subject NP must move into the agent by-phrase and the object NP must move into subject position deriving "John was examined by a specialist". Again, two operations are involved, so it is not clear that even on the predictions of the D.T.C. there should be any difference in the complexity of the two forms of passive (assuming that the deletion and the agent post-posing are equal in complexity).

3. Though the relationship between grammar and parsing does not have to be one of transformational complexity (see, for example, Fodor (1978), Frazier (1979) on parsing complexities that arise from the grammar incorporated into a model of processing), transformational complexity is not necessarily excluded, just because of the failure of early attempts to relate processing complexity to transformational complexity. For example, if parsing difficulties occur with "gaps" in incoming sentences (Fodor, 1978), then problems may arise in derivations where a number of movements have taken place. In other words, the more moves there have been, the more gaps (marked perhaps by traces) are likely which could

give rise to parsing ambiguities. Thus, derivational complexity might still be a source of parsing complexity, though by no means the only one.

4. In fact, Wasow suggests that some passives may have both an adjectival and a verbal form. This would not, however, involve the child in any extensive restructuring. That is, if he acquires the adjectival passives first, he does not lose that analysis but simply adds a transformational passive rule to his grammar.

5. A brief examination of some other experiments on the acquisition of passives (e.g. Slobin, 1966b; Strohner and Nelson, 1974; Turner and Rommetveit, 1967) suggests that it is the verbal passive that has most commonly been investigated. However, since many of the experimenters do not report a full list of their stimulus sentences but only give a few examples, it is difficult to establish this point conclusively.

6. This is not the same as the argument that the adult grammar may be non-optimal because of the adult's (supposed) inability to restructure, so that only the next generation can produce the optimal grammar (Kiparsky, 1965). In this case, the adult is deemed incapable of change, whereas Fodor and Smith are saying that the child is able deliberately to ignore possibilities for grammar change.

7. Karatsos is somewhat unfair in his criticism of T.G.G. in relation to acquisition data. He points out the superiority of grammars which do not have abstract deep

structures, and hence to the superiority of Bresnan's (1978) approach, without making it clear that current E.S.T. also restricts abstract derivations. Either of these approaches will reduce the amount of restructuring that must be imputed to child grammars. Maratsos's remarks are relevant to those who are still working within the "Aspects" framework but do not, in fact, exclusively support Bresnan's account.

CHAPTER 6

THE ACQUISITION OF THE OPTIMAL GRAMMAR

6.1. The relationship of grammars to data.

It has been argued in the previous chapter that it is senseless to dissociate the psychologically real grammar from the linguist's optimal grammar. Indeed, within the framework adopted here, it must be assumed that the grammar that the child learns is by definition optimal. This is a methodological assumption which has important consequences for the way I shall view data from acquisition. By the "optimal grammar" is meant the best description, the one which accounts most elegantly for the facts, for primary data from grammatical sentences as well as non-primary data concerning knowledge of ungrammaticality, paraphrase and ambiguity relations, scope of quantifiers, etc., interacting with a suitable performance model and the requirements of learnability.

The term "optimal grammar" may be slightly misleading in that it seems to imply that there is only one such grammar for any language, and this is often how it is understood. However, this depends on what is meant by "language" in this context. I shall argue that any set of data will be describable by a grammar which is optimal for that set of data and that set of data alone. The child at one stage will construct a grammar for the data of that stage; at another stage

he will construct a different grammar. Both will be optimal. Similarly, facts about the adult's language will be dealt with by an optimal grammar[1] which is not the same as the child's optimal grammar. Various adult dialects will each have their own optimal grammars. Thus, while there is only one optimal grammar for any language in the sense that "language" refers to a uniform set of data, any language, in the more general meaning of that word, will have as many optimal grammars as it has dialects, and a child will progress through as many optimal grammars as he or she has stages of acquisition.

A grammar, then, is optimal with respect to the data which it encompasses. Hence, it makes no sense to compare a grammar directly with another grammar, irrespective of the data, in order to determine which is optimal. Yet this is often what is done with child grammars, which are sometimes compared with adult grammars and sometimes with other child grammars of different stages. But it is only in the later stages of acquisition that the child's view of the data is likely to coincide with the adult's, unless one adopts an "instantaneous" view of acquisition[2] (see Chapter 7). In most cases, the direct comparison of grammars for optimality will be inappropriate.

Pit Corder (1967) makes a useful distinction between the learner's "input" and his "intake". For first language learners, this means that all the language uttered by the adult is not absorbed by the child. The child's perception of the

output of the adult will vary. For example, assuming that the child's output in some way reflects his competence, although it is by no means the only way of determining that competence, if the parent of a young child produces sentences like:

6.1. Here is your teddy

and the child responds initially with:

6.2a. Teddy

but at a later stage with:

6.2b. Here Teddy

the data provided by the adult has not, apparently, changed.[3] Nevertheless, it may have done so for the child in that he is maturationally more capable, at the later stage, of attending to more data (Slobin, 1973). Sometimes, the child appears to have an ability to "switch off" and attend to only part of a corpus (Braine, 1971; Shipley, Smith and Gleitman, 1969; Waterson, 1971). This ability does not seem to be confined only to early stages of acquisition. Roeper (1978) suggests that the child is equipped with an input filter which excludes unknown material and admits material which fits in with the child's current hypotheses. By ignoring problematical material, the child may be able to make data fit a hypothesis which he already has. For example, by excluding relative pronouns, the child may be able to treat relative clauses as instances of conjoined clauses (Tavakolian, 1978). What the child fails to notice in the language samples provided by the adult will depend partly on the hypotheses he has already

adopted, and as the child's hypotheses develop, so his view of the data will change and different things may be ignored.

If the child has a different view of the data from the adult, then clearly his grammar will not be constructed with a view to covering the same range of data but only to deal with his own perception of it. This raises the potentially difficult question of how one determines what the child's perception of the data is. However, if a grammar is psychologically real, a grammar of competence, as I have argued it must be, then the child's linguistic behaviour provides evidence of what his perception of the data is: we can tell from the forms he produces and the judgments he makes what his view of the data must be.

It may,at first sight, seem to be departing from standard usage to define "data" to include the child's perception of the data rather than something "objective" like "the sentences of a language". However, even in writing grammars for adults, it is the case that the data to be accounted for implicitly include the way the data are perceived. This is quite clear in the cases of ambiguous sentences. For instance, the sentence:

6.3. Flying planes can be dangerous

can be perceived in two distinct ways, each of which must be accounted for by the grammar. One account will be correct for one interpretation of the sentence and one for

the other and the accounts are not interchangeable. Thus, incorporating the child's perception of the data into our understanding of the data to be accounted for by the grammar is merely making explicit somthing that is already done in the writing of grammars.

McCawley (1976, 1979) also argues that the child's perception of the data differs from the adult's, and from his own at other times, but he appears to consider that this causes a problem for grammatical theory, that the data are actually inconsistent. That is, he finds a conflict in the fact that "cookie on plate" may be grammatical for the child at one stage but ungrammatical at another. This is only a problem if one feels that there must somehow be one grammar to cover all the data the child ever perceives. If, instead, two different grammars are constructed, one for the stage when the child says "cookie on plate" and another for the stage when the child says "the cookie is on the plate", then no problem arises. All the data that the child ever perceives cannot constitute a single corpus, so it is pointless to discuss the inconsistencies that arise in trying to construct a single grammar for that non-corpus.

A similar misconception as to the scope of the optimal grammar has given rise to the idea of the "blinkered learner", already mentioned in Chapter 5, by which Fodor and Smith (1978) mean that the child might fail to notice the optimal analysis. They argue that there are three possible ways to analyze "have

got" as it is used in certain dialects of English. They further argue that it is impossible to say which analysis is optimal by a simplicity-counting type of evaluation measure. In trying to evaluate these grammars for simplicity, they compare them with each other, neglecting the fact that a grammar must account for data and if it is simple but does not account for the facts, it is useless and, therefore, not optimal. They also seem to consider that the linguist would want to propose only one optimal grammar to cover the various dialects that are involved. However, each dialect provides a set of facts which must be described by a grammar and there is no reason to assume that a grammar proposed for one dialect can be optimal for another.

Fodor and Smith argue that although they cannot choose an optimal grammar for "have got", it is possible to show from acquisition which grammar is psychologically real. Therefore, they claim, the psychologically real grammar need not be the optimal one.

Their three analyses are: (a) taking "have got" as a perfective of "get" with nonperfective meaning; (b) considering "have got" to consist of a main verb "have" followed by a meaningless morpheme "got " transformationally introduced; (c) analysing "have got" as a main verb "to get" with meaningless "have" transformationally inserted. They then suggest that analysis (c) is psychologically real for the child on the basis of forms such as:

6.4. Tommy gots one
6.5. I don't got one
6.6. Do you got one?

(their 21-23)

They fail to argue that this must also be the optimal grammar for the child because they have considered optimality only in terms of a simplicity count, rating grammars against grammars without consideration of the data. But if they had adopted a stricter definition of optimal, such that the optimality of a grammar could only be considered with respect to the data that it must cover, and that data from several dialects of adults and children can hardly count as one corpus to be covered by one grammar, then this notion of "blinkered" learner could never have arisen. Evidence of forms like (6.4-6) suggests that analysis (c) is the child's grammar, both psychologically real and optimal in our terms. But the fact that this is the optimal grammar for the child says nothing about what is optimal for the adult dialects.

A similar "blinkered" learner argument is made by McCawley (1976, 1977) and it is open to the same objections. He argues that certain speakers, in their "mental" lexicons, may not relate words like "right/righteous", as an optimal grammar (i.e. Chomsky and Halle, 1968) would do, and that they may relate words like "moth/mother", which such an optimal grammar would not relate. He argues from this that the psychologically real lexicon is not the same as the optimal lexicon, that the child may fail to notice certain optimal analyses and create

non-optimal ones. However, the optimal grammar for a person who does not see the relationship between "right" and "righteous" and who does make a connection between "moth" and "mother" must capture these facts, just as the optimal grammar for people who do relate the former and do not relate the latter must capture those facts. These are, in fact, two different dialects and cannot be dealt with by the same optimal grammar.

Such misconceptions stem from the fact that many linguists act as if they all speak the same dialect and they try to determine the optimal grammar for that dialect. This is then referred to as the optimal grammar. But the aim of linguistic theory is not just to provide an optimal grammar for English but, rather, to determine universal constraints on any optimal grammar, the form of such grammars, etc. Obviously, in order to achieve such an aim, it is necessary to have detailed knowledge of at least one optimal grammar and, therefore, many people have concentrated on this, but nothing is lost by accepting that there will, in fact, be as many optimal grammars as there are perceptions of data.

The idea of a child failing to notice an optimal analysis, then, stems from a misunderstanding of what the optimal grammar would be in particular cases. If a grammar is optimal only with respect to a particular set of data as understood by a particular speaker, then, in the cases discussed above, the child cannot be said to have failed to notice the optimal solution; there is no way in which he could fail to perceive

it. This leaves the other possibility raised by Fodor and Smith, and also by Maratsos (1978), McCawley (1977), that the child is a "lazy" learner, that he can see an optimal analysis and yet fail to adopt it. (This will be discussed in more detail when restructuring is considered in Chapter 7). In such a case, the child would see that "right" and "righteous" are related and yet not capture that fact in his grammar, because he can manage quite well without doing so. It is not clear that such a position is logically possible: surely, if the child "knows" that certain forms are related, then this is part of his competence already, and therefore rather difficult to exclude.[4]

6.2. The optimal grammar is not necessarily "simple".

These problems have arisen through an idea that one grammar must somehow account for material which in fact will be dealt with by several different grammars. Another misconception in the literature is due to a mistaken identification of "optimal" with "simple". Baron (1973) and Traugott (1977) argue that the child's grammar at any one stage is more elaborate than his grammars of earlier stages, a fact which few would dispute, and go on to suggest that, in consequence, the later grammar cannot be considered to be optimal, since it is less simple than the earlier ones.

However, there is no reason to assume that an elaborate grammar is not also optimal. It is true that, in the context of language change, it has been argued that the child's

grammar is both less elaborate than the adult's and optimal with respect to it (Halle, 1962). However, such a comparison can be made only if the data being encompassed by the child's grammar and the adult's are the same. In the context of change between grammars in the same child, on the other hand, it is clearly nonsense to compare only the grammars involved. As already discussed, the child's grammar changes because his view of the data changes. The reason why a child's grammar at stage B is more elaborate than his grammar at stage A is that his perception of the data is also more elaborate. It is pointless to compare grammar A and grammar B in vacuo, to claim that A is simpler than B and, therefore, that B cannot be optimal. A is optimal with respect to data A and B is optimal with respect to data B. This is something which Baron and Traugott apparently fail to realize.

Comparing the child's grammar with another grammar of a previous stage or with the adult grammar irrespective of the data leads to a useless definition of optimality. Such a definition would imply that, when comparing a grammar which has only a regular past-tense formation rule (as children's grammars do at a certain stage) with one that has regular and irregular forms, the former grammar should be optimal. But such a grammar is not optimal if it is considered in the light of data which contains irregular forms. A child at stage A who has a rule of past-tense formation "Add/-ed/ to the base form of verbs" has noticed only this particular aspect of the data, namely the regularity of certain endings, and

has an appropriate rule in his grammar (which he over-generalizes). At stage B, he notices that the data contain forms which are exceptions to such a rule. He alters his grammar to encompass the exceptions. Now, grammar B is certainly more elaborate than grammar A, in that it contains exceptional statements and has to have more than one rule of past-tense formation, or more lexical entries, or whatever. Nevertheless, it is the optimal grammar for data B and there is no way that grammar A could be optimal for data B. If the data contains "marked" forms, and if the child has noticed them, the grammar must encompass them. A grammar with marked rules, required by the presence in the data of material which cannot be dealt with in a general way, may be less desirable in abstract terms, may be harder to learn, more subject to change, and yet it will be the optimal grammar for a particular situation, for the data that require those marked rules. Thus, evaluating a grammar in terms of simplicity alone will achieve little as far as assessing its optimality is concerned.

Although it is not possible to compare grammars in isolation in order to determine whether or not one is <u>optimal</u> with respect to the other, it is possible to assess grammars for markedness relative to one another, and it may be the assumption that "unmarked" means "optimal" which has led to some of the confusions I have been discussing. But an unmarked grammar, or a less-highly marked grammar, can only be considered optimal or non-optimal with respect to a particular set of data. Nevertheless, a conflict may sometimes arise for

the language learner between the requirement of providing an optimal grammar for the data and the difficulty of learning a highly-marked grammar. If the data are such that the optimal grammar will also be marked, the child may ignore such data and continue with a previous grammar which does not cover the data but which is less marked. In such cases "imperfect learning" (Kiparsky, 1965) results. That is, the child's output and the adult's will differ. Both grammars will be optimal, but for different sets of data. The grammars can be compared for markedness and, presumably, it is the pressures of markedness which have caused the child to change the data so that a marked grammar is no longer required. (See Chapter 8 for further discussion of markedness and acquisition).

6.3. The effect of the non-availability of certain grammatical principles.

So far, I have argued that one must look at a grammar in connection with the data that it is supposed to account for. An optimal grammar, then, is the best account of the relevant data. This presupposes that when the child's perception of the input changes, due to maturational factors, switching of attention to certain aspects of the data, etc., the child's grammar changes too. However, a change in the way the child sees the data may not be the only way that change is brought about in children's grammars: some new principle of U.G. may become available to the child which causes him to reconstruct his grammar in the light of it. The availability or non-

availability of such principles is also relevant when trying to consider whether or not a grammar can be said to be optimal.

Fischer (1976) tries to show that the child's grammar cannot always be considered optimal at each stage, even if one strictly considers each grammar only with respect to the relevant data. She finds that, in the acquisition of particle-movement and dative-movement,[5] children aged $3\frac{1}{2}$ to 4 years do not accept any sentence where a direct object pronoun is separated from its verb. They reject sentences with unstressed or stressed pronoun objects which are separated from the verb by a particle, although adults, speaking the dialect the children are exposed to, accept the separation of stressed object pronouns from the verb. The children show a strong preference for dative constructions of the form V NP PP. Five year-old children, on the other hand, accept sentences such as:

6.7. I saw a pretty flower and I gave my mommy it
6.8. I saw a pretty flower and I picked it up

That is, like the younger children and the adults, they do not allow a particle to intervene between the verb and the unstressed pronoun object, but, unlike these two groups, they do allow datives resulting in the separation of the direct object pronoun from the verb.

For various reasons, Fischer proposes that the optimal solution for the younger children would be to have an output condition:

6.9. A direct object pronoun may not be separated from the verb with which it is in construction.

(her OC 111)

This generalization can account for verb-particle constructions, with unstressed and stressed pronouns, and datives in their speech. For the older children, however, she proposes that the optimal solution would be to have a transformational condition on particle movement:

6.10. Particle movement is obligatory if the direct object is a pronoun.

(her TC 1)

and no restriction on dative-movement. She then goes on to reject her optimal solution for the $3\frac{1}{2}$ to 4 year olds and to propose, instead, that they also have the transformational condition (6.10) and another condition on dative-movement which they later drop.

I do not wish, at this point, to question Fischer's analysis but, rather, her arguments for rejecting her optimal solution. She is worried about how, under the proposals she makes, a child would change from the grammar of the $3\frac{1}{2}$ to 4 year old to that of the 5 year old. She says:

> "The transition is between two stages, the first of which supposedly mentions no transformations and the second of which mentions one. There is no mechanism that I know of in language diachrony which would account for such a jump."

(p. 87)

On the contrary, the formal nature of change between stages in diachrony is such that a great variety of change is possible, including the addition of transformations or transformational

conditions (Lightfoot, 1979) and, as will be argued in Chapter 7, the same is true of acquisition. As long as the two grammars she proposes fall within the range laid down by the theory of grammar, there can be no formal reason for rejecting the optimal solution (if it is, indeed, optimal). There is no reason to suppose that modifications to child grammars must be gradual, as Fischer herself notes elsewhere.

On the other hand, there may be developmental reasons to reject her optimal solution. Fischer suggests that a child of $3\frac{1}{2}$ to 4 may not yet be ready for output conditions. If these children do not have such conditions at all, if they are a part of U.G. which has not yet been activated for them, then, in the context of the linguistic mechanisms available to them, their supposedly non-optimal grammars, with the transformational condition (6.10), are in fact optimal. Two grammars which differ as to the availability of some principle of U.G. cannot be compared directly as far as optimality is concerned. The optimality of (6.9) becomes irrelevant if the children do not have output conditions at all.

Fischer's argument that the younger children's grammar is non-optimal involves the assumption that Output Condition (6.9) is indeed optimal and that the children do not have this condition. In fact, the analysis involving such a condition may never have been optimal in the first place, irrespective of whether children do or do not have output conditions in early stages. In Chapter 4, the learnability of dative con-

structions was discussed and the proposals of Oehrle (1976) and Baker (1978, 1979a) were preferred, that is the suggestion that dative-movement does not exist as a transformation. Instead, the structures V NP NP and V NP PP are both base-generated and verbs are subcategorized to take either or both. In that case, Fischer's attempt to provide an optimal solution for the younger children which covers particle-movement and dative-movement as transformations falls. The difference in the younger and older children's treatment of datives would be accounted for by the fact that the former have only acquired the base-generated V NP PP forms, whereas the latter had both, (though one might still have to explain why V NP PP is learned before V NP NP). Condition (6.10) could explain particle-movement in both cases.

Fischer herself comes close to this approach, though retaining dative-movement as a transformation, when she notes that four-year-olds do not have full mastery of dative-movement and, therefore, that it may be inappropriate to propose Output Condition (6.9) to account for a construction which they do not, in fact, use i.e. which does not need accounting for at this stage. Yet, having noticed that her proposal is not suitable for the four-year-olds' data, she still calls that analysis "optimal", falling into the trap of assuming that something which is maximally general or simple is optimal irrespective of the data under consideration.

6.4. Co-existing systems.

So far, I have considered apparent problems that arise

for the claim that the child learns the optimal grammar which turn out not to be problems at all if one confines oneself to considering a grammar only in relation to the data it is meant to account for. There are other problems that occur even within this more restricted definition of optimal grammar. For instance, if the child learns the optimal grammar for his perception of the data, the implication is that he will have only one grammar for any particular construction at any one time. There are a number of cases reported in the literature which provide apparent counter-examples, where the child seems to entertain more than one grammar at a time. However, on investigation, such cases are not particularly problematical.

Bever (1975) reports one such case, of a child who is aware of the difference between his own speech and that of his father:

> "Child: Mommy goed to the store.
> Father: Mommy goed to the store?
> Child: No, Daddy, I say it that way, not you!
> Father: Mommy wented to the store?
> Child: No!
> Father: Mommy went to the store.
> Child: That's right. Mommy wen....Mommy goed to the store."
>
> (p. 72)

What would be the optimal grammar for the child in this case? Does the fact that the child appears to understand a grammar which he does not himself use cause problems for the concept of optimal grammar? This seems simply to be a case where the child appreciates a different dialect from his own, without

having the grammar of that dialect to use productively. The child's grammar contains a rule of past-tense formation using the /-ed/ suffix only. This is optimal for his perception of the data, since, while he notices forms like "went" in his father's speech, he does not yet seem to appreciate that they have any relevance for him. (In the same way, a British speaker and a North American speaker can be aware that their uses of the verb "have" differ, without considering that that different usage has to be included in their own grammar).

Usually, the child will go on to see the relevance of adult forms and will adjust his own grammar accordingly. This may involve him in periods of transition between grammars, which have often been noted (e.g. Cazden, 1968; Brown, 1973) and which may cause problems for the concept of the optimal grammar, since, if the child learns the optimal grammar by definition, he should not have to spend time trying to decide what the optimal grammar is, change should not be gradual. However, it is not necessarily the case that grammar change itself is gradual; rather, these transition periods may affect the child's output only. For instance, at a certain stage, children produce no plural markers at all; at later stages they produce them consistently. In between, they sometimes use them and sometimes do not. If the in-between stage is considered as a case of optional or variable (Labov, 1969) rule operation, then the change is between two grammars, one of which has no rule of plural marking and the other of which does. This rule itself changes from being optional to being

obligatory.[6] The apparent gradualness of acquisition of the rule is due to its optional operation but the grammar change has been discrete.[7]

A greater problem is raised by the case discussed by Klima and Bellugi (1966). They propose an analysis of negation whereby at the earliest stage, their Stage 1, negation consists of a negator external to a nucleus (a very basic sentence). At a later stage, Stage 3, children's negation can be described by a complex set of rules which resemble those for adult negation. In between, at Stage 2, the systems of Stage 1 and Stage 3 appear to co-exist, in that forms representative of both stages are found. At Stage 2, what could be said to be the optimal grammar? There is no way in which Stage 3 can be analysed as an addition to Stage 1 with certain rules functioning optionally, as in the cases of morpheme acquisition mentioned above. It is completely different from Stage 1 and requires extensive reanalysis of that grammar. Yet, at Stage 2, where the child sometimes uses forms suggestive of the Stage 1 grammar and sometimes uses those suggestive of Stage 3, the child appears to have two co-existing grammars. What exactly is his perception of the data at this stage and what would the optimal grammar be? It may be that the child's perception of the data genuinely fluctuates in such cases, in which case there is no contradiction in having two grammars if they are each for a different perception of the data.

Alternatively, the fault may lie with their analysis of

negation, criticisms of which are given by Bloom (1970). In Bloom's analysis, there are no co-existing grammars. Negation is accounted for at all stages by the same rule:

6.11. S -------> Nom Neg $\left\{\begin{matrix}NP\\VP\end{matrix}\right\}$

In early stages, a "reduction" transformation operates:

6.12. X Neg Y ------> Neg Y

This rule accounts for the surface structure of negatives in early speech. In deep structure, negative is incorporated within the sentence, rather than being external to it, as it is for Klima and Bellugi. There is, therefore, no problem over the optimal grammar in Stage 2, where what one sees is the reduction transformation being applied optionally. By Stage 3, this transformation is lost altogether.[8]

Even if the analysis of Klima and Bellugi is correct, and even if the child cannot be said to have two different perceptions of the data at the same time, Stage 2 still need not cause problems. The fact that there are periods of transition does not mean that the two grammars must be thought of as operating productively during such periods. Instead, one grammar may be productive and the forms not covered by this grammar may exist as remnants or as idioms. There are two ways in which an overlap could occur. On the one hand, as argued by Bever, Carroll and Hurtig (1976):

> "Often the child appears to have mastered a new form as an idiom (i.e. not generated by its grammar) and

> only subsequently is the grammar extended to generate that form."
>
> (p. 151)

That is, in the case of Klima and Bellugi's Stage 2, the Stage 1 grammar would be productive and the Stage 3 forms would occur initially as idioms, memorized rather than produced by the grammar. Alternatively, the new grammar might come into operation and the old forms might remain as relics: Stage 2 might consist of a Stage 3 grammar as the productive source of the child's structures and the Stage 1 forms would occur through familiarity. Indeed, Stage 2 might well consist of both of these processes: first, a Stage 1 grammar with Stage 3 forms occurring as idioms; then, a Stage 3 grammar with Stage 1 forms as relics. In either case, there is only one productive, optimal, grammer. (This same explanation could also account for the cases of morpheme variation explained above by means of optional rules).

6.5. Methodological considerations.

I have argued so far that the term "optimal" has been too loosely defined, to mean either "simple" or else to apply to a grammar which bears no relation to the data under consideration. Instead, a grammar must be considered optimal with respect to a particular set of data and grammars can be compared for optimality only if the same data is under consideration and the same principles of U.G. are available. The amount of markedness in a grammar is irrelevant to its

optimality. Apparent problems for such a definition of optimal grammar, such as the question of over-lapping grammars, turn out to be slight.

It has been claimed that the young child's grammar cannot in any useful sense be called optimal (Traugott, 1977). On the contrary, the position adopted here, that the child's grammar is by definition optimal and that the optimal grammar is psychologically real, is a methodological assumption which minimises the possibilities of grammar construction open to the child, and ways for the linguist to describe the child's language. For example, suppose that the definition of "optimal" provided by the theory of grammar includes some limitation on abstractness in syntactic derivations (perhaps via a transparency principle (Lightfoot, 1979)). If there are two proposals for deriving adjectives in child language at the stage before the child understands or uses relative clauses, one of which says that the child base-generates adjectives in prenominal position, while the other claims that these are derived from relative clauses, the latter derivation can be ruled out as too abstract, since it proposes structures (relative clauses) for which there is no evidence at this stage. If the child does not learn the optimal grammar, then one is left with the possibility that he might adopt the more complex analysis but if he learns the optimal grammar, then this can be ruled out, and the number of grammars under consideration for child language can be considerably reduced.

It might be argued that no one would want to claim that the child's grammar is more complex than the optimal grammar, and, indeed, most discussion of this point has tried to show that the child's grammar is less so (Fodor and Smith, 1978; Maratsos, 1978; McCawley, 1977). That is, people have tried to demonstrate that an analysis which was suitable when the data was viewed in a relatively simple manner can also be carried on when the data must presumably be perceived in a more complex way. Nevertheless, the implication of any approach which separates the optimal grammar from the psychologically real grammar is that the child might choose an overly complex analysis in preference to the optimal one. If there could really be such a person as a "blinkered learner" who failed to see the optimal analysis, then it would be quite impossible to draw any conclusions about anyone's grammar, since one would be under no obligation to account for the data in the most adequate fashion. I maintain that it is by definition impossible for the child to be a "blinkered" or "lazy" learner (see also Chapter 7). The only way a child's grammar can be considered non-optimal is if the linguist has made a mistake. The child never makes a mistake, never chooses a non-optimal grammar, whether by default or laziness, because he does not have any choice in the matter. The principles of U.G. available to him and his perception of the data will result in his learning a grammar suitable for the data and that grammar will be optimal. The fact that linguists cannot always decide on

the optimal grammar in any particular circumstance does not mean that the child is also faced with such dilemmas.

FOOTNOTES TO CHAPTER 6

1. In the earliest discussions of optimal grammars, it was assumed that adults could not necessarily construct an optimal grammar, whereas children could do so (Halle, 1962; Kiparsky, 1965; King, 1969). Adults were thought to be able to modify existing grammars only by means of minor rule addition, whereas children were considered able to simplify by means of rule-loss, reordering and restructuring. The arguments for the adult's supposed inability to restructure were based on observations of an apparent critical period for language learning (Lenneberg, 1967; Penfield and Roberts, 1959). However, if an adult, in response to new data, does something to his grammar which results in appropriate output, it is very difficult to prove that he has not done this by means of an optimal grammar. (Of course, he may be incapable of responding to new data altogether but this shows a limitation on his perception and not on his grammar). Nothing of great importance for linguistic theory hinges on whether or not children alone are the locus of linguistic change. If adults are not capable of producing optimal grammars, then one must accept that the psychologically real grammar and the optimal grammar may not coincide in these cases. However, in the absence of convincing evidence to the contrary, I shall continue to assume that adult grammars, like children's are both optimal and psychologically real.

2. As was the case for Halle (1962), Kiparsky (1965) and King (1969) (see note 1). In their comparison of the child's grammar with the adult's, they adopt an idealization of learning as instantaneous and they assume that child and adult are dealing with the same data.

3. Where the adult does alter his output, for instance by simplifying his language to suit the early needs of the child and gradually using more complicated structures, a change in the data does take place, so that the child's attempt to construct a new grammar in these circumstances is to be expected.

4. It may nevertheless be the case that the adult who knows that certain forms are or are not related does not capture this fact. For example, McCawley (1976, 1977) reports that he often catches himself writing the word "hierarchy" as "higherarchy". This analysis of "hierarchy" as related to the word "high" is presumably part of his competence although he now knows that it is incorrect. But all this shows is that the adult is unable to alter his perception of the data, not that he cannot produce an optimal grammar. The persistence of the spelling error suggests that McCawley's knowledge that it is an error is not really part of his competence but something that he has consciously to remind himself of.

5. Fischer accepts a movement analysis for dative constructions. The alternative analysis, whereby both forms are base-generated, is discussed in Chapter 4, as is the problem of the learnability of a restriction against forms such (6.7).

6. However, as pointed out by Baker (1978, 1979a), Roeper (1978), and discussed in Chapter 5, there is a learnability problem with optional rules. If the child assumes that a rule is optional, he requires negative evidence to find out that it is in fact obligatory, so the change suggested here would require evidence that does not seem to be available. In that case, the explanation that follows, in terms of grammars where only one set of forms is productively generated by the grammar and the others exist as idioms or relics, is preferable.

7. As Brown (1973) notes, the impression of gradualness may also arise from considering several different changes together (in this case the acquisition of various different inflectional endings):

> "It looks as if performance improves gradually and rather slowly rather than abruptly. However, the percentages quoted sum across all morphemes and a gradual rise in these percentages is not inconsistent with an ordered series of abrupt changes in the many particular morphemes."
>
> (p. 256)

8. In spite of the comments of Bowerman (1973), Fodor, Bever and Garrett (1974) that Bloom's reduction transformations are due to performance factors and ought not to be included in a grammar of competence, an interaction of the rule of grammar (6.11) with the performance factors that lead to (6.12) would yield the same results.

CHAPTER 7

STAGES OF ACQUISITION

7.1. Two models of linguistic change.

In Chapter 6, I have argued that the child not only learns an optimal grammar eventually but that his grammar is optimal at every stage that he goes through. It is now time to consider such stages in more detail. In the past, acquisition was often assumed to be instantaneous, although this was always seen as an idealization (Chomsky, 1965; Chomsky and Halle, 1968; Chomsky, 1975). It was felt that more would be achieved by examining the grammar the child eventually acquires than by investigating the stages he goes through before attaining that grammar. Nevertheless, data from child language can add to the data base that the theory of grammar is responsible to; by determining the various grammars that children may construct, one is able to gain further insight into the form of possible grammars, constraints on such grammars, and so on.

What, then, is the relationship of each of these stages to the others? What happens when the child proceeds through a series of grammars? By a "stage", I do not here mean a description of the child's output in terms of the number of

words or the style he uses, as in "two-word stage", "telegraphic stage". Rather, the child will be considered to be in a different stage every time there is any change in his grammar, as evidenced by his comprehension or output. Change may be quite local: grammar X may differ from grammar Y with respect only to its treatment of one structure, but these will be treated as different stages here.

Two models of grammar change have often been discussed in the literature. On the one hand (e.g. King, 1969), language change is seen as a succession of discrete stages, the grammar of each stage relating directly to the relevant data rather than to any previous grammar:

7.1. Stage 1: Data --> LAD --> Grammar 1 --> Output 1

Stage 2: New data --> LAD --> Grammar 2 --> Output 2

The input to Grammar 2 is the output of Grammar 1, that is, data in the form of the child's utterances, and new data, i.e. the utterances of adults and peers. Grammar 1 does not itself serve as input to the subsequent grammar. While this model was originally proposed for change between generations, it is equally appropriate for grammar change within one individual.

On the other model (McCawley, 1968, 1977; Derwing, 1973), each stage is constructed on the basis of feedback from the pre-

ceding stage, feedback not just from the child's utterances but from his grammar. The input to any new grammar includes the previous grammar itself as well as new data:

7.2. Data -------> LAD -------> Grammar

As Chomsky (1975) points out, in the second case, one would expect to find "substantial differences in the result of language learning depending on such factors as order of presentation of data, time of presentation, and so on". Differences, that is, in speakers of the same dialect. But although children are not presented with the same data in the same order, the end results are substantially similar, suggesting that children's grammar construction does not depend on previous hypotheses but on their perception of the data. In the final stages of acquisition, all children come to see the data in much the same way, regardless of the history of their previous grammars.

Roeper (1978) proposes a combination of these two models. He notes that certain hypotheses, as to word-order for instance, are formed early on and remain unchanged through various grammars, suggesting that the child must have access to the previous grammar(s) which made those hypotheses. However, the difference between the two approaches may well be trivial at this point; if, as it appears, the child does not change his grammar in wholesale fashion but only deals with that section of the grammar where he has noticed a change in the data, then,

if there is nothing in the data to suggest a change in word-order, the child has no need to reconstruct word-order each time he makes some other change in his grammar. Thus, the fact that certain hypotheses appear irreversible, that the child appears to have access to previous grammars, means nothing more than that that section of a previous grammar is still appropriate for his current perception of the data, that it is still his current grammar.

7.2. Restructuring in acquisition.

If a grammar is optimal with respect to its data, then there can be no reason to change it unless there is a change in that data or in the child's perception of that data which renders that grammar insufficient. Thus, the major cause of change in children's grammars will be their changing view of the data, their realization that a particular grammar is too limited to cover the facts. Since, according to model (7.1), change involves starting afresh in grammar construction for any particular structure under consideration, without reference to the way this structure has been dealt with before, there are certain important implications for the kinds of change that can take place. The relationship of one grammar to another will be no more than that each is a possible grammar, as defined by the theory of grammar, though similarities in the data to be dealt with by each grammar will lead to similarities in the grammars. In theory, any kind of change allowed by the theory of grammar should also be possible between the

grammars of different stages of any individual child. The formal relationships between successive grammars fall into a wide range: changes such as rule addition, rule loss, re-ordering and restructuring, proposed by Kiparsky (1965), could, for syntax, involve any of the following: phrase-structure rules, lexical rules, strict sub-categorization requirements, transformational rules, interpretation rules, amongst others (Lightfoot, 1979). That is, formally, one can expect change in any part of the grammar and change can involve any grammatical operation.

In particular, the assumptions that the child learns the optimal grammar and that the optimal grammar is psychologically real lead to the position that grammar change in acquisition must sometimes involve restructuring. Restructuring is a possible form of linguistic change (Kiparsky, 1965). Formally, the relationship between the grammars of a single child and those of different generations is the same, and the same kinds of change can result. If the child constructs an optimal grammar for a particular set of data and then comes to see that data in a different light, he will have to construct a further, optimal, grammar to account for it. Between the first and second grammars, the relationship might well be one of restructuring, or of any other possible kind of change. Since model (7.2) is explicitly postulated to deny the role of restructuring and to show that children only add minor rules to existing grammars (McCawley, 1977) and since others have

also recently denied a role to restructuring in child language (Fodor and Smith, 1978; Maratsos, 1978), this matter needs to be considered in more detail.

Some of the objections to restructuring stem from an idea that it somehow involves a psychological cost. Thus, McCawley (1977):

> "Acquisition would be proceeding according to a principle of least effort (i.e. it is less of an effort to make a minor change in the grammar than a major change)."
>
> (p. 8)

Similarly, Fodor and Smith (1978) state:

> "Restructuring the grammar itself has an associated cost...there is a degree of inertia in the acquisition device that must be outweighed by the advantages to be gained by restructuring. The assumption that restructuring is automatic and cost free is really very implausible. For the child to recognize that a grammar is more optimal than the one he is currently operating with, he must retain his current grammar while formulating the alternative and testing it against the available language data to determine that it provides a better account of (the) facts."
>
> (p. 62)

However, it is just as much an unproven assumption that minor change (if that is what rule addition is) is less costly than major. On the model I have adopted, when the child constructs a new grammar, whether by restructuring or any other form of change, he considers the data alone. That is, Fodor and Smith's objection that he must retain his current grammar while formulating the alternative does not hold. As soon as the child realizes that his old grammar cannot account for some aspect of the data, he constructs a new grammar to deal with

that point; that new grammar is automatically optimal, so no comparison with the previous grammar is necessary. The "lazy" learner model, on the other hand, the one favoured by McCawley and by Fodor and Smith, does involve a cost: the child, apparently, actually sees the potentially optimal analysis but chooses to ignore it.[1] In that case, he must surely be comparing two grammars and the cost of rejecting an optimal analysis would seem the same as that of adopting it, in their terms. In fact, any model which proposes that grammars somehow add on to existing grammars involves more comparison of grammars than a model where the child does not have to think about what went before. Certainly, on McCawley's model, minor rule addition may be less costly than major, because the child actually has access to the previous grammar and so could make such comparisons. But if the child has no feedback at all from previous grammars, then it is completely pointless to talk about differences in effort between different kinds of change.

Evidence from second-language learning may also be relevant here. It is frequently noted with what ease young children can learn a second language, and with very little interference from their mother tongues (Dulay and Burt, 1974). This will surely involve the construction of grammars often radically different from the grammars of their first languages. Yet this seems to occasion little difficulty. Restructuring is nothing more than the construction of a grammar which differs

by more than minor rule addition or loss from a preceding grammar. If children have the flexibility to construct grammars for other languages, they surely have the psychological ability to restructure their own without one's having to think of this as a costly operation.

Maratsos (1978) argues against reanalysis on the grounds that it requires the child to substitute complex analyses for hypotheses that are already completely adequate. That is, in certain cases, he proposes that there is no need for grammar change at all:

> "The proposal of reanalysis encounters various theoretical difficulties. For the constructions discussed here, children apparently make initial hypotheses that are adequate to expression and comprehension - or... models are available which do not require extensive change of the stored representations. In all these cases, and others as well, extensive reanalysis for the purpose of attaining uniform grammatical representations would require that the child greatly complicate analyses of forms he had already captured with analyses closer to surface structure."
>
> (p. 257)

By arguing in this way, Maratsos makes it appear as if there is something specifically at fault with restructuring as such and that it should be ruled out on psychological grounds. But, as has already been discussed in Chapters 5 and 6, this stems from a mistaken dissociation of the optimal grammar from the psychologically real one. Maratsos objects to the kind of situation that would, supposedly, arise with prenominal adjectives, for example. In early versions of T.G.G., prenominal adjectives were derived from relative clauses (Smith,

1961). Since children produce prenominal adjectives long before relative clauses, any grammar not involving a great deal of abstractness will propose base-generated prenominal adjectives for the child. When the child produces relative clauses, this will involve the restructuring of prenominal adjectives to derive them from such clauses, if that is really the optimal grammar in those circumstances. But if the optimal grammar for the adult does not derive prenominal adjectives from relative clauses either, then no such restructuring will be necessary.

It now appears that the optimal grammar for adults will also have base-generated prenominal adjectives (Baker, 1975), and evidence from experiments on sentence perception (Fodor and Garrett, 1967) suggests the same, as does evidence from learnability (Baker, 1978, 1979a). Therefore, what is wrong in these cases is not the principle of reanalysis as such but the failure to identify the optimal grammar for the adult, which has led to the assumption that certain reanalyses would be required which do not, in fact, have to take place at all. If the optimal grammar for the adult involves analyses which are relatively close to surface structure, then many cases of restructuring will turn out not to be necessary, without one's having to rule out reanalysis in principle.

One is surely going to require the concept of reanalysis in many cases; for example, lexical restructuring appears necessary to explain the child's varying interpretations of

verbs like "promise" and "ask" and their interaction with the Minimal Distance Principle (M.D.P.) (Rosenbaum, 1965).[2] C. Chomsky (1969) reports that in the acquisition of these verbs initially children are consistent in treating the verbs as if they obey the M.D.P. Finally, they realize the exceptional nature of "promise" and of "ask" in certain situations. This in itself will involve modifications to the lexical entries of these verbs. In addition, there are in-between stages, where the children are not only unsure how to treat these verbs, but other verbs, such as "tell", which should consistently follow the M.D.P., are also affected, so that some temporary restructuring appears to occur in their lexical entries too. (See Chapter 8 for further discussion).

Other cases of lexical restructuring are discussed by Carey (1978). It would seem odd to retain reanalysis for the lexicon but to rule it out on psychological grounds from other parts of the grammar. Rather, one should accept reanalysis as a legitimate form of change, both formally and psychologically, but recognize that the need for it will be less now that less abstract analyses are proposed by the optimal grammar.

What, then, would constitute a reasonable proposal for restructuring in child language? In Chapter 4, the case of the acquisition of German word-order has been discussed. There, the position was taken that the principle of structure-preservation has nothing to say about the word-order likely to be adopted when the child is only dealing with root sentences

and is not aware of sentences containing embeddings. If the child is confronted with data that is largely in the form of SVO sentences, his initial hypothesis may well be that German word-order is SVO. As we have seen, the evidence from child language is somewhat ambiguous on this point. As soon as true embeddings are encountered, the structure-preserving hypothesis does become relevant and the child would have to restructure to underlying SOV, as a requirement of the optimal grammar which takes Emonds's principle into consideration. Maratsos (1978) argues that the evidence from early acquisition of German supports an analysis of free word order to begin with. In that case, since presumably Maratsos does not think that German children continue to exhibit free word-order, he must also accept restructuring from free to fixed order.

There have been other attempts to limit reanalysis in child language, particularly to confine the child to proceeding one step at a time in grammar change. Bever and Langendoen (1972), for example, note that children's grammars at any stage show only "minimal changes in highly articulated grammatical rules" and they propose the following principle to account for this observation:

> 7.3. The child's grammar at one stage is a minimal change from the grammar at the preceding stage.
>
> (their (85))

It is not clear what this is meant to be a principle of; it cannot be a principle of grammar, since the theory of grammar

does not impose any restriction on possible changes (Lightfoot, 1979), nor would it be possible on the kind of model I have adopted (7.1), where the child has no direct feedback from his previous grammars, where he is incapable of comparing grammars to see whether a change is minimal or not (if "minimal change" is even a definable term). A similar proposal is made by Hamburger and Wexler (1975), Wexler, Culicover and Hamburger (1975) who, in their discussion of learnability, propose that grammar change should take place only rule-by-rule, that is, that the learning procedure should hypothesize at most one new transformation at a time or reject at most one.

I do not deny that the child's grammar construction is constrained in some way; certainly, one does not find the extensive changes that might be expected if formal considerations alone applied. My quarrel is with the kind of restriction that has been proposed to account for this fact. Hamburger et al. get closer to the point when they suggest that the constraints may be due to restrictions on the learning procedure and not on change itself, though I suspect that their restriction is too strong (as they do too). If one places a limit on reanalysis itself, as Maratsos and McCawley do, then one still has to place some restriction on rule addition as well, to prevent massive additions to the grammar taking place at once. It does not seem to be the <u>type</u> of change which is at stake here at all.

There seem to be two ways in which one could reasonably

assume the child to be limited, and both concern the handling of data, rather than any formal constraints on the grammar. On the one hand, it seems quite plausible to suppose that there is some limitation on how much data the child can handle at once, that the child can concentrate on only a few structures at a time. Roeper (1978) proposes that change in children's grammars is local, that children focus on a limited range of data at a time.[3] If the child is only focussing on one thing at a time, then one would only expect grammar change in that one area. This will limit the range of changes that might occur but still allows for the possibility that there might be considerable restructuring within that limited area.

The other way in which there may be limitations on grammar construction is that all children may perceive the data in similar ways at similar stages. It has frequently been noted that children go through the same stages in the same order, though not necessarily at the same rate or age (C. Chomsky, 1969; Slobin, 1970; Brown, 1973). Grammars will be similar to one another if the data they account for is perceived as similar; this applies both across various children and between different grammars in the same child. Various factors may influence the child's perception of the data: maturational and cognitive factors, aspects of universal grammar, the ability to learn simple constructions before complex. These factors may cause the child to proceed through grammars which are relatively close to one another and cause different children

to go through the same sequences. Again, the restriction on grammar construction that this provides is not a limitation on any kind of grammar change as such but on the child's perception of the data.

It is not clear how a model like (7.2) can explain the similar stages and orders of acquisition that are often found across different children. According to this model, it appears to be a matter of chance whether people end up with the same grammars and, indeed, McCawley (1976) claims that they often do not. On this account, a child with grammar X preceded by grammar Y will construct a different grammar from the child whose preceding grammar is grammar A. In that case, it seems rather odd that so many children in fact go through similar stages. This is no problem if the child always goes back to the data and if different children's perceptions of the data are governed by similar factors so that they construct similar grammars. (Of course, McCawley could also claim that different children perceive the data in similar ways, so that they all add similar rules on to similar grammars, but this would rather destroy his point about differences in grammar construction).

Thus, attempts to limit restructuring in child grammars are misconceived in so far as they attack reanalysis itself as a possible form of change. Reanalysis can occur, just like any other form of change, subject to limitations on the child's data-handling abilities. A major methodological problem arises

if one excludes restructuring in principle; if the child only adds on to existing grammars, the linguist can never have any idea of what the child's or adult's grammar looks like at any stage without knowing his previous history, all the data he has encountered in the order he encountered it and the grammars he constructed for that data. The linguist is most unlikely to have access to this kind of information. If, on the other hand, reanalysis is possible, if we allow that the child's grammar must be optimal at each stage, no such problem arises; we are in a position to gain insight into the child's grammar of any particular stage and, hence, into the form of grammars in general.

7.3. The relevance of child grammars to adult ones.

I have argued so far that the child's grammar must only be considered in terms of the data which it accounts for. It has, therefore, no direct relationship with the adult grammar. Child grammars and adult grammars are possible grammars, as defined by the theory of grammar, and the only relationship they have to each other is that of being members of the class of possible grammars, since they do not account for the same range of data. However, where child and adult perceive the data alike, their grammars will be alike; children's later grammars may come to look very like adult grammars because of their attempts to deal with the same range of data as the adult, but their early grammars, where their views of the data may be very different, may bear little resemblance to the adult grammar.

In the field of acquisition, this independence of child and adult grammars is sometimes recognized:

> "The relevant standard of complexity is not the adult grammar but the child's own grammar."
>
> (Brown and Hanlon, 1970, p. 41)

More often, however, it is not. For example, McNeill (1966), Bloom (1970), and Brown (1973) express a preference for children's grammars that can be related to adult grammars. Thus, Brown remarks, with reference to pivot-open grammars:

> "The pivot-open distinction...seems to make no particular linguistic sense. If these are primal classes why are they so? They have nothing obvious to do with adult standard languages."
>
> (1973, p. 95)

And Bloom (1970) justifies her analysis of negation in child speech in the following way:

> "The specification of the negative particle outside the sentence is inconsistent with sentence negation in the adult model of English, where negation is an inherent semantic fact of English sentences.
> .....
> Bellugi (1967) suggested that, in Period A, negation outside the sentence represented a syntactic structure that has no relation to the adult model of the language - 'a primitive abstraction which later drops out'. In contrast, the development of Kathryn, Eric and Gia suggested that the earliest system of negation was more similar to the adult model than it was different."
>
> (p. 163)

Brown and Bloom, who both recognize the need to describe child grammars in their own right and see acquisition as a progression through different grammars, nevertheless use the closeness of an analysis to the adult grammar as an argument in its favour, a

step which is surely incompatible with the view that the child's grammar must be considered on its own terms. Bellugi (1967) is more consistent in allowing the possibility that the child and adult grammars are radically different. This is not to say that I prefer Bellugi's analysis over Bloom's. But the arguments which Bloom uses to reinforce her position, which she reaches on the basis of the child data alone, are based on a misconception. The fact that the child grammar may closely resemble the adult's is not because this is in itself a virtue but because both must be possible grammars and the data they are covering happens to be seen in a similar light. Therefore, attempts to discredit pivot-open grammars, or sentence external negation, or whatever, must not be based on statements that they are not like adult grammars but must show that they do not correctly account for the data or that they are not possible grammars. It may be that this is what Brown and Bloom are trying to capture but the comparison with adult grammars is misleading.

Part of the confusion arises from a feeling that the child has as his goal or aim the achievement of a "terminal state" (Klima and Bellugi, 1966), namely the adult grammar. Linguistic development is seen as increasing mastery of the adult grammar (Fodor, Bever and Garrett, 1974). This gives one an image of the child proceeding through a sequence of grammars towards the ultimate grammar; one cannot speak of the grammar of Old English as "aiming" at the grammar of Modern

English, and it is similarly inappropriate to think of the child's activity in this way, as if he is teleologically directed towards the adult grammar. This would imply a "racial" memory, some means for the child to know what he is aiming at. But how could he possibly know this? Rather, the child's aim must be expressed in terms of the data. He seeks to deal with the data that he hears and to produce language of his own. As he deals with more and more data, and matures so that he is capable of understanding and producing more, his grammar approximates more closely to the adult one, because of his increasingly adult-like perception of the data. As Andersen (1973) puts it:

> "The learner who formulates a grammar on the basis of the verbal input of his models has as his goal a grammar that will produce that output. Whether his grammar actually is identical to or different from that (those) of his models has no practical relevance in the speech community, which can only be concerned with observable usage."
>
> (p. 789)

This makes it quite clear that the child can only be said to aim to produce a grammar to deal with the facts. He does not proceed through a series of sub-grammars of the adult grammar but through a series of possible grammars attempting to encompass the data. In so far as the data is the same, the child and adult grammars will be similar,[4] and this follows from the fact that both must be possible grammars and optimal. But where the child's perception of the data differs from the adult's, there is no reason to suppose that the child grammar will

necessarily closely resemble the adult's and no merit to arguments for particular grammars which rely on this claim.

The assumption that language learning proceeds in stages and that these stages must be considered independently of one another, that the child cannot be seen as aiming at the adult grammar, has certain consequences as far as the use of acquisition data is concerned. If learning were instantaneous, one would be able to use evidence from child language as direct support for proposals concerning adult grammars. But since this is not the case, one must accept that rules which are present in the child's grammar are not necessarily present in the adult's and vice versa. Even the claim of the irreversibility of certain hypotheses (Roeper, 1978), that in certain areas the early grammar remains adequate and therefore does not get altered, does not allow direct comparison with adult grammars, since one cannot necessarily tell which parts of a grammar are irreversible. Referring back to model (7.1), the rules of grammar 1 cannot be used to account for the output of grammar 2 and the rules of grammar 2 cannot describe the output of grammar 1, despite attempts along these lines for diachronic change by Traugott (1969), Keyser (1974) and others.

Nevertheless, a number of people invoke acquisition data to support synchronic analyses of adult grammars directly. This is granting such data a special status which it is not at all obvious that it has. For example, Solan (1978) argues

that data from language acquisition support a tough-movement analysis (Postal, 1971) for sentences like:

7.4. John is easy to please.

On the movement account, such sentences have as their source the following:

7.5. It is easy to please John

On the other hand, sentences like (7.6) are derived via complement object deletion from (7.7):

7.6. Mary is pretty to look at
7.7. Mary is pretty to look at Mary

An alternative analysis is to derive both (7.4) and (7.6) in parallel fashion by complement object deletion; that is, there is no movement involved in either case (Lasnik and Fiengo, 1974).[5]

The acquisition data apparently show that children acquire the "easy" sentences before the "pretty" ones. Solan claims that this must lead one to select the movement analysis for the adult, that psychological reality is somehow at stake if one does not do so. However, I have argued that the grammar which is psychologically real for the acquirer at one stage will not necessarily remain unchanged. Solan has not, in fact, shown that a deletion analysis can be ruled out as psychologically real for the adult, although he has shown that it is insufficient to account for a certain stage in the acquisition process. Since reanalysis is logically possible, it might be that the child originally acquires the "easy" sentences as movement rules and later restructures them along the lines of the "pretty" deletions. The interpretation of

the acquisition data from a particular stage as being crucial is too strong, since one does not know in this case, or any other, whether the stage under consideration is the final stage the child will attain for the rule in question. I make no decision here as to which analysis is the correct one. I merely wish to emphasize that one must be cautious of the strength one affords to acquisition data to support arguments in other areas. Such data can be suggestive but not conclusive. Solan's findings suggest a difference in the two sentence types for the child, though whether this is a difference between movement and deletion as he claims is open to question. If this difference must be captured in the child's grammar of a particular stage, it is possible, but not absolutely necessary, that such a difference exists in the adult grammar. Other data must be brought to bear on the issue.

The particular way in which acquisition data can be suggestive, then, is to provide insights into possible grammars. For example, child language may offer evidence on the structure of the lexicon. Jackendoff (1975a) proposes that inflected words are entered in full in the lexicon, and are related by lexical redundancy rules, like derivationally related words. Children appear to include inflections in their lexical representations; this is suggested by the findings of Berko (1958) who shows that children can productively and correctly pluralize nonsense words requiring /s/ or /z/ but that they cannot add /əz/ where it is required by a nonsense word, although they can

correctly produce real words with plurals forms ending in /əz/, such as "glasses". In such cases, it seems that words like "glasses" must be entered in full in the child's lexicon, rather than their having a productive rule to form the /ə z/ plural. If children have lexical entries fully specified as to inflections, it is possible that adults do too. But decisions as to the correct grammar for the adult will depend on a number of factors and child language cannot be considered as providing conclusive evidence.

Nevertheless, there are times when child grammars may offer direct evidence of what adult grammars must look like. In Chapter 4, the question of learnability has been discussed. Baker (1978, 1979a) proposes to restrict the class of possible grammars by requiring that analyses can only be proposed on the basis of positive evidence, to rule out the kind of overgeneralization that would require negative evidence in order to be corrected. For example, if there is a transformation of "to be" deletion, producing (7.8.a) from (7.8.b):

7.8.a. John seems sleepy
b. John seems to be sleepy

then (7.9.a) ought to be possible from (7.9.b) but it is not:

7.9.a. *John happens sleepy
b. John happens to be sleepy

If the child had such a rule, one would expect him to generate forms like (7.9.a) and then be corrected, thus learning the exceptions to the rule. Children neither seem to get corrected for syntax nor to produce such errors in the first place, and Baker therefore proposes to exclude such grammars altogether.

If they are excluded in principle, then there is no way the adult could later acquire them (even if negative evidence is available at later stages of learning) and the fact that adults also do not come up with aberrant forms like (7.9.a) confirms this. Thus, considerations of learnability provide constraints on all possible grammars.

Some of the proposals for reanalysis which Maratsos (1978), Fodor and Smith (1978) object to will be rendered unnecessary if grammar construction is limited to consideration of positive data only; the more abstract grammars which allow overgeneralization will automatically be ruled out. For example, Baker (1978) shows that learnability criteria must exclude a relative clause source for prenominal adjectives. If adjectives are derived from relatives, then (7.10.a) ought to be possible from (7.10.b):

7.10.a. *the awake man
b. the man who is awake

The fact that children do not produce such forms and so could not be corrected, and that in any case they are rarely corrected or ignore corrections, means that this analysis must be excluded. This leaves us in the position that a relative clause source is unacceptable even when the child has relative clauses in his grammar and we have already excluded it from grammars of stages before he has them. Therefore, the child and adult grammars will be alike in this respect. But this does not mean that child and adult grammars will always be alike, since there are other grammars that children construct

on the basis of positive data which are clearly different from adult grammars.

7.4. The relevance of acquisitional data to diachronic change.

So far, I have discussed cases where such a close relationship between child and adult grammars is presupposed that child data are considered direct evidence for adult grammars. Another area where a close relationship is often unjustifiably assumed is between child language and diachronic change. Child language is felt to offer special insights into linguistic change, over and above other kinds of evidence. But, again, if the grammar for any stage for a child relates only to the specific data that the child is dealing with, there is no reason why change in this grammar should relate directly to diachronic change, where the data concerned are likely to be quite different.

Two kinds of claim have been made regarding child language and linguistic change. One proposal is that the state of a child's grammar at a particular stage can be used to predict the direction of future change in the speech community (Fischer, 1976). Another is that acquisition in some way "recapitulates linguistic change" (Baron, 1973; Traugott, 1977).

Both of these claims are too strong. In discussing the acquisition of particle movement and dative movement, Fischer notes that children go through a stage (dialect C) which is different from the dialect of their parents (dialect A), though it is an acceptable dialect in other parts of the U.S.A.

She suggests that the fact that children have this dialect for a time is strong evidence that eventually dialect C will predominate over dialect A. However, the children also go through other grammars which Fischer has not picked on as likely directions of change. In order to decide which of the many grammars that children go through is the one likely to be adopted (if any are; after all, many child grammars are never adopted into the speech community as a whole), one has to take into consideration facts about the formal nature of the adult grammar, whether it is ripe for change by being highly marked in some area. The interest in acquisition data arises because only children are considered able to initiate change[6] but it is factors beyond their control which will dictate whether adults will accept the new grammar. Given a range of child grammars attempting to cover a particular structure and a description of the adult grammar of that structure, and no other information, one would not be able to show which of the children's grammars, if any, would replace the adult one, though one might be able to predict from markedness in the adult grammar that some kind of change would be likely (i.e. the fact of the need for change but not the direction the change would take). In fact, Fischer backs up her claim with synchronic evidence, and this is crucial. The child data alone is not sufficient to predict the direction of change, though in conjunction with other data it may perhaps help to do so.

Baron (1973) is interested in child language in order to look back at changes that have occurred, rather than forward to ones that might happen. She investigates the ontogenetic development of the causatives "have" and "get" and compares this with their diachronic development:

> "In traditional historical accounts, the verbs 'have' and 'get' are said to develop causative interpretations from earlier non-causatives. The ontogenetic data analyzed support this general hypothesis for both verbs. However, the acquisition data argue that the usual descriptions given for how the causative interpretations arose may not accurately reflect history."
>
> (p. 84)

The fact that the ontogenetic data show something is entirely irrelevant to what happened historically. The data for children learning "have" and "get" today is quite different from what it was in the past, and unless children have knowledge of the grammars of their ancestors and of the data of past times, which clearly they do not, there is no way in which the child language can either support the derivation of "have" and "get" from non-causatives historically or suggest that usual diachronic accounts are wrong. Grammars of past stages of English must relate to those stages; grammars in acquisition must relate to the child's perception of the data. Unless the data happen to coincide, there will be no resemblance between the grammars at all, and none in the sequences of grammars that result.

I do not deny that there are any similarities between language change as it occurs diachronically and in acquisition. But this is not surprising. Both acquisition and

diachrony involve change between grammars; the grammars in both cases fall into the class of possible grammars. Grammar construction is constrained by U.G. and directed at particular sets of data. Where there are similarities in the data, it is not surprising that the same kinds of grammar should be proposed to account for them or that the same kinds of change can occur between grammars. If the class of possible grammars is reasonably restricted, it is inevitable that the same kinds of change will be seen again and again. What is surprising is that so much should be read into this similarity between language change historically and in acquisition, to the extent of believing that one can explain the other.

In conclusion, it appears that a study of the stages of language acquisition is useful to the linguist in that it can provide further insights into what a possible grammar is; it provides a means of increasing the data base to which the theory of grammar is responsible. But, given that child grammars, like any other grammars, relate directly only to data and that the grammar of one stage is independent of another, facts about child language cannot be directly relevant to descriptions of other grammars, whether these are the grammars of children at other stages, or those of adults, or ones describing earlier stages of a language. Rather, the data bear directly on theories of grammar.

FOOTNOTES TO CHAPTER 7

1. See also Chapter 6 for discussion on whether this position is logically tenable.

2. In most cases where a verb is followed by a verb in infinitive form, the understood subject of the latter is the NP most closely preceding it, for example:

i. John wanted to leave
ii. John told Bill to leave

where the underlined NPs are the understood subjects of "to leave". "Promise" and "ask", in one of its readings, do not obey this minimal distance principle:

iii. John promised Bill to leave
iv. John asked Bill what to do

3. An implication of this is that when one says that the child acquires the optimal grammar, it will be locally optimal for the structure in question and not globally optimal. That is, when one particular structure is being dealt with, the child does not look at the rest of the grammar as well.

4. Indeed, the similarity in grammars that results from attempts to encompass the same data might be sufficient to account for "mutual comprehensibility" between different generations, without the need for a functional limitation on change, such as the one proposed by Halle (1962).

5. Baker (1979a) also argues against a movement analysis for "easy"-type sentences and in favour of base-generated adjective phrases.

6. I have already expressed my doubts about this, earlier

in this chapter and in Chapter 6. I suspect that adults are far more able to change their grammars than is normally allowed them.

CHAPTER 8

ACQUISITION AND MARKEDNESS

8.1. Evaluation measures.

In the literature, the assessment of grammars has been seen in two different ways, though both are considered to be part of an evaluation metric which allows a choice between "competing" grammars. On the one hand, one must account for the learning of the best possible grammar for any particular set of data, as opposed to a less satisfactory grammar for the same data. On the other hand, the theory must account for the spread of possible grammars, irrespective of the data that they describe. The kinds of comparison and evaluation that are involved differ, since only in the former case are grammars which cover the same data compared to one another.

The choice between several grammars for the same data is made by a "simplicity" metric. The problem of evaluating grammars is frequently expressed as though the child is faced with the need to choose between grammars, but this is the linguist's problem rather than the child's. While the linguist may have difficulty in formalizing a notion of simplicity and then assessing grammars relative to that definition, the child is not a "little linguist"; he does not face the same

problem because he already has the properties that the linguist is trying to discover. I have argued in Chapter 6 that the child must select the optimal grammar for the data in question, that his grammar is optimal by definition. This means that any "simplicity" metric, or whatever it is that decides optimality, is such that the child can only come up with one grammar in the first place. The definition of optimality, then, is built in; it is part of the a priori knowledge that children bring to the acquisition task, so that they never have to construct a grammar and then wonder whether or not it is optimal; they cannot help producing the best one.

The need for such a simplicity metric to choose between different grammars for the same primary data is discussed by Chomsky (1965), when he considers the logical possibility that the class of grammars could be sufficiently restricted and the primary data sufficiently rich to allow only one grammar compatible with the data (see Chapter 1). In such a case, no evaluation procedure would be necessary to single out a grammar, since only one grammar would be produced in the first place. Rejecting this possibility, he goes on to say:

> "It is rather difficult to imagine how in detail this logical possibility might be realized, and all concrete attempts to formulate an empirically adequate linguistic theory certainly leave ample room for mutually inconsistent grammars, all compatible with primary data of any conceivable sort. All such theories therefore require supplementation by an evaluation measure if language acquisition is to be accounted for and selection of specific grammars is to be justified."
>
> (p. 37)

However, in many cases, a more restrictive theory of grammar has ruled out some of the grammars that apparently were competing for consideration, i.e. has removed the need for a simplicity metric in certain cases. For example, we do not need a simplicity metric to tell us that that Det-Nom account of One-substitution is better than the NP-S analysis (see Chapter 3) since these grammars turn out to be non-equivalent once the non-primary data are taken into consideration and only the Det-Nom account is adequate to describe our competence.

Similarly, if one considers two grammars for pre-nominal adjectives in English, one might reject the grammar which derives these adjectives from relative clauses on the grounds that it is not as "simple" as the grammar which base-generates them, since it is more abstract, involving more operations and a greater distance between deep and surface structure, and requiring that certain adjectives are marked as exceptions to the rule of relative clause reduction/modifier shift. However, this is not the only means to exclude the relative clause analysis, since, as we have seen, it can be excluded on the grounds of non-learnability (see Chapter 4). Again, with two grammars for the same primary data, one does not have to rely on a simplicity metric to rule out one of them. Though some need for a simplicity-type of evaluation may remain, this becomes less important as a restrictive theory of grammar is developed.

The other kind of evaluation of grammars is that provided

by the concept of markedness. Here, rather than trying to eliminate certain grammars which account for the primary data but in a less than simple fashion, one is trying to establish the limits of possible grammars in general and the accessibility of grammars within that range. When a grammar is considered as more or less marked than another, this comparison is made in isolation from the data. For example, where grammatical constraints are relative, a spread of grammars may occur, with each grammar describing a different set of data: $\bar{X}$ proposals do not constrain languages absolutely; some languages, such as Dutch, do not observe the parallel ordering of specifiers and major categories across all categories. Thus, a theory of grammar must account for the occurrence both of grammars which observe $\bar{X}$ constraints and of those which do not. Such grammars may be evaluated with respect to each other, irrespective of the fact that both adequately account for the data they encompass. This evaluation is made by the theory of markedness, which claims that the more marked grammars are less accessible to the child, or harder to acquire.

The requirements of the data dictate whether or not a marked grammar is needed and a marked grammar will also be optimal, in that it is the best account of the (complex) data that it has to describe. (As markedness is a form of complexity, this creates confusion if the assessment of optimality is thought of as being carried out by a "simplicity" metric, since a grammar may be both marked and optimal, or, apparently,

complex and simple at the same time). Where one grammar is marked and another less so, the data under consideration will be different. Consider, for example, the case of the filters proposed by Chomsky and Lasnik (1977). Their filter (20), here (8.1):

8.1. * $[_{NP}$ NP tense VP$]$

is proposed in order to rule out subject relative pronoun deletion in Modern English. The filter accounts for the non-occurrence of forms like:

8.2.a. *I saw a man has a bald head
b. *The man has a bald head saw me

In Middle English, where forms such as (8.2.a and b) were permitted, the filter did not operate. In order to overcome learnability problems, the filters operate in the absence of positive evidence to the contrary and this constitutes the "unmarked" case. One can compare a grammar where the filter operates with one where it does not and assess the former as less marked than the latter. In other words, one can compare grammars even though they deal with different data. Thus, the grammar of Modern English is less marked than the grammar of Middle English in this particular respect.

The evaluation of grammars in terms of markedness, then, is rather different from evaluation in terms of simplicity, in that it is not concerned with choosing between a number of possible grammars for the same primary data. If the primary data require a marked grammar, such a grammar will be

constructed. However, the claim of markedness is that this is difficult for the child, that the necessary triggering experience is more elaborate, and that marked alternatives will only be considered after less marked ones have been tried and have failed.

8.2. Some predictions from markedness for acquisition.

Claims about markedness, then, are claims about acquisition in real time, about the likely course of acquisition, with the more accessible grammars being attempted first. It is often noted that children acquire complex structures after simple ones and markedness may provide a more adequate definition of linguistic complexity than previous attempts (for example the D.T.C.), such that a correlation may be found between acquisition orders and the predictions of markedness.

Markedness can also make claims in other areas, for example on the likelihood of linguistic change: a highly marked grammar is liable to change (Lightfoot, 1979). For some reason the predictions of markedness with respect to acquisition and change have sometimes been interpreted as being paradoxical. Baron (1973) and Traugott (1973) find it inconsistent that children's grammars should proceed from unmarked to marked, whereas diachronically the change is in the other direction.[1] However, I see no dilemma here: children's grammars become more marked because they deal with less complex aspects of the data first. If they are learning

a language which requires a high degree of markedness in certain areas, they will eventually come up with the necessary marked structures, after proceeding through less marked alternatives which will not be adequate for the adult's perception of the data. On the other hand, there is pressure on highly marked grammars to change - for the data itself to change so that the high level of markedness is no longer necessary, so that the grammar is more easily attainable. It is possible, but by no means necessary, that the late acquisition of marked forms contributes to their loss historically; if the child initiates change, he allows a less marked grammar to predominate, ignoring the actual data (on the lines of "imperfect learning" (Kiparsky, 1965)). If adults can initiate change, then they may alter the data so that the child does not have the problematical structures in his input and does not have to account for them in his grammar. Either way, there is no paradox but an interaction which follows logically from predictions of markedness made in the areas of acquisition and change.[2]

Any predictions that can be made about the course of acquisition in real time will depend not only on one's theory of grammar but also on one's theories of speech production and processing and one's theory of real-time acquisition. If the grammar is psychologically real, it must be part of such theories, though the relationship does not have to be as close as the one-to-one relationship between grammatical transforma-

tions and parsing operations that was assumed by the D.T.C. Recently, some interesting proposals have been made for a theory of parsing which retains the grammar as a component of the processing mechanisms (Fodor, 1978; Frazier, 1979) and for a theory of acquisition where the grammar inter-relates with other factors (Roeper, 1978). With the grammar as a component of processing/acquisition models (as it must be on any view of "psychological reality"), it is reasonable to expect some kind of link between grammatical complexity, as defined here by markedness, and processing or acquisitional complexity.

This is not to say that all sources of difficulty in acquisition will arise from the grammar. The grammar can only be expected to make predictions concerning grammatical complexity; other areas will be outside its scope. Thus, it is not the place of the grammar to account for perceptual difficulties, although these may affect the parser and indirectly the grammar in the long run. In some cases, linguistic complexity may under-determine parsing problems; sentences which do not seem at all complex as described in the grammar may give rise to temporary problems of ambiguity for a left-to-right parser (Fodor, 1978; Frazier, 1979). In other cases, sentences which appear grammatically complex may not create parsing problems at all due to the operation of particular parsing strategies (Frazier, 1979). Thus, it is possible that the prediction that markedness will cause acquisition difficulties

will turn out to be too simple, or that particular parsing strategies will vary the effects of markedness. Nevertheless, in the absence of any alternative, I shall take it as a working hypothesis that the theory of markedness does say something about acquisition in real time. In this way, acquisition data provide a further area on which the predictions of markedness can be tested: if marked rules are indeed acquired late, this is confirming evidence for their marked status, whereas if no problems are involved in their acquisition, this will indicate the need to reconsider the proposals.

Recent proposals for markedness in syntax arise from the concept of "core" grammar (Chomsky, 1977; Chomsky and Lasnik, 1977; Chomsky, 1978, 1979). A language will consist of core rules, which are maximally general, and which are unmarked (or least marked; see Kean (1979) for a discussion of the distinction and the implications for core grammar), and peripheral rules, which do not fit into the generalisations of the core grammar and which are marked (see Chapter 2 for details). A number of possible predictions for acquisition follow from these proposals: if core rules are less marked than non-core rules, they should be less complex. Children working within the core framework and hypothesizing more accessible grammars before less accessible ones might first of all treat all marked structures as if they could be dealt with by the core grammar, only later realising that the generalisations of the core grammar do not apply, or they might

avoid them altogether, not being able to interpret the data within any framework as yet.

Within the core grammar itself, there are possible predictions for acquisition. Chomsky (1977) proposes that English has only two core transformational rules. One of them, Wh-movement, accounts for a number of structures which were previously dealt with by separate rules of grammar. Comparative deletion, topicalization, clefting and tough-movement are all considered to be instances of Wh-movement, as well as the more obvious cases where a Wh-element is explicitly present, as in direct and indirect questions and relative clauses. This is proposed for adult grammars and may or may not be true of child grammars, since one does not know that the child necessarily has the same core grammar as the adult.

Since children appear to have Wh-movement relatively early in their speech, as shown in direct questions, for example, it is reasonable to assume that this rule is psychologically real for them. If they see it not just as a rule to deal with questions but as a core rule like the adult, then for them also all the constructions covered by this core rule should be learned together, all other things being equal. A number of factors make it difficult to compare the acquisition of the various sentence types subsumed under the one rule. Children do not learn all the Wh-constructions together: direct questions, for example, are produced before relative clauses

or indirect questions. However, there are obvious reasons for this. Simple sentences are produced before ones containing embeddings, so one would expect direct questions to emerge earlier than relative clauses or indirect questions. Different interpretation rules may be involved for the various sentence types, so that a difference in acquisition might be due to factors to do with interpretation rather than syntax. Perceptual factors might also contribute difficulties to one structure rather than another.

Nevertheless, it is possible to test aspects of the claim that structures accounted for by the core rule of Wh-movement are acquired together. For instance, once a child can embed sentences, he ought to be able to produce indirect questions. Since he already has Wh-movement as a rule applying in simple sentences, and the interpretation rules for direct and indirect questions are the same, embedded questions should occur at the same time as other embedded structures. Similarly, once a child can produce complex NPs, he ought to be able to produce relative clauses, since these are a particular kind of complex NP involving Wh-movement.

There is some evidence from acquisition that bears out such predictions. Sheldon (1972) notes that full relatives enter children's speech at the same time as other sentential embeddings, as do Klima and Bellugi (1966). Limber (1973) reports that wh-adverbials and indirect questions emerge together.

Since I do not claim that the child and adult grammar need be alike, indeed that they can only be so if they see the data in the same way, it is certainly possible that the child constructs one rule to deal with Wh-questions and other rules to cover the other structures, not seeing them as potential cases of Wh-movement. The child might notice Wh-questions in the data and arrive at the rule of Wh-movement to account for them. Later, he might notice, say, topicalization and construct another rule to deal with it, not realising at that point that the same rule will do for both.

There is some evidence from Gruber (1967) which can be interpreted to show that this is not the case, that the child does have Wh-movement as a general core rule which can apply outside an overtly Wh-structure. Gruber reports that the child he studied spontaneously produced topicalized structures although there were none in his input data.[3] He also produced direct questions beginning with Wh- words at the same time. This suggests that once the child has acquired a core rule, he is capable of extending its use beyond the immediate data, to produce other structures which the rule allows but for which there is no apparent primary evidence.

While it may be difficult to reach any definite conclusions on the effect of core rules on acquisition, due to the variety of other factors that must interact with such proposals, it is easier to assess predictions about the acquisition of non-core, or marked rules, in the grammar. Chomsky and Lasnik (1977)

remark that one might think of marked rules as "the syntactic analogue of irregular verbs". As such, one would expect problems with learning them.[4] One of the ways in which marked rules are irregular is that they are able to bypass the general conditions on rules such as S.S.C. and P.I.C. (see Chapter 2). Rules can be "immunized" against conditions by specifying constants instead of variables in their conditioning environments. Since the least-marked, core rules do not "violate" conditions, one would expect the child to assume that all rules obey the conditions before working out that specific formulations can avoid them. When attempting to learn a marked rule, the child might first try to make it conform to some more general, less marked alternative (even though this would not be adequate for the data requiring the marked rule) or he might avoid the structure altogether.

Thus, the general rule of L-tous (2.14), here repeated as (8.3):

8.3. W V Q Z ----→ W Q V Z

should be learned before the marked rule (2.17), here (8.4):

8.4. vbl V* Q que ∝ Pro vbl

The learning of (8.3) should involve no violations of conditions such as S.S.C. or P.I.C., since the rule is unmarked and subject to those conditions. The learning of (8.4), on the other hand, involves the realization that one is dealing with a special case which does not fall under these conditions. The child will have to learn to formulate the rule in terms of

constants in the conditioning environment, meeting that special case. The rule is marked because of its lack of generality. Forms such as (2.16), here (8.5), should, therefore, be acquired late:

8.5.a. Il faut toutes qu'elles s'en aillent
b. Il faut tous qu'on se tire

This is a testable prediction.

Chomsky and Lasnik (1977) propose a rule of reflexive-deletion and note that the rule is outside the core grammar, as it involves a "high degree of uncertainty and variation". They argue that forms such as:

8.6. We want to win

are derived from:

8.7. We want for ourselves to win

The rule of reflexive-deletion operates in the context:

8.8. $\langle_{\overline{S}}$ for _____ to .....$\rangle$

This results in the following:

8.9. We want for to win

which is, in turn, excluded by the "for-to" filter for those dialects, such as Standard English, which do not permit such forms. Now, presumably, since this rule is not a core rule, it ought to be acquired late. Yet structures such as:

8.10. I want to go

do not seem to be particularly difficult for the child and emerge before forms with specified subjects (Maratsos, 1978), such as:

8.11. I want Fred to go

It can be argued that forms such as (8.10) are originally acquired as VP complements, in which case the reflexive-deletion rule will not be relevant at this stage. To maintain the reflexive-deletion account, the child's grammar must restructure when the child becomes aware of sentential complements. In that case, one would expect forms such as:

8.12. I want for myself to go

to occur at a stage when the child has restructured to S-complements but has not yet learned the deletion rule. While there are some dialects that allow forms like (8.12), it is by no means clear that children learning all dialects of English produce them. There is no indication that they ever postulate a reflexive which they must subsequently delete via a marked rule. This cannot be dismissed by saying that the children are avoiding the marked form altogether, since the structures like (8.12) which include reflexives are not themselves marked; rather it is the deletion rule which is not in the core grammar. If considerations of acquisitional complexity are taken into account, it looks as if there is something wrong with the proposal for reflexive-deletion, though the acquisition data cannot pinpoint the source of the problem, since they are consistent either with the rule's existing but being a core rule rather than a non-core rule, or with there being no such rule at all.[5]

Markedness may also interact with $\overline{X}$ proposals (see

Chapter 2) to produce predictions about acquisition. $\bar{X}$ theory limits the form of phrase-structure rules, stipulating that all specifiers must either precede or follow their major categories but not both. Since there are languages where specifiers occur in both positions, such as Dutch and German, these proposals cannot be seen as absolute prohibitions but must rather be considered a kind of markedness proposal, with cases where specifiers and major categories are consistent relative to one another constituting the least marked case. If the child is predisposed to start with the least marked alternative, he is likely to look for specifiers always in the same position relative to their major categories. For example, a child learning Dutch might first try to place all specifiers before $\bar{X}$, then all after $\bar{X}$, or vice versa, working within the general framework:

8.13. $\bar{\bar{X}} \dashrightarrow \langle \bar{S}pec\ \bar{X} \rangle\ \bar{X}$

or:

8.14. $\bar{\bar{X}} \dashrightarrow \bar{X}\ \langle \bar{S}pec\ \bar{X} \rangle$

The mixed positions, that is with some categories being preceded by specifiers and others followed by them, should be attempted later. Presumably, also, it should take Dutch children longer to work out the specifer systems of Dutch than it takes children to work out those of consistent languages. $\bar{X}$ theory has wider implications than just the acquisition of correct order of specifiers and X within phrases. It is possible that this kind of order helps the child to determine

sentence word-order as well, so that there will be greater problems learning the base word order of a language which is marked as far as $\overline{X}$ is concerned.[6]

Kean (1979) raises some problems that arise in a theory where markedness is defined solely as deviations from core grammar. For example, where two rules both deviate from core, they will both be counted as marked with no means of telling if one is more marked than the other. There are many cases where one would want to show that one rule is more marked than another, although both are more marked than core. Indeed, if one can determine relative markedness, this has important consequences in the field of acquisition since it would predict certain orders of acquisition, with the child proceeding from least marked to most marked through an ordered sequence. If there is no such thing as being marked relative to another marked form, such ordered sequences should not occur, or, if they do, markedness can offer no explanation of them.

It has often been noted in the literature that many children do seem to go through the same sequences in acquisition, that is, in learning a particular structure they proceed through the same hypotheses in the same order, though not at the same rate or age. C. Chomsky (1969) and Brown (1973) have shown that such orders follow if linguistic complexity is taken into account. It is worth seeing whether such results can be described in terms of markedness and whether markedness makes more specific predictions about relative complexity than "non-core" would imply.

C. Chomsky investigates the acquisition of verbs which violate the minimal distance principle (M.D.P.), which states that the unexpressed subject of an infinitival verb is understood to be the NP most closely preceding it (Rosenbaum, 1965). In the following sentences, the underlined NP is the implicit subject of the complement verb:

8.15.a. John wanted to leave
b. John wanted Bill to leave
c. John advised Bill to leave
d. John told Bill to leave
e. John told Bill what to do

The verb "promise" violates the M.D.P. in that the understood subject is not the nearest NP but the subject of the matrix verb:

8.16. John promised Bill to leave

The verb "ask" also violates the M.D.P. in certain cases. When it means "request" it does not violate the M.D.P., as in:

8.17. John asked Bill to leave

However, when it means "question", it does violate the principle, and the subject of the matrix verb assigns control:

8.18. John asked Bill what to do

Chomsky (1978) discusses the control of subjectless infinitives and proposes that all cases fall under the core rule of construal, namely "Coindex". Any prohibitions on coindexing will fall out from the conditions on logical form. Application of the rule of Coindex is "governed by lexical properties of the verb" in the matrix clause. As most matrix verbs assign control to NPs in their complements,[7] this

appears to be the least marked situation, the most common case. "Promise" is marked in that it assigns control to subject, "ask" is marked in that it sometimes assigns control to subject and sometimes to object, depending on its meaning.

If one considers "promise" and "ask" as simply two marked verbs both departing from the core grammar, one would not predict any difference between them as far as acquisition is concerned. However, C. Chomsky shows that one can make different predictions for each verb and that these predictions are borne out by the acquisition data, showing that as well as being just marked, a form can be marked relative to another marked one. She suggests, firstly, that the correct use of "promise" should be acquired after the correct use of verbs like "tell", where control is assigned by the object of the matrix verb. She also suggests that "ask" is more complex than "promise", since it is not consistent. Once children have worked out that "promise" assigns control by subject, this is always the case, but for "ask" they have to work out when subject control applies and when object control applies.

C. Chomsky's results with children aged 5 to 10 bear out her predictions on the relationship between linguistic complexity and order of acquisition. As far as "promise" is concerned, she tested this verb together with the verb "tell", which assigns object control in the regular way. She found four different stages of development. At stage 1, children assigned control of the infinitive to the matrix object in all cases,

including "promise". In stage 2, they assigned control randomly in all cases, both for "promise" and "tell", so that sometimes matrix subject had control and sometimes the object. At stage 3, they got "tell" right again, assigning object control, but were still haphazard with "promise". At stage 4, they get both verbs right, consistently applying subject control for "promise" and object control for "tell".

Thus, part of her prediction, that "promise" should be acquired in its correct usage after "tell", is borne out. In addition, from the way the children started out with the hypothesis that "promise" is like "tell", it seems that they try out the least marked case first, and only when they realize that this will not work for the data do they go on to look at more highly marked alternatives. (Furthermore, the stages the children go through to attain the adult usage of "promise" suggest that some restructuring must go on in the children's grammars before subject control is attained for this verb. That is, the details of the lexical entries as far as control is concerned must change considerably in the various stages. If restructuring is avoided by the child because of some kind of cost (Fodor and Smith, 1978; Maratsos, 1978) or if the child can only add on to grammars (McCawley, 1977), it is difficult to see how the child gets from stage 1, where "promise" assigns control by object, to stage 4 where control is assigned by subject, or to explain the changing of control assigned by "tell" in the process of learning "promise". (See also Chapter 7).

As far as the verb "ask" is concerned, she found that at first children interpreted all cases of "ask" as if it meant "tell". Later they understood "ask" as a question in certain contexts and as a command in others. When they had got the question interpretation correct, they still made errors of control, before finally reaching the stage where "ask" was both correctly understood in its "question" sense and control was correctly applied to the subject of the matrix clause. This stage was reached after the children acquired control of "promise".

There is one sense of "ask", in its request meaning, not its question meaning, where subject-control is possible. In:

8.19. John asked Bill to leave

some people interpret the sentence to mean that John asked for permission to leave. This usage ought to be particularly marked since the question of control is otherwise clearly divided between the two meanings of the verb. The request meaning normally assigns object control, so to have it assign subject control as well is very confusing. Chomsky only found one child who interpreted such sentences with subject control rather than object. Such sentences provide a clear case of where the grammar has reached its limit. This use of "ask" is so marked that there should be pressure to change it, in this case by interpreting the same data on the assumption that object control applies, as it usually does with the request sense of "ask". This is what most of the children did.

Thus, there does seem to be a correlation between grammatical complexity and difficulty of acquisition, as shown by the later acquisition of the correct usage of "promise" and "ask". That is, the children require more input data, more triggering experience, to figure out the properties of the exceptional forms. There is also a correlation with order of acquisition, as shown by the fact that the children went through similar stages in the acquisition of these verbs and that "ask" was harder than "promise". It would be a pity if degrees of complexity could not be captured in some way by the concept of markedness.[8]

8.3. Markedness and learnability - the case of filters.

Another area in the grammar where proposals for markedness have been made is in the surface filters (Chomsky and Lasnik, 1977). Surface filters are required in the grammar to rule out some of the overgeneralizations which potentially result from the operation of core rules. For example, the rule of COMP-deletion allows the deletion of subject relative pronouns, which is not permissible in Modern English, and this is in turn prevented by the operation of filter (8.1), here repeated:

8.1. * $\lceil_{NP}$ NP tense VP $\rfloor$

Also, the rule of Wh-movement into COMP could result in structures where both a Wh-complementizer and another complementizer would be present, if the optional rule of COMP deletion had not

applied. That is, structures such as (8.20) can be generated by the grammar but are not permissible in Modern English:

8.20. The man who that is tall is bald

These forms are ruled out by another filter:

8.21. *[$_{COMP}$ Wh phrase + complementizer]

(Chomsky and Lasnik's filter (18)).

Such filters appear to be language specific, since Middle English allowed the deletion of subject relative pronouns and had COMP slots consisting of [$\overline{Wh}$ + that]. On the face of it, it might seem a complication of the grammar for the child to have to add such filters, since the effect of the filters is to make a very general rule slightly less general. In Chapter 4, I have already discussed the learnability problems that arise if the child is seen as having to learn the filters: if children were to overgeneralize the rule of COMP-deletion to produce structures like those in (8.2) or if they failed to delete and produced forms like (8.20), they would need negative evidence in order to acquire the relevant filters. As children do not get significant negative evidence, it is hard to see how such filters could be acquired.

In order to avoid this learnability problem, Chomsky and Lasnik (1977) and Baker (1979b) propose that the occurrence of a filter in the grammar should not be seen as a complication at all. Rather, the filters should be considered to be present universally specified, and to operate unless there is

positive evidence to the contrary. That is, filters (8.1) and (8.21) represent the "unmarked case". In Modern English, the data provide no evidence of subject relative pronoun deletion or of a doubly-filled COMP at surface, so the filters are activated. In Middle English, on the other hand, the positive data provide instances of both of the above, so the filters do not apply, this being the "marked" case.

Obviously, this answers the learnability problem. Indeed, the proposals are made precisely for that reason. Is this a case of using U.G. to solve an inconvenient problem when other alternatives might be available? If the operation of the filters really constitutes the unmarked case, if the more highly valued grammar has the filters operating, and "more highly valued" means easier to acquire, then, as we shall see, certain predictions follow from an account which claims that the filters constitute the unmarked case which will not follow from alternative accounts where markedness is not under consideration.

Baker (1979b) provides a suitable alternative analysis which will make different claims with respect to acquisition of the filters from the theory that they constitute the unmarked case. Because of the large number of filters that a universal filter component would require to ensure that the unmarked case is acquired, Baker proposes what amounts to base-generating the permissible complementizer configurations on the basis of positive evidence. That is, the grammar of Modern English would have the following "permitting filters":

deletion and then overgeneralized to subject relatives (leaving the learnability problem of how the child would find out that the latter is not permissible in adult English). If the case, discussed in Chapter 4, of my son's systematic omission of subject relative pronouns is at all representative, it could provide evidence that the marked value of the filter is not triggered by positive data and, hence, against the marked/unmarked distinction in this particular case. Since permitting filter (8.22.d) could be overgeneralized to include subject relative pronoun deletion, the occurrence of such deletions is consistent with Baker's proposals. (However, see Chapter 4 for alternative explanations of this particular case).

It appears that markedness proposals in the theory of grammar are being used in two logically distinct ways to account for acquisition data. On the one hand, where the least marked case is equivalent to core proposals, markedness is a claim about a loss of generality, which is translatable into claims about real-time complexity and its effect on acquisition. Both unmarked and marked rules are hypothesized by the child on the basis of positive evidence in the input data e.g. there is evidence that "tell" assigns control to an NP in the complement, while "promise" assigns control to its subject, though the child requires more evidence in the latter case to establish its exceptional nature. In the case of filters, on the other hand, the presence of the filters, the unmarked case, is motivated by the <u>absence</u> of

something in the data and only the marked case requires positive evidence.

It is not enough to use markedness proposals as a means of illegitimately getting round problems raised by learnability. Certainly, one can make a filter learnable by proposing that the unmarked case is normally operative but it must also be shown that the claims of markedness with respect to filters are equivalent to the claims of markedness in other areas: the least marked alternatives should predominate over more highly marked ones in the initial hypotheses of the child and should be easier to learn. If markedness is to account for the scatter of potential grammars by showing the range from least to most highly marked and if the most highly valued, least marked, grammar is the more accessible, the child should start out by testing the more accessible hypotheses first. Data on the acquisition of filters could be used to determine whether the claims of Chomsky and Lasnik (1977) for marked values of the filters are equivalent to claims resulting from markedness elsewhere in the grammar, in which case we are no longer faced with two apparently different uses of "markedness", or whether the acquisition of filters can be explained without the concept of markedness, as Baker (1979b) proposes.[11]

FOOTNOTES TO CHAPTER 8

1. This claim involves an over-simplification. Since grammar change is local, a change which is in the direction of less marked in one area of the grammar may nevertheless cause a build-up of markedness in another area (Lightfoot, 1979). Thus, no grammar ever reaches the state of being entirely unmarked.

2. Traugott (1973) argues that explanation of acquisition and change in terms of natural processes instead of markedness will resolve what she sees as a dilemma. In many cases this turns out to be giving a different name to the same phenomena. Furthermore, the appeal to natural processes leads to some inconsistencies. Traugott claims that children might come up with marked categories which are not present in the input data, for example the category "dual". In order to do this, they, apparently, dip down into natural, universal hierarchies. However, if language learning consists of suppression of natural processes, as she claims, following Stampe (1969), and if the marked category dual is a natural process that is suppressed in many languages, then it ought to occur early as an unsuppressed form. If it emerges late, then she can hardly argue that it was first suppressed and later unsuppressed. (I have argued in Chapter 3 that one cannot speak of language being created in the absence of input data and I shall argue further in this chapter that marked

grammars arise because of the requirements of the data, so it makes no sense for the child to try out a less highly valued grammar unless the data specifically require it).

3. See also Chapter 3 for discussion of whether this reveals U.G. "creating" language in the absence of input data.

4. Brown (1973) claims that past tense forms of irregular verbs are learned before regular ones and that they may be less complex than regular forms. However, it is arguable that when the child first produces forms like "went" and "came", he is not thinking of them as the past tenses of "go" and "come" but as lexical items in their own right.

5. Other arguments for VP complements based on acquisition, though not on markedness, are given by Bresnan (1978), Maratsos (1978).

6. See Chapter 3 note 6 for discussion on the interaction of $\bar{X}$ proposals and the learning of word order.

7. That is, where there are NPs present in the complement of the verb. If the matrix verb has a subject but no object or indirect object, then the coindexing rule will have to coindex with the matrix subject, giving the right results. Thus, in:

i. John asked to go
   John wanted to go

the only available NP for coindexing is the matrix subject, which is understood as the subject of the infinitive.

8. Jackendoff (1972) suggests that the controller of the

infinitive is not dependent on grammatical relations such as subject or object of the matrix verb but on thematic relations. The controller for "promise" is the Source, whereas the controller for "permit" is the Goal, and so on. This account as it stands fails to explain why "Source" causes more problems to the language learner than the other thematic relations. The acquisition data suggest that the child starts off by assuming that all thematic relations assign the same control and that that control is grammatically defined, not thematically.

9. See White (1976) for arguments that the structure of COMP in Middle English included a /wh + that/ slot but that this has been restructured in Modern English.

10. Unless one configuration facilitates parsing more than another. Bever and Langendoen (1972) discuss the relationship of subject relative pronoun deletion to perceptual strategies. Chomsky and Lasnik (1977), Frazier (1979) discuss the relationship of filters to parsing strategies. However, while there may well be a relationship, while the unmarked case of filters may facilitate parsing, it does not require the presence of filters in order to be true. That is, while Frazier points out how filter (8.1) facilitates her parsing principle of Minimal Attachment, that same principle, without any filter, would predict problems with an empty COMP slot in the case of subject relatives. Thus, the parsing strategies cannot distinguish between a grammar with a prohibiting filter component and a grammar with a permitting filter, or base-generated

complementizer configurations. Furthermore, while there may be an interaction with parsing strategies in the case of some of the filters, it is by no means true of all of them; it is not clear that the presence of two complementizers at surface would cause parsing problems, for example.

11. In fact, Baker does propose a universal filter component as well, but not for the kind of filter which has been discussed in this chapter.

CHAPTER 9

CONCLUSION: THE INTERACTION OF GRAMMATICAL THEORY AND ACQUISITION

The relationship between grammatical theory and language acquisition is a two-way one; on the one hand, grammatical theory can explain the fact of acquisition and certain aspects of acquisition data, while, on the other hand, acquisition data can be used in support of proposals for the theory of grammar, and consideration of the circumstances of acquisition can provide means of restricting the theory.

The theory of grammar explains the fact of acquisition, the logical possibility of learning a language, by assuming that grammatical principles are innate and that the child comes to the learning task in some sense equipped with those principles. Thus, children are able to work within particular limits, as they seem to do, and they do not have to try out all logical possibilities in grammar construction, as would be expected on a purely inductive approach. Apparent shortcomings in the input data, such as a lack of evidence on ungrammaticality, paraphrase, ambiguity, etc., turn out not to cause problems to the child, since the a priori principles interacting with the data are sufficient for grammar construction without such non-primary evidence.

By assuming not only that there are a priori principles but that these principles are part of a restrictive theory of grammar, the "projection problem" is considerably reduced, since the number of analyses that a restrictive theory will allow are limited, and, hence, the child does not have to be thought of as choosing between competing alternatives for the same data. The more restrictive the theory, the closer one gets to a unique analysis for any set of data, and the less the need for a choice between grammars in terms of a simplicity-counting evaluation metric.

With the child's grammar construction constrained by grammatical principles, the class of grammars that the child constructs falls within the bounds set by the theory of grammar, falls within the definition of a "possible" grammar. This explains certain things about acquisition data, such as the non-occurrence of forms which would suggest that the child has gone beyond those limits. Since children's grammars are restricted by the theory of grammar, that theory can make predictions for acquisition, just as it makes predictions in other areas, though one must bear in mind that numerous other factors interact in the acquisition process, such as questions of maturation, cognitive development, increasing memory, parsing, etc. These factors sometimes make it difficult to determine exactly what the theory does predict in certain circumstances and one must be careful not to expect the theory to make predictions in areas outside its scope.

The output of child grammars is part of the data base to which the theory of grammar is responsible. In spite of the interest of the theory of grammar in explaining acquisition, acquisition data are not privileged, in that proposals for particular grammars or universal grammar can be and usually are made without recourse to child language. Data from acquisition can, of course, be used to test proposals of the theory of grammar, just as data from other areas can be.

Data from child language provide evidence about particular grammars, namely those child grammars that produced the language, and only those grammars. However, in so far as any grammar is a possible grammar, it provides evidence about the form of grammars in general, and child grammars are no exception to this. Nevertheless, acquisition data can only provide useful evidence about child grammars or the form of grammars in general if it is accepted that they are both psychologically real and optimal. Without this requirement, it is impossible to establish what child data are data of.

While the child's output provides data which rank with data from other areas, such as synchrony, diachrony, language pathology, and so on, as data to which a theory of grammar must be responsible, consideration of what constitutes the input data to a child can be of particular importance for the theory. That is, the primary data that serve the child as input are "deficient", in that information that is available to the linguist in constructing grammars is not available to

the child. A consideration of exactly what is missing can help both to suggest the need for certain a priori principles and to provide a learnability constraint on the grammars that the theory allows, such that the theory accepts as possible grammars only those which can be learned on the basis of the input data interacting with the a priori principles. This provides a means of further restricting the theory of grammar.

The theory of grammar is not itself a theory of language acquisition. Its aim is to explain how acquisition is possible in principle, by characterising the a priori principles that are available to the child, rather than to explain the course of acquisition itself. Nevertheless, interacting with other domains, the theory of grammar does make certain predictions about how acquisition proceeds in "real time". For example, the theory predicts the types of change that can be expected in child grammars, and the fact of change follows from the assumption that the child learns the optimal grammar and that his perception of the data changes over time. The assumption that complexity causes acquisition difficulties and that markedness is a form of complexity leads to certain predictions from markedness theory about things that will be hard to learn and acquired late.

A theory of acquisition which attempts to explain not only how acquisition is possible in principle but also how it actually proceeds is likely to involve an interaction of theories from many different areas: theories of language

processing and production, theories of maturation, cognitive development, etc. Above all, given the need to explain the logical possibility of acquisition and given the assumption of the psychological reality of grammars, such a theory must presuppose the acquisition of a particular grammar, or grammars, and the availability of a priori grammatical principles to the child.

BIBLIOGRAPHY

Andersen, H. 1973. Abductive and deductive change. Language 49, 765-793.

Aronoff, M. 1976. Word Formation in Generative Grammar. (M.I.T. Press; Cambridge, Mass.).

Baker, C. L. 1975. The role of part-of-speech distinctions in generative grammar. Theoretical Linguistics 2, 113-131.

Baker, C. L. 1978. Introduction to Generative-Transformational Syntax. (Prentice Hall; New Jersey).

Baker, C. L. 1979a. Syntactic theory and the projection problem. Linguistic Inquiry 10.

Baker, C. L. 1979b. Remarks on complementizers, filters and learnability. Paper presented at the Sloan Foundation Workshop, Stanford University, January, 1979.

Baron, N. S. 1973. Diachrony and Child Language: Phylogenetic Reflections of Ontogenetic Processes. (Indiana University Linguistics Club; Bloomington, Indiana).

Bellugi, U. 1967. The Acquisition of Negation. (Doctoral dissertation, Harvard University).

Berko, J. 1958. The child's learning of English morphology. Word 14, 150-177.

Bever, T. G. 1970. The cognitive basis for linguistic structures. In J. R. Hayes (ed.), 1970.

Bever, T. G. 1975. Psychologically real grammar emerges because of its role in language acquisition. In D. P. Dato (ed.), Developmental Psycholinguistics: Theory and Applications, Georgetown University Round Table on Languages and Linguistics, 1975. (Georgetown University Press; Washington, D. C.).

Bever, T. G., J. M. Carroll and R. Hurtig. 1976. Analogy or ungrammatical sequences that are utterable and comprehensible as the origins of new grammars in language acquisition and linguistic evolution. In T. G. Bever, J. Katz and T. Langendoen (eds.), An Integrated Theory of Linguistic Ability. (Thomas Crowell Inc; New York).

Bever, T. G. and T. Langendoen. 1972. The interaction of speech perception and grammatical structure in the evolution of language. In R. P. Stockwell and R. Macauley (eds.), Linguistic Change and Generative Theory. (Indiana University Press; Bloomington, Indiana).

Bickerton, D. 1975. Creolization, linguistic universals, natural semantax and the brain. Working Papers in Linguistics 6.3, 124-141. (University of Hawaii).

Bickerton, D. 1977. Pidginization and creolization: language acquisition and language universals. In A. Valdman (ed.), 1977.

Bloom, L. 1970. Language Development: Form and Function in Emerging Grammars. (M.I.T. Press; Cambridge, Mass.).

Bloom, L. 1975. Language development review. In F. D. Horowitz (ed.), Review of Child Development Research, vol. 4. (University of Chicago Press; Chicago).

Bowerman, M. 1973. Early Syntactic Development: a Cross-Linguistic Study with Special Reference to Finnish. (Cambridge University Press; Cambridge).

Braine, M. D. S. 1963a. The ontogeny of English phrase structure: the first phase. Language 39, 1-13.

Braine, M. D. S. 1963b. On learning the grammatical order of words. Psychological Review 70, 323-348.

Braine, M. D. S. 1971. On two types of models of the internalization of grammars. In D. Slobin (ed.), The Ontogenesis of Grammar. (Academic Press; New York).

Bresnan, J. 1972. Theory of Complementation in English Syntax. (Doctoral dissertation, M.I.T.).

Bresnan, J. 1978. A realistic transformational grammar. In M. Halle, J. Bresnan and G. A. Miller (eds.), 1978.

Brown, R. 1973. A First Language: the Early Stages. (Harvard University Press; Cambridge, Mass.).

Brown, R. and U. Bellugi. 1964. Three processes in the child's acquisition of syntax. In E. Lenneberg, (ed.), New Directions in the Study of Language. (M.I.T. Press; Cambridge, Mass.).

Brown, R., C. Cazden and U. Bellugi. 1969. The child's grammar from 1 to 111. In J. P. Hill (ed.), The 1967 Minnesota Symposium on Child Psychology, vol. 2. (University of Minneapolis Press; Minneapolis).

Brown, R. and C. Hanlon. 1970. Derivational complexity and order of acquisition in child speech. In J. Hayes (ed.), 1970.

Carey, S. 1978. The child as word learner. In M. Halle, J. Bresnan and G. A. Miller (eds.), 1978.

Cazden, C. 1968. The acquisition of noun and verb inflections. Child Development 39, 433-488.

Chomsky, C. 1969. The Acquisition of Syntax in Children from 5 to 10. (M.I.T. Press; Cambridge, Mass.).

Chomsky, N. 1957. Syntactic Structures. (Mouton; The Hague).

Chomsky, N. 1965. Aspects of the Theory of Syntax. (M.I.T. Press; Cambridge, Mass.).

Chomsky, N. 1967. The formal nature of language. Appendix to E. Lenneberg, Biological Foundations of Language. (Wiley; New York).

Chomsky, N. 1970. Remarks on nominalization. In N. Chomsky, Studies on Semantics in Generative Grammar, 1972. (Mouton; The Hague).

Chomsky, N. 1973. Conditions on tranformations. In S. Anderson and P. Kiparsky (eds.), A Festschrift for Morris Halle. (Holt, Rinehart and Winston; New York).

Chomsky, N. 1975. Reflections on Language. (Pantheon Books; New York).

Chomsky, N. 1976. On the biological basis of language capacities. Lenneberg Memorial Symposium, Cornell University, May 1976.

Chomsky, N. 1977. On Wh-movement. In P. W. Culicover, T. Wasow and A. Akmajian (eds.), Formal Syntax. (Academic Press; New York).

Chomsky, N. 1978. On binding. (Unpublished mimeo, M.I.T.).

Chomsky, N. 1979. Markedness and core grammar. Paper presented at GLOW, Pisa, April 1979.

Chomsky, N. and M. Halle. 1968. The Sound Pattern of English. (Harper and Row; New York).

Chomsky, N. and H. Lasnik. 1977. Filters and control. Linguistic Inquiry 8, 425-504.

Cook, V.J. 1973. The comparison of language development in native children and foreign adults. I.R.A.L. 11, 13-28.

Corder, S. P. 1967. The significance of learner's errors. IRAL 5, 161-170.

Derwing, B. L. 1973. Transformational Grammar as a Theory of Language Acquisition: a Study in the Empirical, Conceptual and Methodological Foundations of Contemporary Linguistic Theory. (Cambridge University Press; Cambridge).

Drach, K. M. 1969. The language of the parent: a pilot study. Working Paper no. 14, 53-69, Language-Behavior Research Laboratory, University of California, Berkeley.

Dresher, E. and N. Hornstein. 1976. On some supposed contributions of artificial intelligence to the scientific study of language. Cognition 4, 321-398.

Dulay, H. C. and M. K. Burt. 1974. Natural sequences in child second language acquisition. Language Learning 24, 37-53.

Emonds, J. 1976. A Transformational Approach to English Syntax: Root, Structure-Preserving and Local Transformations. (Academic Press; New York).

Fillmore, C. 1963. The position of embedding transformations in a grammar. Word 19, 208-231.

Fischer, S. 1976. Child language as a predictor of language change. Working Papers in Linguistics 8.4, 71-104. (University of Hawaii).

Fodor, J. A. and M. Garrett. 1966. Some reflections on competence and performance. In J. Lyons and R. J. Wales (eds.), Psycholinguistics Papers. (Edinburgh University Press; Edinburgh).

Fodor, J. A. and M. Garrett. 1967. Some syntactic determinants of sentential complexity. Perception and Psychophysics 2, 289-296.

Fodor, J. A., T. G. Bever and M. Garrett. 1974. The Psychology of Language: an Introduction to Psycholinguistics and Generative Grammar. (McGraw Hill; New York).

Fodor, J. D. 1978. Parsing strategies and constraints on transformations. Linguistic Inquiry 9, 427-473.

Fodor, J. D. and M. R. Smith. 1978. What kind of exception is "have got"? Linguistic Inquiry 9, 45-66.

Frazier, L. 1979. On Comprehending Sentences: Syntactic Parsing Strategies. (Indiana University Linguistics Club; Bloomington, Indiana).

Gold, E. 1967. Language identification in the limit. Information and Control 10, 447-474.

Goodluck, H. and L. Solan (eds.). 1978. Papers in the structure and development of child language. University of Massachusetts Occasional Papers in Linguistics 4.

Greenberg, J. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg (ed.), Universals of Language. (M.I.T. Press; Cambridge, Mass.).

Gruber, J. 1967. Topicalization in child language. Foundations of Language 3, 37-65.

Halle, M. 1962. Phonology in generative grammar. Word 18, 54-72.

Halle, M. 1973. Prolegomena to a theory of word-formation. Linguistic Inquiry 4, 3-16.

Halle, M., J. Bresnan and G. A. Miller (eds.). 1978. Linguistic Theory and Psychological Reality. (M.I.T. Press; Cambridge, Mass.).

Hamburger, H. and K. Wexler. 1975. A mathematical theory of learning transformational grammar. Journal of Mathematical Psychology 12, 137-177.

Hayes, J. (ed.). 1970. Cognition and the Development of Language. (Wiley; New York).

Hooper, J. and S. Thompson. 1973. On the applicability of root transformations. Linguistic Inquiry 4, 465-497.

Hornstein, N. 1975. S and the $\bar{X}$ convention. Montreal Working Papers in Linguistics 4, 35-71.

Jackendoff, R. 1972. Semantic Interpretation in Generative Grammar. (M.I.T. Press; Cambridge, Mass.).

Jackendoff, R. 1975a. Morphological and semantic regularities in the lexicon. Language 51, 639-671.

Jackendoff, R. 1975b. "Tough" and the trace theory of movement rules. Linguistic Inquiry 6, 437-447.

Jackendoff, R. 1977. $\bar{X}$ Syntax: a Study of Phrase Structure. (M.I.T. Press; Cambridge, Mass.).

Katz, J. and P. Postal. 1964. An Integrated Theory of Linguistic Descriptions. (M.I.T. Press; Cambridge, Mass.).

Kay, P. and G. Sankoff. 1974. A language-universals approach to pidgins and creoles. In D. DeCamp and I. A. Hancock (eds.), Pidgins and Creoles: Current Trends and Prospects. (Georgetown University Press; Washington, D. C.).

Kayne, R. 1975. French Syntax: the Transformational Cycle. (M.I.T. Press; Cambridge, Mass.).

Kean, M-L. 1979. On a theory of markedness: some general considerations and a case in point. Social Sciences Research Report 41. (School of Social Sciences, University of California, Irvine).

Keyser, S. J. 1974. A partial history of the relative clause in English. University of Massachusetts Occasional Papers in Linguistics 1, 1-33.

King, R. D. 1969. Historical Linguistics and Generative Grammar. (Prentice-Hall, New Jersey).

Kiparsky, P. 1965. Phonological Change. (Doctoral dissertation, M.I.T.).

Kiparsky, P. 1968a. How abstract is phonology? (Indiana University Linguistics Club; Bloomington, Indiana).

Kiparsky, P. 1968b. Linguistic universals and linguistic change. In E. Bach and R. T. Harms (eds.), Universals in Linguistic Theory. (Holt, Rinehart and Winston; New York).

Klima, E. S. and U. Bellugi. 1966. Syntactic regularities in the speech of children. In J. Lyons and R. Wales (eds.), Psycholinguistics Papers. (Edinburgh University Press; Edinburgh).

Labov, W. 1969. Contraction, deletion and inherent variability of the English copula. Language 45, 715-762.

Lasnik, H. and F. Fiengo. 1974. Complement object deletion. Linguistic Inquiry 5, 535-571.

Lenneberg, E. 1967. Biological Foundations of Language. (Wiley; New York).

Lightfoot, D. W. 1979. Principles of Diachronic Syntax. (Cambridge University Press; Cambridge).

Limber, J. 1973. The genesis of complex sentences. In T. Moore (ed.), 1973.

Maratsos, M. 1978. New models in linguistics and language acquisition. In M. Halle, J. Bresnan and G. A. Miller (eds.), 1978.

Maratsos, M. and R. Abranovitch. 1975. How children understand full, truncated and anomalous passives. Journal of Verbal Learning and Verbal Behaviour 14, 145-157.

Matthei, E. 1978. Children's interpretations of sentences containing reciprocals. In H. Goodluck and L. Solan (eds.), 1978.

McCawley, J. D. 1968. Review of "Current Trends in Linguistics, vol. 3", ed. by T. A. Sebeok. Language 44, 556-593.

McCawley, J. D. 1976. Some ideas not to live by. Die Neueren Sprachen 75, 151-165.

McCawley, J. D. 1977. Acquisition models as models of acquisition. In R. W. Fasold and R. W. Shuy (eds.), Studies in Language Variation: Semantics, Syntax, Phonology, Pragmatics, Social Situations, Ethnographic Approaches. (Georgetown University Press; Washington, D.C.).

McCawley, J. D. 1979. Towards plausibility in theories of language acquisition. Paper presented at the Sloan Foundation Workshop, Stanford University, January 1979.

McNeill, D. 1966. Developmental psycholinguistics. In F. Smith and G. A. Miller (eds.), 1966.

McNeill, D. 1970. The Acquisition of Language: the Study of Developmental Psycholinguistics. (Harper and Row; New York).

Menyuk, P. 1971. The Acquisition and Development of Language. (Prentice Hall; New Jersey).

Menyuk, P. 1977. Language and Maturation. (M.I.T. Press; Cambridge, Mass.).

Moore, T. (ed.). 1973. Cognitive Development and the Acquisition of Language. (Academic Press; New York).

Oehrle, R. T. 1976. The Grammatical Status of the English Dative Alternation. (Doctoral dissertation, M.I.T.).

Olmsted, D. L. 1966. A theory of the child's learning of phonology. Language 42, 531-535.

Park, T. Z. 1971. The acquisition of German syntax. Working paper, Psycholinguistic Institute, University of Berne.

Parker, F. 1977. Typology as an inadequate explanation for language change. Paper presented at a conference of the M.L.A. Change Division, Chicago, December 1977.

Penfield, W. and L. Roberts. 1959. Speech and Brain-Mechanisms. (Princeton University Press; Princeton, N.J.).

Peters, S. 1972. The projection problem: how is a grammar to be selected? In S. Peters (ed.), Goals of Linguistic Theory. (Prentice-Hall; New Jersey).

Pfuderer, C. 1969. Some suggestions for a syntactic characterization of baby-talk. Working Paper 14, 30-52, Language-Behavior Research Laboratory, University of California, Berkeley.

Postal, P. 1971. Cross-over Phenomena. (Holt, Rinehart and Winston; New York).

Prideaux, G. D. 1978. The acquisition of English relative clauses. Paper presented at the Canadian Linguistics Association Meeting, 1978.

Roeper, T. 1973. Connecting children's language and linguistic theory. In T. Moore (ed.), 1973.

Roeper, T. 1978. Linguistic universals and the acquisition of gerunds. In H. Goodluck and L. Solan (eds.), 1978.

Rosenbaum, P. S. 1965. A principle governing deletion in English sentential complementation. IBM Research Paper RC-1519. (Yorktown Heights, New York).

Sampson, G. 1978. Linguistic universals as evidence for empiricism. Journal of Linguistics 14, 183-206.

Sankoff, G. and S. Laberge. 1974. On the acquisition of native speakers by a language. In D. DeCamp and I. A. Hancock (eds.), Pidgins and Creoles: Current Trends and Prospects. (Georgetown University Press; Washington, D.C.).

Sheldon, A. 1972. The Acquisition of Relative Clauses in English. (Doctoral dissertation, University of Texas at Austin).

Sheldon, A. 1978. Assumptions, methods and goals in language acquisition research. (Indiana University Linguistics Club; Bloomington, Indiana).

Shipley, E. R., C. S. Smith and L. R. Gleitman. 1969. A study in the acquisition of language: free responses to commands. Language 45, 322-342.

Slobin, D. 1966a. The acquisition of Russian as a native language. In F. Smith and G. A. Miller (eds.), 1966.

Slobin, D. 1970. Universals of grammatical development in children. In G. Flores d'Arcais and J. Levelt (eds.), Advances in Psycholinguistics. (North Holland Publishing; Amsterdam).

Slobin, D. 1973. Cognitive prerequisites for the development of grammar. In C. A. Ferguson and D. Slobin (eds.), Studies of Child Language Development. (Holt, Rinehart and Winston; New York).

Smith, C. 1961. A class of complex modifiers. Language 37, 342-365.

Smith, F. and G. A. Miller, (eds.). 1966. The Genesis of Language: a Psycholinguistic Approach. (M.I.T. Press; Cambridge, Mass.).

Snow, C. 1972. Mothers' speech to children learning language. Child Development 43, 549-566.

Solan, L. 1978. The acquisition of tough movement. In H. Goodluck and L. Solan (eds.), 1978.

Solan, L. and T. Roeper. 1978. Children's use of syntactic structure in interpreting relative clauses. In H. Goodluck and L. Solan (eds.), 1978.

Stampe, D. 1969. The acquisition of phonetic representations. In Papers from the 5th Regional Meeting of the Chicago Linguistic Society. (University of Chicago Press; Chicago).

Strohner, H. and K. Nelson. 1974. The young child's development of sentence comprehension: influence of event probability, non-verbal context, syntactic form and strategies. Child Development 45, 567-576.

Tavakolian, S. L. 1978. The conjoined-clause analysis of relative clauses and other structures. In H. Goodluck and L. Solan (eds.), 1978.

Traugott, E. C. 1969. Toward a theory of syntactic change. Lingua 23, 1-27.

Traugott, E. C. 1973. Some thoughts on natural syntactic processes. In C. J. N. Bailey and R. W. Shuy (eds.), New Ways of Analyzing Variation in English. (Georgetown University Press; Washington, D. C.).

Traugott, E. C. 1977. Pidginization, creolization and language change. In A. Valdman (ed.), 1977.

Turner, E. and R. Rommetveit. 1967. The acquisition of sentence voice and reversibility. Child Development 38, 649-660.

Valdman, A. (ed.). 1977. Pidgin and Creole Linguistics. (Indiana University Press; Bloomington, Indiana).

Wanner, E. and M. Maratsos. 1978. An ATN approach to comprehension. In M. Halle, J. Bresnan and G. A. Miller (eds.), 1978.

Wasow, T. 1977. Transformations and the lexicon. In P. W. Culicover, T. Wasow and A. Akmajian (eds.), Formal Syntax. (Academic Press; New York).

Waterson, N. 1971. Child phonology: a prosodic view. Journal of Linguistics 7, 174-211.

Watt, W. 1970. On two hypotheses concerning psycholinguistics. In J. R. Hayes (ed.), 1970.

Wexler, K., P. W. Culicover and H. Hamburger. 1975. Learning-theoretic foundations of linguistic universals. Theoretical Linguistics 2, 215-253.

White, L. 1976. Changes in the structure of COMP. Montreal Working Papers in Linguistics 7, 201-221.

White, L. 1977. Is it possible to do all semantic interpretation off surface structure? Montreal Working Papers in Linguistics 8, 163-182.