

EARLY SYNTACTIC ACQUISITION

K. L. Kelley

November 1967

EARLY SYNTACTIC ACQUISITION

K. L. Kelley*

The RAND Corporation, Santa Monica, California

*Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of The RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The RAND Corporation as a courtesy to members of its staff.

Except for editorial changes, the text of this paper is that of the author's doctoral dissertation in linguistics, presented to the University of California, Los Angeles. It was written while the author was a participant in a seminar on computational linguistics sponsored by the Advanced Science Seminar Program of the National Science Foundation and held at The RAND Corporation.

ABSTRACT

The linguistic system acquired by a child is multifaceted; it is common to distinguish phonological, syntactic and semantic components. This Memorandum is concerned with accounting for the earliest acquisition of syntactic structure, a stage which is taken to be pre-transformational. A formal model of the mechanism for acquisition is presented, the model being stated as a computer program which is executed. To the extent to which the model is consistent with what is known about syntactic acquisition it serves as a hypothetical explanation of that process.

The principal assertions of the model are that syntactic acquisition proceeds by the testing of hypotheses about the language, these hypotheses reflecting the initial structure of the acquisition mechanism as well as the language data to which the child is exposed. It is further argued that semantics is crucial in acquisition as it provides for the confirmation or disconfirmation of grammatical hypotheses as well as comprising part of the initial structure available to the child. And finally, on this active view of language acquisition, it is maintained that sentences are used as data by the acquisition

mechanism only when they are to some degree understood.

In other words, the child responds differentially to acceptable and unacceptable sentences (unacceptable sentences are those which are either radically ungrammatical or overly complex relative to the current level of the child's language competence).

ACKNOWLEDGMENTS

The author is indebted to both Professor Barbara Partee and Dr. David Hays for considerable help and encouragement in the development of the ideas herein. In addition to helpful criticism of the publication itself, Professor Martin Kay contributed incisive solutions to programming problems encountered. The author also benefitted from comments by Professor Margaret Jones and Professor Earl Hunt on different aspects of this research.

CONTENTS

| | |
|---|-----|
| Abstract | v |
| Acknowledgments | vii |
| Introduction | 1 |
| I. Methodological Comment | 4 |
| II. Data for Language Acquisition | 10 |
| 2.1 The Methodological Problem | 10 |
| 2.2 Possible Input Data | 13 |
| 2.3 Sequence of Presentation | 27 |
| III. The Product of Language Acquisition | 29 |
| 3.1 The Methodological Problem | 29 |
| 3.2 Observations and Speculations | 34 |
| IV. The Process of Language Acquisition | 47 |
| 4.1 Braine - Contextual Generalization | 48 |
| 4.2 Jenkins and Palermo - Mediation Theory | 55 |
| 4.3 McNeill - Nativism | 64 |
| 4.4 Implications of Unacceptable Input Data | 82 |
| V. A Model of Language Acquisition | 92 |
| 5.1 Flow Diagram | 92 |
| 5.2 The Process | 95 |
| 5.3 An Experiment with Ungrammatical Input Data | 143 |

| | |
|-----------------------------|-----|
| 5.4 Summary | 146 |
| VI. Summary | 148 |
| Appendix | |
| A. Tables | 153 |
| B. Computational Discussion | 159 |
| C. A Program Trace | 168 |
| Bibliography | 191 |

INTRODUCTION

This Memorandum is concerned with the initial stages of syntactic acquisition during first language learning. It will not deal directly with later syntactic acquisition nor the acquisition of phonology or semantics, although semantics will be considered inasmuch as it relates to syntactic acquisition.

Research on the process of language acquisition has multiplied in the last decade. Much of the stimulus for this can be credited to the rediscovery by both linguists and psychologists that language is structured, and that the speech and behavior of a child may be studied with the goal of describing what language system the child has internalized at a given point in time. The trend toward mentalism in linguistics has interacted with an increased appreciation for the formal apparatus of linguistic theorizing (machinery for grammatical description) with the result that considerable attention is now being directed toward the study of language competence in children.

The problem of language acquisition has been characterized as the problem of describing (1) the input to the psychological mechanism whereby the child acquires a language; (2) the result of the acquisition process (the

language competence acquired by the child); and (3) the process whereby the child gets from (1) to (2) (Fodor 1966: 106). This tripartition has been used in this report: Chapters II through IV deal with (1) through (3) respectively. The literature relevant to each section is reviewed in that section. Chapter V is a detailed description of the language acquisition model, the general characteristics of which will have been hinted at in the previous chapters. Prior to consideration of these topics, however, there is in Chapter I some methodological comment on the research problem.

In discussion of language acquisition or universal theories of grammar a distinction between formal universals and substantive universals is often drawn (Chomsky 1965:28). A formal universal specifies a condition upon a theory while a substantive universal specifies one or more items, sets, or rules that are components of a theory. The distinction is not always easy to apply, and nothing hinges upon it in this report, but in an intuitive sense the research described here has been concerned principally with the specification of some formal universals. English has been used as the language upon which most of the discussion and results are based, but my speculation about substantive universals and/or details of English syntax acquisition should be

regarded as far more tentative (and in my view less important, given the present state of the art) than the conclusions about formal properties of the mechanism for language acquisition.

CHAPTER I

METHODOLOGICAL COMMENT

This chapter is concerned with the general problem of psychological theorizing. Much of what follows has been said elsewhere, but as this is the first use of a computer program in this particular domain, as a model of the language acquisition mechanism, it seemed wise to describe the technique.¹ The questions to be considered are: (1) What makes a computer program a psychological model? (2) How does one test or validate the model? (3) What are the advantages and disadvantages of using computer programs as models?

1. Put simply, a computer program is a psychological model if it is intended to perform some psychological task in the way the organism being modeled (hereafter a person) performs the task. But the phrase "the way a person performs a task" needs explanation, as computers do not do anything exactly the way a person does, one reason being that the

¹There are a number of ways in which the term model has been used; here the term refers to analogue models in the sense of Max Black (1962), where there is something actually constructed that is isomorphic to the object being modeled.

hardware is different; people use nerve and brain cells while computers use diodes, ferrite cores, and so on. But this difference between a model and the object being modeled always exists to some degree; the two are never exactly identical. The requirement is that, for some object X to be a model of an object Y, it must be the case that at some level of detail there is an isomorphism between X and Y. And just what isomorphism is intended to hold between X and Y must be stated for one to be a model of the other; the isomorphism is never established by the computer program itself. The correspondence between a program and a behavioral mechanism is not, typically, at the level of individual computer instructions, but at a higher conceptual level, perhaps in terms of program segments or subroutines. But the level is determined not by the program but by the model builder, who must, so to speak, append a description to his program that says just what the units, processes, and relations are in the program that are intended to correspond to psychological units, processes, and relations. In short, a computer program is a psychological model if one has stated an isomorphism between the program and the psychological mechanism being modeled.

2. There are two questions to be asked with respect to a model: (1) Is a particular assertion about the model true?; and (2) is the putative isomorphism between the object and its model valid? The truth value of an assertion about a model may or may not be easy (or even possible) to ascertain. Typically, when dealing with computer programs as models one is concerned with systems that are not solvable mathematically. Thus, verification of a claim about a computer model usually proceeds by executing the program and inspecting the program's operation and results to see if the claim is supported. When done carefully the psychological confidence one has in one's conclusion is nearly the same as if the conclusion had been derived mathematically.

Certainly in most psychological investigations the more difficult question is to validate the psychological reality of the theoretical constructs (the objects, processes, and relations) which the model supposedly represents. But verification here is no different from verification in any other psychological study: objects, processes, and relations which may be studied directly are so studied, while those which may be studied only inferentially are studied through the making of inferences.

Hypothesized objects or mechanisms which are not observable are confirmed (disconfirmed) to the extent to which they do (do not) provide explanations for the behavior of the person. Selection between alternate formulations is done by comparing the degree to which they can be generalized to account for other behavior, the extent to which they are parsimonious, and the extent to which they lead to other verifiable observations.

3. The point of a model is to make the study of a system more tractable by establishing an isomorphism between that system and a second system more easily studied than the first. And the particular advantage of computer programs as models is that the content or implications of the model may be studied by executing the program.

To illustrate, take the case of first language acquisition. Suppose that somehow one has a hypothesis about what input data are necessary and what data are extraneous for language acquisition. To verify the hypothesis one would like to be able to take a child and expose him to just the "necessary" input data and see if in fact he acquires language competence. But this is not likely to be possible -- children are not, after all, laboratory animals -- and so it will not usually be possible to control

for all extraneous variables (in terms of this hypothesis) that might confound one's results. A solution would be to create an acquisition model where one can control the input data and, equally importantly, inspect the output that results. Because of the specificity of the mechanism there is not likely to be a ready-made analogue that can serve as a model, and so one is built by writing a computer program. As will be seen later, the controls on input variables and experiments that are possible make this approach a powerful one.

There are reasons, however, for being cautious in building computer program models. First, one should not overlook the possibility of determining mathematically some of the properties of a model in place of simply making inductions on the basis of some program executions. When mathematical study is possible it is to be preferred, for when a system's properties are formally deducible rather than informal inductive generalizations, the system can in a certain sense be said to be controlled. Second, there is the danger of confusing study of the properties of a model with empirical verification. Validation of an assertion about a model is not the same thing as validation of the isomorphism between the model and the object being modeled.

One might be able to demonstrate very convincingly the truth of certain claims about a model without, however, being able to show that the model is in fact an accurate representation of the object being studied. Third, creating programs is an expensive and time-consuming enterprise. What one gets from it is a model that can be studied and perhaps some ideas that were suggested during the process of formalization. But it would be wise to consider the possibility of creating a model that did not require such an expenditure of time and money, perhaps a paper and pencil model. If the system being studied is sufficiently simple, it may be possible to obtain the same results one would get from writing a computer program by recourse to one's ingenuity and intuitive overall understanding of the system. The point is simply that it is not usually very interesting to have a computer program that will model some psychological process per se unless it contributes to an understanding of that process.

CHAPTER II

DATA FOR LANGUAGE ACQUISITION

How one investigates what data a child uses in language acquisition is the first topic in this chapter; it is followed by a look at some possible types of language data, and then a comment about the sequence in which language data is presented.

2.1 The Methodological Problem

The concern of this chapter is to discriminate between data which are normally used by the child in the process of language acquisition and data which are used not at all or only incidentally or under exceptional circumstances. One would like to run controlled experiments varying the types of data available, but because of the lengthy time during which a child would have to be withdrawn from normal social intercourse (probably several years) this is not possible. There appear to be four more or less devious routes to obtain the information unavailable through direct experimentation.

1. Experiments performed by nature. One might inquire whether the kind of experiment out of bounds to psychological investigators has not been performed by nature. Deaf or disabled children, and children raised

in very restricted (verbal and nonverbal) environments, may provide information about the variables that contribute to language acquisition.

2. Indirect psychological experimentation. One can run very limited experiments with children, using artificially constructed languages, and use the results of such experiments to make inferences about the process of first language acquisition. Other nonlinguistic psychological experiments may be performed and the results generalized to language acquisition.

3. Direct psychological study. One can attempt to study the course of language acquisition directly without, however, hoping to control for all possibly relevant input variables. There is not a great deal of certainty attached to one's conclusions in this type of study, but nonetheless this has still been the most productive mode of investigation.

4. Modeling. One can attempt to model the mechanism and process of language acquisition in the hope that hypotheses and results will be more easily obtained by using the model than without it.

The first three techniques have all been used with some success. However nature has only run a few of the

experiments whose results are of interest, and the unsettling question about artificial language and nonlinguistic experiments is whether the mechanism or behavior being studied in the experiment has any relation to the mechanism of first language acquisition.¹ That it may not have is suggested by some of the inductive experiments where subjects were able to learn languages with properties not present in natural languages (e.g. a cyclical language); apparently the inductive mechanism used is not constrained in the way in which many linguists suppose the language acquisition mechanism is constrained.

Direct psychological study may be either observational or experimental. In observational studies one makes inferences about the relevance of one or another variable for language acquisition on the basis of facts about the child's environment and his verbal behavior collected over a period of time. In experimental studies the child's environment is modified in one way or another, and conclusions are drawn on the basis of what happens.

¹See G. Miller (1967:178) for a negative discussion of this question.

An investigation of the fourth type is the concern of this study. In this chapter tentative discriminations between important and unimportant sources of data will be made. One step in establishing the validity of the discrimination will be to show that the categorization of important sources of input data is sufficiently rich to allow for acquisition of an appropriate grammar. (See Chapter V.)

2.2 Possible Input Data

In this section possible input data will be partitioned into three classes: (1) verbal data that are not contingent upon the child's speech; (2) data (both verbal and nonverbal) that are contingent upon his speech; and (3) nonverbal noncontingent data.

2.2.1 Verbal Noncontingent Data

Verbal data that are not contingent upon the child's speech would include:

- (a) examples of sentences from the language;
- (b) examples of nonsentences from the language identified as such;²

²A nonsense presented to the child would have no positive value (and indeed would probably be confusing) if he used it in the same way he used an example of a

- (c) pairs of sentences syntactically related
(in the sense of transformational grammar)
identified as such;
- (d) pairs of sentences which are syntactic or
semantic paraphrases identified as such.

The reason for distinguishing (c) and (d) from (a) is that one can imagine a child, in the process of formulating a grammar, using two sentences known to be syntactically or semantically related in a manner distinct from how he might use two unrelated sentences. For this to be true it must be possible for the child to recognize the relationship between the sentences prior to the relationship being signaled by his grammar. How the child might do this is not clear.

Most examples of nonsentences identified as such are contingent in the sense that they occur in connection with the child's speech attempts (see (a) and (b) of 2.2.2 below).

Although all of the above are possible types of information about the language, only (a) is used by the model. (See Chapter V.)

grammatical sentence. I mean, therefore, to refer to the case where the child knows that he has been presented with a nonsentence. Similarly, (c) and (d) assume extra-sentence knowledge.

2.2.2 Data Contingent Upon Child's Speech

Contingent data (both verbal and nonverbal) of possible importance in language acquisition would include:

- (a) corrections of a child's speech;
- (b) expansions of a child's speech³;
- (c) differential reinforcement of a child's speech.

The evidence that bears on the role of corrections and expansions will first be considered followed by evidence relative to the role of differential reinforcement. None of these data, however, will be input to the model described in Chapter V.

2.2.2.1 Corrections and Expansions

- (1) Experiments performed by nature

- (i) Lenneberg (1962) has reported a case study of an eight year old boy with a congenital disability for the acquisition of motor speech

³By expansion of a child's utterance is meant the repetition of that utterance with the addition of missing function words, auxiliaries, and inflections. Accordingly this is simply a special type of correction. If a child were to say "That Christopher car" an adult might respond with the expansion "That's Christopher's car."

(anarthria). Although this boy, who Lenneberg says is representative of many similar people, has never spoken a language, he nonetheless has a "normal and adequate understanding of spoken speech". Clearly corrections and expansions are not relevant in these cases.

(ii) Ervin (1964:163) has reported (without references, unfortunately) that children "who were believed for years to be mute have been found employing relatively mature grammatical patterns." The observation, if it can be supported, would be quite inconsistent with any theory in which corrections and expansions are taken to be crucial.

(2) Indirect psychological experimentation

(i) Braine (1963a) has run a series of experiments with children (age range 4 to 11) to test their acquisition of productive rules given "sentences" from an artificial language. In two of the experiments the children were provided with negative instances (non-sentences), identified as such, as well as positive instances, with resulting impairment of learning.

(ii) In an experiment in concept formation by Whitman and Garner (1963), a task

formally akin to grammar induction, it was found that presentation of negative instances, identified as such, interspersed with positive instances, impaired concept acquisition when compared to the condition when positive instances were presented alone.

(3) Direct psychological study

(i) In her doctoral dissertation (1965) Cazden tested the hypothesis that expansions would materially aid language acquisition over and above their value as models. She divided twelve children into three groups, one group receiving "intensive and deliberate expansions", a second group receiving qualitatively equal exposure to well-formed sentences that were not expansions, and the third group receiving no treatment at all. The period of exposure was twelve weeks. Her results did not show the expected difference between groups, however, and so her conclusion was that expansions probably do not have any particular value in language acquisition by virtue of their being expansions.⁴ This conclusion was also supported by an additional study (Cazden

⁴However her conclusion was based on a failure to reject a null hypothesis and so is only suggestive.

1967) where the child whose speech had the smallest percentage of adult expansions was the most advanced grammatically (with respect to inflections.)

(ii) There appears to be a fair consensus among researchers who have collected data that there is little direct syntactic correction of child speech by adults (Braine 1963a:334; Braine 1966b; Brown and Fraser 1963:203). Brown and Bellugi (1964:140)⁵, however, report observing expansions of child's speech by their mothers about 30% of the time.

2.2.2.2 Differential Reinforcement

The claim that language abilities and capacities develop by the careful molding of a child's attempts at speech by differential reinforcement⁶ has been often suggested; in this section some evidence is considered.

⁵See also Slobin comment in discussion of McNeill (1966d:86).

⁶This claim should be distinguished from the more general assertion that differential reinforcement of something, not necessarily speech, is crucial to language acquisition. See Chapter IV.

(1) Experiments performed by nature

Both of the observations previously mentioned with respect to corrections of a child's speech apply as well here to disconfirm a theory that emphasizes the role of reinforcement of a child's speech efforts.

(2) Indirect psychological experimentation

In a cleverly designed experiment by Earl Hunt (1965) active learning was contrasted to passive learning in a concept acquisition task. An artificial language was created by the experimenter. Subjects in the first of two groups could select possible sentences and be told whether they were or were not in the language. Each subject in the first group was matched with one in a second group, and the information available to the subject in the first group about whether an example was or was not in the language was given to the matched subject in the second group. In spite of receiving exactly the same information, subjects in the second group did poorer in acquiring the concept than subjects of the first group. Perhaps this result is explainable in terms of reinforcement theory, but Hunt suggests that a more reasonable learning theory would deal with hypothesis-testing instead of the passive establishment of associations.

(3) Direct psychological study

(i) In a study concerned with infant vocalization Weissberg (1963) found that the quantity of vocalization could be increased by reinforcement via contingent social stimulation. However, this result, as well as the many verbal conditioning experiments with adults, while having implications for an explanation of verbal performance, has no clear application to an explanation of language acquisition.

(ii) That there is more to language acquisition than simply the confirmation of acquired forms is shown by the phenomena whereby a child initially acquires some correct forms (which he is presumably reinforced for using) but later loses them by virtue of overgeneralizing. Still later the correct forms will be reacquired. Examples from the literature include conjugations of strong verbs (Ervin 1964:178)⁷, negation (Bellugi 1964), and some Russian conjugations (Slobin 1966a:138). It might turn out, however, that the phenomena could be accounted for by using mediational paradigms.

⁷Charles Ferguson and Hans Kalmus made observations similar to Ervin's about French and German respectively in the discussion following McNeill's paper (1966d).

To summarize, there appears to be no evidence supporting the contention that differential reinforcement of the child's speech efforts is necessary for language acquisition, and some evidence that suggests the contrary, the strongest of which are the observations about children who have learned a language without such reinforcement. As noted, however, this conclusion is not a general disavowal of any possible reinforcement theory of language acquisition.

2.2.3 Nonverbal Noncontingent Data

By nonverbal noncontingent data is meant semantics. The two questions that might be considered are: (1) Do semantic data play a central role in language acquisition?⁸ and (2) if so, what is the nature of the semantic support for acquisition? The second question is considered in Chapters IV and V; here evidence is reviewed that bears on the first question.

⁸The use of semantics in providing the motivation for syntactic acquisition without, however, affecting the acquisition mechanism (see Chomsky 1965:33) is here considered a supporting (non-central) role.

(1) Experiments performed by nature
None.

(2) Indirect psychological experi-
mentation

Children and adults are able to learn artificial content-free languages of varying complexity, as shown by a variety of experiments (Braine 1963a, Esper 1925; K. Smith 1963; G. Miller 1967). In several cases the languages were given some content (the "words" had referents) with somewhat contradictory results. In one of the Harvard "Project Grammarama" experiments, Norman compared language learning with a language where words had no referents with the same language where words did have referents (G. Miller 1967). Acquisition appeared to be unaffected, although in the latter case the tendency to think, prematurely, that one "knew" the language was not present, although it was in the former. In an experiment of Neal Johnson's (1962), reported by Jenkins and Palermo (1964:167-8), there was a marked contrast in the extent to which a subject could master an artificial language with semantic reference in contrast to the same language without reference. The conclusion was that at least in this case correct categorial assignment of words depended upon the existence of semantic

cues in addition to syntactic cues.

(3) Direct psychological study

(i) There have been a number of observations that focus on the parallel development of semantic concepts and the formal syntactic mechanisms in the language that express those concepts, although what can be inferred from the parallelism about the relationship between semantics and syntactic acquisition is not entirely obvious. The general conclusion of such studies is that categorial distinctions are acquired (a) in direct relationship to the obviousness of the semantic properties of the category, and (b) at the same time as control of the semantic distinctions implied by the category (Brown 1957; Slobin 1966a:141-2; see the references in Braine 1966b). For example, Slobin establishes the following order for the acquisition of morphological distinctions in Russian: first, classes with concrete references (e.g. number distinction), followed by classes based on relational semantic criteria (e.g. case, tense). Grammatical gender distinctions are learned last, being almost void of semantic correlates in Russian.

Another observation on the parallel development of semantics and syntax was made by Warner and Kaplan (1950). They investigated the formation of word meanings by introducing a meaningless word in a variety of contexts and asking for definitions of the word. There seemed to be several types of definition children would make: (a) one where the meaning of the word was identified with the entire sentence context in which it occurred; (b) another where there was an identification of meaning with some context less than the whole sentence; and (c) the usual (adult) word meaning. Occurrences of (a) or (b) decreased with the child's age, but at age 11 there was an abrupt rather than a gradual decrease in examples of type (a) but not type (b). Without attempting to explain this Warner and Kaplan remarked that there was a parallel discontinuity at the same age with respect to the child's manipulation and use of sentences (the sentences in which the meaningless words occurred were often modified by children up to age 10 or so and infrequently thereafter.) Apparently something of moment occurs around that time, and since it affects both semantic and grammatical structure, this is regarded

as supporting the "genetic interdependence of meaning and structure". It is not clear what Warner and Kaplan meant by genetic interdependence, but Jenkins and Palermo cite this experiment to support their view of language acquisition in which semantic properties of classes and constructions are the basis for the acquisition of the classes and constructions in that they are what the child "recognizes and responds to". It does not seem, however, that the Warner and Kaplan experiment bears on the theory of language acquisition because of the character of the supposed changes in grammatical structure. It is unlikely that these changes reflect shifts in the acquired grammatical structure rather than simply a shift in the mechanism that uses that structure since the child must have attained nearly full grammatical competence long before age 10, for otherwise his early facility in speaking and understanding is unexplainable.

(ii) Roger Brown (1957)

has suggested that the basis for the child's early classes

are semantic properties.⁹ He observed that children's nouns are more likely to name concrete things than adults' nouns, and children's verbs are more likely to name actions. Furthermore these semantic characteristics are nonaccidental properties of the child's classes in the sense that the child associates the property with the class. This was demonstrated by an experiment in which a meaningless word was used in some context that identified it as a mass noun, count noun, or verb. By having the child make a choice among three pictures it was determined whether the child had associated the meaningless word with a mass object, a count object, or an action; the performance of the children showed that they had indeed made semantically consistent associations. Showing that the child's early categories have semantic properties does not establish that the classes were created by reference to those properties, but it adds a certain plausibility to the hypothesis.

⁹ A suggestion incorporated in the Jenkins and Palermo (1964) theory, and also in the model described in Chapter V.

2.3 Sequence of Presentation

Since language acquisition is not an instantaneous accomplishment, the order in which linguistic data are presented to the child may have some importance. One can distinguish two notions: (1) that an external ordering of linguistic data is sufficiently important so that language acquisition would not take place if the ordering were not observed; and (2) that somehow the child imposes an internal ordering on the linguistic data he is exposed to, which data may themselves be unordered.

The first proposition can probably not be immediately rejected, for it is the case that in most speech to children one does not use grammatical structures which are known to be beyond their understanding except when such structures are obligatory (eg. number and tense in English.) Could children acquire a language by being exposed only to discourse between adults? I have been told of cases of children who have learned their language simply by continual exposure to television with no significant verbal interchange with parents or other children, but I have found no confirmation of this in the literature. At any rate, in the model described in Chapter V, an external

sequencing of input data is not assumed.

Discussion of the second proposition is deferred until Chapter IV except for noting that E. Hunt's (1965) experiment contrasting active with passive learning indicates that within some problem domains the ability to impose a sequence on the data being considered does affect performance (favorably).

CHAPTER III

THE PRODUCT OF LANGUAGE ACQUISITION

3.1 The Methodological Problem

Investigation of child language has difficulties not present in the investigation of adult language. It is not possible to become one's own informant by learning to speak the child language, nor can one get intuitive judgments about the language as is possible with adult informants (Brown, Fraser, and Bellugi 1964:91; Brown and Bellugi 1964:135). At present almost all theories of what children have acquired are based upon collected records of observed speech (Braine 1963b; Brown and Fraser 1963; W. Miller and Ervin 1964; Gruber 1967; Weir 1962), although there have been a few manipulative studies where elicitation was attempted (Berko 1958; W. Miller and Ervin 1964; cf. the Russian references cited in Braine 1966b), and an occasional study where comprehension of language rather than speech was investigated (Brown 1957; C. Smith 1966; cf. the references in Braine 1966b). There are at least two potential sources of error in basing a grammatical theory upon a corpus of observed speech:

- (1) The underlying grammatical structure internalized by

the child may be obscured by various performance effects on the child's speech (attention shifts, forgetting, interruptions, etc.); (2) the underlying grammatical structure may be fully used only in the process of speech comprehension, part being unused during speech production.

That the linguistic system of a child will be only imperfectly revealed by his performance (productive or receptive) is something that one has more or less to live with. An investigator will of course use his judgement in disallowing examples confounded by obvious performance effects. An interesting objective technique for sorting out agrammatical utterances from a corpus of a child's speech was that of Menyuk's (1963) where she had children repeat sentences (some grammatical and some not, at least by adult standards) that they had spoken on another occasion. The repetition was correct in general for the grammatical sentences but incorrect for the agrammatical ones; the children usually corrected the agrammatical sentences and made them grammatical. The experiment was done with nursery school and kindergarten children and rather lucidly demonstrates both the existence of syntactic rules in very small children and the fact that children deviate from

those rules when speaking.^{1,2}

The second possible source of error, that comprehension might be more advanced than production, has been the subject of much discussion. On the one hand it is certainly a priori possible that much of a child's grammar would not be used in his speech; the limiting case would be that of a child who did not speak at all (cf. Lenneberg's case study.) Furthermore this would be consistent with a gradualistic theory of language acquisition in which there are intermediate stages between the absence of a mental construct and its full integration into the system, and where the system itself need not be entirely self-consistent. That is, there is a certain asymmetry in such a theory between the child's

¹The results reported by Menyuk do not, unfortunately, distinguish between those children who correct an ungrammatical utterance which exemplifies a grammatical error they make from those children who make the same correction but who do not themselves make the error in question. To the extent to which it is only children making corrections of other children's errors the result is less interesting than it would otherwise be, for this would not illustrate the ability of children to recognize their own performance deviations.

²C. Smith (1966) has also reported an experiment that involved, in part, the repetition by children of ungrammatical sentences created by the experimenter. Very frequently the children corrected ("normalized") the stimulus sentence in their imitation.

production and his grammatical competence and the child's comprehension and his grammatical competence. To illustrate with a hypothetical analysis of some actual observations, suppose a child learns some strong verb conjugations in English; e.g. throw-threw, swim-swam, etc. And then suppose later the child discovers the more general verb conjugation; e.g. walk-walked. This conjugation, being a far more general pattern than the strong conjugations, perhaps gets more fully integrated into the child's system than the rules for strong verb conjugation. Consequently in speaking one would expect the child to use the more fully integrated rules, and so at this stage strong verbs would be improperly conjugated. If one then based his description of the child's grammatical competence simply on the child's speech one would posit that the rules for strong verb conjugations had been lost. But suppose the child is required to comprehend a sentence with a strong verb. If it were possible for him to understand the sentence using his very general verb conjugation rule he would--but it is not, and

so he must use less well integrated rules. And in this case his strong verb conjugation rule does allow him to understand the sentence, although at this stage he would not produce such sentences.

Considerations of usage might also suppress the child's utilization of some construct in speech which construct would, however, be used in comprehension when necessary. Simple adult examples might be either somewhat rare words whose meaning one is not sure enough of to use but which are understood in context, or somewhat awkward multiply embedded sentences.

The small amount of experimental and observational data that has been reported supports the view that comprehension is more advanced than production, but leaves open the degree to which this is so. In an experiment by Fraser, Bellugi, and Brown (1963) the ability of children to imitate, comprehend, and produce speech was tested, with the expected result that the children could comprehend speech beyond their ability to produce (and they were most able at imitation.)

Chomsky (discussion following W. Miller and Ervin 1964:39) and Lees (discussion following Brown, Fraser, and Bellugi 1964:94) have suggested that the sudden

appearance in a child's speech of a full system indicates that the system had been controlled prior to its use in production. The examples they cite are a phonological distinction and the system of auxiliaries in English, although the supposed facts on which they base their conclusions have been contested (Braine 1966b).

The point of these examples is to suggest that speech and comprehension are not necessarily symmetric with respect to the grammar underlying performance. The implication of this for grammatical theories based on a speech corpus may not, however, be too serious. The data presented above suggest that the child may possess rules used in comprehension that are not used in production. But this is only to say that the grammar inferred to underlie speech production may be incomplete, and that if one had good tests of the structures underlying comprehension one would discover some rules left out. But for a grammatical theory to be incomplete is not as serious a defect as it would be if a grammatical theory based on speech production were essentially inconsistent with a theory based on speech comprehension.

3.2 Observations and Speculations

In this section some of the earliest grammatical

structures revealed in the speech of children around the age of 2 will be described. One word, two word, and some three word sentences will be considered in turn, followed by a discussion of grammatical relations. It is not claimed that progress in language learning can be reasonably characterized by the length of the child's spoken sentences; however as the subject is not being treated comprehensively this organization is convenient for making some selected points.

3.2.1 One Word Sentences

There is rather little to say about these sentences except that the words typically refer to concrete objects or situations. It sometimes requires a bit of imagination to interpret all one word sentences as having concrete references, however, as when a child says "want". It does not seem likely, however, that the word is meant simply to be an expression of an emotional state; there is a subject and an object involved. The development of this particular word might be as follows: (i) When first acquired the word refers to an entire situation in which ego desires a particular object or event. The word therefore refers to a subject, emotion, and object all at once. (ii) Soon, however, the child learns to apply this word

when there are different objects of his desire. He may say "want" to indicate a desire for some food, a toy, or whatever; the object typically is spatiotemporally contiguous to the child and psychologically prominent. When this development occurs the utterance "want" should no longer be considered a single word sentence, but a two word sentence with an unexpressed object. One could argue that unless he sometimes overtly expresses objects this should still be considered a one word sentence; however this would suggest that the important linguistic step is taken when the child produces a two word sentence with an object, whereas it seems that the important linguistic step is in fact the fractionation of the holophrastic concept "want" into two parts. (iii) Still later the child learns to apply the word in situations with other subjects, and now its development is more or less complete. Usually by this stage, however, the object of "want" is overtly expressed, and so one would not get the situation where the single word utterance "want" was in actuality a three word sentence.

Notice, by the way, the problems this view entails for the analysis of child speech. In place of a reasonably straightforward taxonomic analysis of a speech

corpus, one is concerned with uncovering the meanings of words and the functional relationships in terms of which the child analyzes and produces speech. And neither meanings nor functional relationships may be studied in the straightforward way in which word combinations may be studied.

3.2.2 Early Two Word Sentences

Descriptive studies of early child speech tend to agree in distinguishing two classes of words, what Braine (and McNeill) call the pivot and open classes. The open class is large and new members are easily added, while the pivot class has many fewer members. Pivot words tend to be used much more frequently than open class words. A pivot word typically occurs in either initial or final position in two-word sentences, but not in both positions. If one uses an adult's grammar as a measure of these classes, each may be heterogeneous (e.g. some of the descriptive studies included adjectives and verbs in both pivot and open classes; see 4.3). Two word sentence patterns are pivot + open or open + pivot. Open + open sentences also occur, although Braine has argued that this is a later pattern that actually subsumes more than one sentence type (Braine 1966b). Along with these two word

sentences there are still one word sentences consisting of an open word; pivot words rarely occur alone.

What this characterization actually says is that there are two word sentences consisting of any two words except that there may be two separate subsets, one of which has members that can occur only in sentence-initial position (initial pivots) and a second which has members that can occur only in sentence-final position (final pivot). One might state this by the rules:

$$\begin{array}{ll} S \longrightarrow O + P_f & O = \text{open class} \\ S \longrightarrow P_i + O & P_f = \text{final pivot class} \\ S \longrightarrow O + O & P_i = \text{initial pivot class} \\ S \longrightarrow O & \end{array}$$

An alternate characterization would be:

$$\begin{array}{ll} S \longrightarrow X + X & X = \text{all words} \\ S \longrightarrow X & \end{array}$$

where the lexical entry for each sentence-final pivot has the "subcategorization" feature (see Chomsky 1965) + $X_{__}$, while the lexical entry for each sentence-initial pivot has the feature + $__X$. The two characterizations are not exactly equivalent, however, for the one using idiosyncratic lexical subcategorization does not suggest

that the pivot class has a psychological unity for the child whereas the characterization in terms of pivot and open classes does. And if the pivot class is a psychological unit one would suppose that it might have some consequences for later stages of development; but in fact the putative pivot class seems to exist only evanescently. McNeill, and Bellugi and Brown, have described the differentiation of the pivot class which, if it were the only categorial revision going on, might confirm the unity of the class. However it is not -- for example a class of adjectives is formed by drawing its members from both the supposed open and pivot classes -- and so their description is taken as a consequent of their pivot and open distinction and not support for it.

There are several different types of functional relations expressed in two-word sentences. In Table 1 of Appendix A some typical sentences are listed with suggestions about what relations might hold, but all of these suggestions are very speculative. One should not interpret the functional labels as expressing their usual adult meaning; they are only approximate names. The point of the examples is to show how very difficult it is to guess at what functional relations the child is expressing in his

speech, and also how diverse formally similar structures may be when one speculates about functional relations. In the face of this "mushy" data from which to infer functional relations one can easily understand the motivation for choosing to base a grammatical theory on distributional data alone. But if the expression of functional relations is essential in the production and derivation of meanings, then one cannot choose to disregard the study of such expression.

One rather well-supported generalization about child speech at this stage is that subjects rarely are overtly expressed. This might be explainable as a consequence of the usual situation in which children speak; their early speech is present tense and directed toward a single person, and the contextual situation would almost always resolve any possible subject ambiguity between the child and the person he was speaking to.

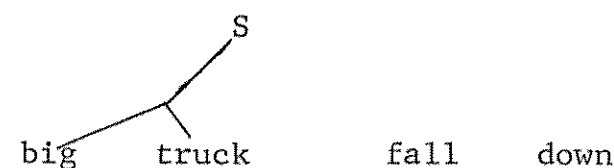
3.2.3 Early Three Word Sentences

The prior comments about the diversity of sentence types are equally relevant in the case of three word sentences. However, as the child matures and his sentences increase in length, it becomes easier to identify the functional relations expressed, for it seems reasonable

to suppose that the relations available are increasingly similar to those of an adult. Also it may be possible by the use of judicious questioning to elicit information that will help in the identification of functional relations (see Braine 1966b). "Replacement sequences" may also be of help (Braine 1966b). The term refers to a sequence of utterances spoken by the child, all presumably equivalent in meaning and occurring in a short time period, the longer utterances being expansions of the shorter ones. Thus when one hears the sequence "Man. Car. Man car. Man in car. Man in the car." (Braine 1966b) spoken by a child, one sentence after another, one can assign the structure noun + locative to the first two word sentence (instead of possibly possessive + noun, or something else).

One of the most important facts about three word sentences is that this is the first stage at which the child's sentences have hierarchical structure. Typically a three word sentence consists of two words that form a unit which combines with the third word to form a sentence. There is internal linguistic evidence supporting this view, the most important being that revealed by replacement possibilities. If a sequence may be replaced by a single word or another sequence itself known to be a unit, one

hypothesizes the unity of the original sequence. Another replacement sequence of Braine's (1966b) demonstrates this: "Truck fall down. Big truck fall down. This fall down." where "truck" is replaced by "big truck" which is later replaced by "this". It appears that "big truck" is a unit with internal structure. But this is to say that the sentence structure is (in part):



Whether "fall down" is itself a unit or only two parts of a tripartite sentence is not known from this example. Brown and Bellugi (1964) have similar examples where a single word noun is replaced by a noun phrase.

In addition to the replacement possibilities, Brown and Bellugi (1964) have mentioned that pauses in a sentence will tend to occur between constituent boundaries rather than within a constitute, and this provides further support for attributing a unity to such constitutes.³

³See Suci (1967) and the references he cites for discussion of pauses as constituent structure markers in adult speech.

Categorial development proceeds concurrently with the development of mechanisms for expressing more complex functional relations. It is not possible to set up a milestone table to show the sequence in which major categorial distinctions are acquired, for there is rather little agreement between the categories hypothesized for different children by different investigators. A class of child words including most words that one would identify as adult nouns is rather quickly distinguished from a class that includes most of what one would identify as adult verbs. Further categorial discrimination is in the direction of the adult system; there is no evidence for the formation of any categorial distinction that is not related to an eventual adult distinction (apart from possibly the pivot-open distinction for those who wish, as is not done here, to call this a distinction of category; but even this distinguishes between a class containing adult nouns (the open class) and one that does not.)

3.2.4 Functional Relations

Characterization of functional relations is taken in this research to be crucial for an accurate statement of a child's linguistic competence. At each stage the child is assumed to possess a small set of such

relations which he will express in his speech and on the basis of which he will interpret the speech of others. To decide just what these might be appears, however, to be an extremely difficult task.

There is probably a stage at which the child has a single holophrastic linguistic concept that we might characterize as his "concrete referent". Most of the child's early one word sentences fit this concept, although occasionally there are apparent counterexamples. The example of "want" was already mentioned; another example would be "hot". Although "hot" refers to a property for adults, for the child it might well refer to an entire situation.

The next stage would involve a fractionation of the previously unitary concept. There might now be two concepts, one indicating the referent and another indicating some qualification or modification of that referent. For example, the child might learn the relationship of possession. A referent would no longer be viewed unitarily but rather as something possessed by something else, and so there are sentences like "My car", "My mommy", and "Daddy coffee" (with the meaning "Daddy's coffee", not "Daddy is drinking coffee".) Perhaps the relationship of attribution is also

learned so that referents are viewed as having properties; sentences like "Pretty boat" would follow. And there are probably other relationships of referent modification learned about the same time.

Some stages of development later, the child differentiates between different subjects performing an action, and the functional relation "subject of a sentence" is now expressed in sentences like "Airplane all gone" and "Calico all gone".

This view of language development suggests research areas that deserve exploration: (1) To develop techniques for interpreting what functional relations are expressed in child speech; (2) To establish how the set of functional relations present in early speech develops over time, it probably being the case that some relations are fractionated, while others shift in meaning, and still others are dropped altogether; (3) To establish invariant developmental sequences over languages in the event that there are any.

Gruber's "Topicalization in Child Language" (1967) is one recent study where functional relationships have been considered seriously. His concern is with the evidence for a particular functional relationship (the

topic-comment relation) in the linguistic competence of a child; the detail of his study indicates rather well the degree to which evidence about functional relationships is inferential and indirect.

CHAPTER IV

THE PROCESS OF LANGUAGE ACQUISITION

This chapter will review three theories of language acquisition (those of Braine, Jenkins and Palermo, and McNeill), and will then consider again the data from which children build their language. There are others who have contributed to the development of acquisition theory, but I have chosen to review these four people because together their proposals encompass much of the field. The theories differ fundamentally in what they assume the child brings to the task of language acquisition.¹ There is on the one hand the associative theorizing of Jenkins and Palermo, in which the basic operations of the child are response and stimulus generalizations. On the other hand there is the very strongly nativistic position of McNeill, in which the child possesses a substantial initial set of formal and substantive linguistic universals. Braine is somewhere in between, or so it appears from his more recent comments about the need for

¹Throughout this paper the term "initial" will be used to refer to a structure a child brings to the task of language acquisition. Whether the structure is innate or acquired by means of extra-linguistic experience will not be considered here.

an initial characterization of the properties of sentences relevant in language acquisition (Braine 1966b).

4.1 Braine - Contextual Generalization

Braine has proposed the theory of contextual generalization as an explanation for part of syntactic acquisition. The basic idea is that a child learns that an item occurs in a certain position by experiencing the item in that position in some set of contexts. The generalization occurs when the child places that item in that position in some different set of contexts. Sequences of units can themselves form higher level units whose position will be learned in the same way. An absolute definition of position appears to be, on the basis of his experiments, easier to learn than relative position; but absolute position need not be stated strictly in terms of first or last position in the sentence, but may be stated relative to such things as intonation markers or closed-class morphemes.

Some objections have been raised to Braine's claim that contextual generalization, when supplemented by the learning of positional contingencies, can explain the acquisition of what used to be called kernel sentences (a large group of sentences in the language with a constant

order.) One objection is that the theory depends upon there being a dominant word order, the point being that if a word does not occur predominantly in the same position, then that position will not be learned. And this constancy of word order has been questioned, as it is apparently not the case that the verbal environment of a child consists mainly of simple declarative sentences (Bever, Fodor and Wexsel 1965a). Furthermore, even in a language with relatively free word order (Russian), the early constructions of children are ordered, a fact which Slobin (1966a: 134-5) explains as a predisposition by the child to learn ordered sequences of categories irrespective of whether there is in fact a dominant order in the language. At any rate the learning of positional constancies is apparently a fact: the question of whether this is simply a generalization phenomenon or whether it reflects an initial syntactic constraint on the types of constructions that can be acquired needs more study.

Another objection to Braine's theory, and one with more far-reaching implications, is based on the linguistic conjecture that what the child acquires are structures for which there are no models at all in adult speech. Bever, Fodor, and Wexsel provide the examples:

(1) the abstract nature of the structure that underlies simple declaratives and other transformed sentences, the number of the verb being unspecified until after application of some transformations; and (2) by positing an underlying order for elements of the auxiliary different than the order ever encountered in actual speech (the correct order being achieved by an affix movement transformation) one simplifies a description of English.

To say that what the child acquires is a structure (or structures) for which there is no model in the language data he is exposed to can be confusing. In one sense there is no model in the language data for any part of a grammar unless the grammar is simply a listing of sentences. All mental constructs (rules, features, semantic markers) are inventions and not overtly present in the data; so this clearly is not what is meant by the absence of a model for some structure. In a discussion of this point in an article directed to some criticism of his work by Gough and Segal, Braine seemed to be saying that a rule (a mental construct) was exemplified in the language data if there was at least one sentence in that data such that in the derivation of that sentence the rule was used. Braine writes, "Even for the child learning the natural

language, it is hard to see that rules could be learned unless they were exemplified in the input, e.g. presumably a child could not learn the passive transform without ever having had any commerce with a passive sentence." (See Braine 1965c:242.) Apparently, however, Braine is not basing his rejection of the Bever, Fodor and Weksel examples on this quite reasonable definition of "exemplification in the data", for in both of their examples the structures do have models in the language data in the sense just explicated; namely that one can, using these structures, derive grammatical sentences of English.

It is not clear what Braine now means by there being a model in the language for some construct. Why is it all right for there to be an underlying order for passive sentences that is never exemplified by those sentences, and not all right for this to be true of simple declarative sentences? Braine's answer would likely hinge on the claim that there are exemplifications of the structure underlying passives (namely simple declarative sentences) and no such exemplifications of some putative abstract structure underlying simple declaratives. But as already noted, equating exemplification of a structure with overt occurrence of that structure fails immediately.

The issue between Braine and his critics seems to be that Braine is apparently willing to allow for the acquisition of only those mental constructs that stand in a fairly direct relationship to observed data, while transformational grammarians would allow for a less direct relationship between the acquired grammar and the data upon which it is based.

There appear to be no reasons a priori for requiring this relationship to be direct. Consider the following example: If an adult is presented with the data

2 , 6

4 , 12

6 , 18

5 , 15

and asked what system underlies these, he would probably say something like, "There is a number followed by a comma followed by three times the first number." The answer is correct, but involves two concepts (the number three and multiplication) for which there is no overt exemplification in the data. What has happened, of course, is that the adult is primed to recognize numeric patterns using the arithmetic operations plus, minus,

multiplication, and so on. But why should it not also be possible that a child is constrained to discover relationships of an equally subtle sort in the language data he is exposed to?

The point is that requiring a priori a fairly direct relationship between the grammar acquired and the data on which it is based is not a sound theoretical assumption. This is not to assert that current linguistic theory has made a strong case for such things as abstract base structures for all sentences, but the theorizing is at least interesting and suggestive. On the other hand Braine (1965a) is quite correct in his argument that considerations of parsimony should be applied relative to both a grammar and the acquisition mechanism, not just one alone.

The theory of contextual generalization has been faulted by McNeill (1966d:29) because it does not account for the supposedly nonrandom membership of the pivot class. In Braine's theory the pivot class is just that set (or sets) of words whose position has been learned, but this would likely mean that the pivot class has members drawn indiscriminately from adult categories. This conclusion Braine would likely not contest, but its factual falsity is

important to McNeill's acquisition theory (see 4.3 below.) As already suggested, there are reasons not to view the pivot class as a psychologically unified class at all, but if one wants to categorize the pivot elements in terms of adult categories, a diversity of categories would appear to be the rule: Gregory (Braine 1963b:4) had one or more pronouns, verbs, adjectives, and greetings as pivot elements; Andrew (Braine 1963b:5) had one or more prepositions, pronouns, verbs, adjectives, adverbs, greetings, and prepositional phrases. The example of relatively homogeneous pivot classes listed by McNeill (1966d:22, in particular Ervin's example) appears to be beyond the early stages of language learning.

To summarize these comments on Braine, the learning of positional invariances appears to be important in early stages of language acquisition, although whether this is simply a generalization phenomenon or whether it reflects initial constraints on the types of constructions that can be learned is open. Nonetheless, since such positional invariances are really only an approximation to the structure of natural language, it is likely that the role of such learning diminishes as acquisition proceeds. The learning of positional contingencies, on the other

hand, while another property of language that the child can recognize, does reflect (in part) the structure of natural language, and so this type of learning presumably continues to play a role throughout language acquisition.

4.2 Jenkins and Palermo - Mediation Theory

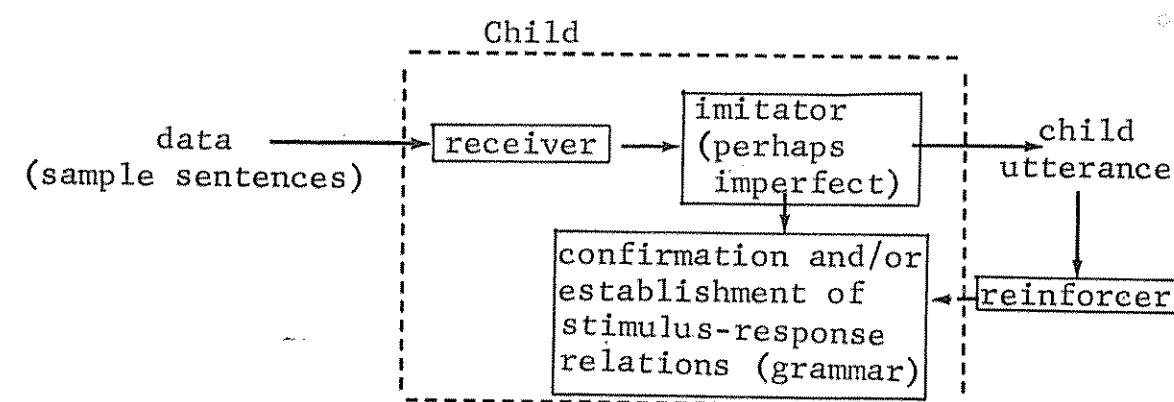
Jenkins and Palermo have outlined a mediation theory to account for language acquisition. Several of the points they make will be considered, but first a difference between mediation theory as applied by Jenkins and Palermo and single-stage reinforcement theory as developed and applied to language behavior by Skinner (1957) should be noted. The difference is a tremendously important one: Skinner is concerned with explaining linguistic behavior, while Jenkins and Palermo are not (apart from some comments on why children imitate); rather, they are concerned with explaining the acquisition of a linguistic system (a grammar). The latter is certainly a more reasonable goal than the former at the present time, and makes the theorizing of Jenkins and Palermo immune to much of the criticism directed toward Skinner (see Chomsky 1959).

The proposal of Jenkins and Palermo will be discussed in terms of what they say about (1) imitation,

(2) development of categories, and (3) development of constituent structure.

4.2.1 Imitation

The child's imitation of speech is taken to be central to language acquisition. The relationship between the verbal environment of a child, his own speech attempts, and his acquired language competence, may be loosely diagrammed as follows (my diagram, not Jenkins and Palermo's):



The important feature about this diagram is that hearing a sentence has no effect on the child's grammar (the network of s-r relations is here called a grammar) unless that sentence is imitated by the child (presumably the imitation need not be externalized but can remain covert and yet be effective; see the Jenkins and Palermo analysis of the Esper experiment.) But this is rather curious--why must an utterance be imitated (perhaps imperfectly) before it can

have any effect on the organism? The answer is that the theory is concerned with the establishment and strengthening of s-r bonds (where both the stimulus and response may be internal events according to mediation theory), and although a stimulus may be something done either by the person involved or by a completely external agency, the response must be made by the learner himself. Thus, to explain language acquisition by this theory, responses by the language learner to stimuli are necessary. Well what, in terms of language, does a child do? He listens and he talks. And since listening (including comprehension) is a somewhat mysterious operation, it apparently was deemed sounder to suppose that it was the child's speech which served as the response in terms of which s-r relations are acquired. This would make application of reinforcement principles straightforward; people would respond differentially to the child's speech efforts, which would provide the required contingent reinforcement.

Historically this concern with imitation rather than listening may have been due in large part to the possibilities for using the production of a speech event as a response to be reinforced, whereas what the response might be in listening was just not known. This

gain, however, in the discreteness and tangibility of behavioral units, is achieved only at the cost of assuming imitation of all instances of language data that yield new information about language structure. It seems that psychological theorizing has been simplified by the imposition of a very strong assumption regarding the child's behavior, an assumption for which there is no independent evidence. Ervin compared the spontaneous imitations of children to their normal speech patterns and found that the imitations were not grammatically progressive. She concludes "... there is not a shred of evidence supporting a view that progress toward adult norms of grammar arises merely from practice in overt imitation of adult sentences."² (Ervin 1964:172) Just why children imitate is an interesting but peripheral question here; for some discussion see J. Hunt and the references he cites (1961:footnote 11 on p. 144).

²There is a type of imitation which is grammatically progressive, and that is the child's imitation of adult expansions of the child's speech (Slobin 1964:32). However a theory that made this the principle mechanism for language acquisition would have difficulty in explaining the language learning of children who grow up in social environments where there is little or no verbal interaction with adults.

This emphasis on the role of imitation, besides being unsupported and requiring imagination to apply in, for example, Lenneberg's case study, where all imitations would have to be internal, is apparently unnecessary. An explanation that does not depend upon covert or overt imitations would make the process of comprehension, not imitation, that which provides the responses to stimuli.³ When a child hears a sentence he takes certain actions to understand it, and when his understanding, based upon a certain set of s-r relations, is correct, one can suppose that that set of relations is strengthened. By correct understanding is meant that the child's interpretation of the sentence taken together with all that he knows about the world is such as not to lead him to reject his putative interpretation of that sentence. If a child hears "Daddy ate his dinner", his interpretation of that sentence would presumably be consistent with the rest of his knowledge if he assigned approximately correct references to the

³The argument that places the emphasis on learning via comprehension as opposed to production is apparently a special case of the more general psychological issue as to whether skill acquisition depends upon perceptual experience or motor experience. For a survey of some conflicting evidence see J. Hunt (1961:99ff.).

content words, interpreted "Daddy" as the subject and "dinner" as the object, and so on. If he tried to interpret the sentence in any of an infinite number of incorrect ways, say by taking "dinner" as the subject and "Daddy" as the object, his comprehension would not be consistent with the rest of his knowledge and he would be led to reject that interpretation (see Chapter V.) Differential reinforcement of correct responses would also be explained in terms of the reinforcing consequences of correct understanding: understanding a sentence might lead him to some action which would be reinforcing, and understanding itself would soon become secondarily reinforcing.⁴

This discussion of imitation has not attempted to show that a mediation-theory account of language acquisition is possible or impossible, but rather has suggested

⁴No brief is held for the looseness with which the term "reinforcing" is being used. Chomsky (1959) has correctly criticized this type of theorizing (using a fairly well defined technical term and extending its application to an extremely ill-defined domain, and assuming that laws stateable using the original term will still apply.) The purpose of the discussion is only to show that if one wants to talk this way he can as well talk about reinforcement in terms of comprehension as in terms of speech production.

that any theory should place the burden of language learning on the process of speech comprehension rather than speech production (and, in particular, imitation). The centrality of comprehension in language acquisition is in fact an important feature of the model described in Chapter V.

4.2.2 Development of Categories

In the view of Jenkins and Palermo, development of categories proceeds by the process of stimulus or response generalization. They describe some paradigms which, when applied to categorial acquisition, explain how words experienced in the same contexts will tend to elicit each other; how, in other words, they will form a unified class. Looking for common distributions of elements to ascertain categorial membership is of course one of the principles of distributional linguistic analysis. It is also assumed that the class will acquire a semantic distinctiveness because of the semantic commonality of the class members.

That categories are defined by experiencing items in the same contexts has an apparent simplicity that makes the notion seem plausible. However, there is a problem, which is that class membership based on a distributional analysis appears to require the storage of

extremely large amounts of data, with concomitant retrieval problems. There is no suggestion in the theory that the child discovers diagnostic contexts and assigns them particular value; rather, all contexts are treated alike, and so the child has no basis for being selective in the contexts he chooses to retain. Perhaps semantic considerations will, in ways not yet known, narrow the storage and retrieval problem. One should also look at Braine (1966b), who raises some questions regarding the experimental verification of the generalization phenomena.

4.2.3 Development of Constituent Structure

Jenkins and Palermo credit the child with the ability to learn sequences where the sequence itself may be abstract. For example, continued exposure to noun + verb + noun sequences, even though the specific nouns and verbs are different from sentence to sentence, will result in the acquisition of this pattern. Although the acquisition of constituent hierarchies is not discussed, the presumable explanation for such major categories as NP and VP would also be in terms of response or stimulus generalization (e.g. a Det + N sequence and an N may both occur in similar contexts and so together they would form a class which one can call NP.) Although this explanation is somewhat

plausible, the problem of the large amount of data needed to support the distributional analysis recurs here as well. Another difficulty is that the theory provides no account of the recursive nature of natural language. Clearly the child does not learn some finite number of patterns, say for adjectival position, of the form:

| | |
|---------------------|----------------------|
| Det + N | "The man" |
| Det + Adj + N | "The old man" |
| Det + Adj + Adj + N | "The strong old man" |
| . . . | |

Rather he has learned a recursive rule that allows for an unbounded number of adjectives in that position. (Other recursive processes are conjunction and relativization.)

In summary, the concern of Jenkins and Palermo with explaining the acquisition of language structure rather than attempting to explain language behavior is to be welcomed. Although their emphasis on the role of imitation in language acquisition is not accepted, perhaps their learning paradigms can be applied to s-r links established in sentence comprehension. At any rate, they appear to have shown how a type of distributional analysis of language data can be accomplished by means of a particular type of associative learning. Their analysis is not strictly

formal, however, say of the type described by Harris (1951), as they allow for semantic correlates of classes and constituents. Although distributional analysis probably plays a role in language acquisition, it is of only peripheral interest in the model described in Chapter V; the semantic correlates of classes are, on the other hand, very relevant in the model.

4.3 McNeill - Nativism

McNeill is one of several students of generative grammar who have taken their direction from the rationalistic framework for language acquisition proposed by Chomsky (1965:47ff.). McNeill's proposal will be discussed in terms of the topics (1) categorial development, (2) functional relations, and (3) hypothesis testing.

4.3.1 Categorial Development

4.3.1.1 The Proposal

McNeill (1966d) has argued that there is a very strong initial⁵ constraint on the development of

⁵McNeill's discussion dealt with innate abilities; however, for ease in integrating his ideas with those incorporated in the model of Chapter V, the term "initial" will be used.

grammatical categories. A child is presumed to possess an initial categorial hierarchy which would "direct the child's discovery of the classes of English" (and the classes of other languages as well.) At the top of the hierarchy two classes might be distinguished, and so a child would be led to discover those two classes in the speech sample he is exposed to (if indeed those classes are exemplified in that sample, but not all hierarchical distinctions must be evidenced in any particular natural language.) The next lower level in the hierarchy would direct the child's differentiation of his previously acquired classes, and so on. According to this proposal, therefore, at or near the top of the hierarchy there is a distinction between what McNeill calls the pivot and open classes; differentiation of these classes proceeds in accord with a lower level in the hierarchy.

4.3.1.2 Comments

(1) The theory would appear to require that all children learning the same language under conditions of comparable language exposure make approximately the same categorial distinctions in the same sequence, but this does not appear to be factually true. McNeill explains this as a consequence of somewhat different starting

points in the categorial hierarchy; not all children need start at exactly the same place and follow exactly the same sequence, the hierarchical sequence being somewhat indeterminate.

(2) One way of justifying this proposal is indirect; namely that it provides an explanation for categorial development which proponents of a distributional analytic scheme may not be able to explain. As McNeill notes, there are indefinitely many possible bases for making categorial distinctions, and how is the child to know which one(s) to use? In other words, how will he discover diagnostic frames for analysis?

(3) Another possible justification of this proposal is that it explains the rather controversial claim of McNeill's that the categorial distinctions of the child will "honor" the eventual adult distinctions. That is to say, at no stage should words that will eventually end up in one adult class be members of two or more classes of the child. This claim has already been mentioned (see Chapter III) with tacit approval as a correct statement about categorial development beyond the very earliest stages. For example, once a "noun-like" category is formed its differentiation will be in the direction of the

eventual adult distinctions (count nouns, proper nouns, etc.); a child will not form a noun subclass that bears no relation to any eventual adult subclass. The problem comes, apparently, with the application of this principle to the very earliest stage of categorial development. As was pointed out in the discussion following presentation of his paper (McNeill 1966d), the original pivot and open class distinction does not honor all eventual adult distinctions (e.g. adult adjectives are members of both the pivot and open class of Gregory (Braine 1963b:4), and adult verbs are members of both the pivot and open class of Steven (Braine 1963b:7).) Perhaps this should count against supposing the pivot elements to actually constitute a class; it might be that the initial categorial hierarchy concept could be applied beginning with the formation of the open class. McNeill has recognized the factual difficulty in maintaining that the child's categories are related generically to the eventual adult categories: the extent of cross-classification is under investigation (McNeill 1966a).

(4) One of the difficulties with this proposal, and indeed with much of the rationalistic theorizing about language acquisition, is to decide what

role a putative universal actually plays. Perhaps all that can presently be done, from the viewpoint of contemporary linguistic theory, is to identify formal and substantive universals and assume that they in some manner or other constrain the acquisition of language. The problem of how they might do so is, perhaps, given the present state of knowledge, simply unanswerable.

However, it would be desirable to interpret how a claim about linguistic universals might be integrated into a model of language acquisition. For example, what implications might there be for an acquisition theory in the claim that the initial categorial hierarchy "directs the child's discovery of the classes of English", or "...a child classifies the random specimens of adult speech he encounters according to universal categories that the speech exemplifies."? (McNeill 1966d:36) Some speculations about the type of information such a categorial hierarchy might contain and how it might be used will be considered directly.⁶

⁶The remainder of this discussion on the categorial hierarchy contains speculations not directly attributable to McNeill.

(a) The possibility that the categorial hierarchy contains information that would make possible a correct distributional analysis (for example, diagnostic frames for the parts of speech) is not reasonable, as such data are language specific.

(b) Perhaps an element in the hierarchy contains information about the semantic correlates of a class when classes have such semantic properties. This hypothesis depends upon the lexical categories having such semantic invariances across languages. The problem of identifying semantic correlates of adult classes is well known, but if the child's categories are only approximations to the adult classes then they may well have semantic invariances (cf. Brown 1957). This hypothesis, however, is difficult to distinguish from the Jenkins and Palermo theory, or at least that part of their theory that deals with categorial formation mediated semantically. In both cases the child is attuned to recognize certain semantic properties of words⁷ and to form classes on the basis of shared semantic features.

⁷Although Jenkins and Palermo do not state this explicitly, it is clearly necessary that the child recognize only a finite subset of the semantic properties associated

(c) A third hypothesis, which is entirely consistent with the second one, is that an element of the categorial hierarchy contains certain definitions of formal properties of a category. For example, the category noun can be defined as that class of words in a grammar which have inherent properties in contrast to the other lexical categories whose features are determined by selectional rules. Other formal properties of lexical categories and major categories such as NP and VP have been tentatively suggested (Chomsky 1965:115ff.). The problem in utilizing these formal definitions in language acquisition is that they themselves are defined relative to a grammar (Chomsky 1965:footnote 2, pp. 208-210). If one had a putative grammar one could use these formal properties to identify the lexical and major categories and could then go on to test the grammar (by seeing if it correctly identifies the universal grammatical relationships, etc.), but it is not clear how one could use these properties in grammar construction. (For a discussion of hypothesis testing see 4.3.3.)

with a word drawn from the infinite set of such properties that is possible. Those that are recognized would no doubt be called the "salient" features.

4.3.2 Functional Relations

4.3.2.1 The Proposal

McNeill has proposed that the basic functional relations such as subject of a sentence, object of a predicate phrase, main verb of a predicate phrase, head noun of a noun phrase, modifier of a noun phrase, and so on, are initially defined for the child and constrain or direct his language acquisition. The child is presumed to have a definition of a functional relation in terms of universal lexical or major categories and a deep structure configuration. For example, the definition of the sentence subject might be "the NP directly dominated by the sentence." The theory has certain obvious entailments: (1) Lexical and major categories must themselves be initially defined and known to the child by the time he uses his initial characterization of functional relationships; (2) the child must somehow distinguish between base structures and transformed structures, for the definitions of functional relations will apply only to the former. This would follow automatically if it were the case that at early stages a child's competence is pretransformational (as McNeill suggests, perhaps "it is not too unreasonable to think of children 'talking' base structures directly" (McNeill 1966d:51)).

4.3.2.2 Comments

(1) A functional relation is first of all a semantic relation. Although the relationship is defined syntactically, it plays no role in the generation of deep or surface structures; it is used only in the semantic interpretation of a sentence.⁸ An initial definition of functional relations might then involve specification of the semantic consequences of particular relations. That is, for example, to say that subject of a sentence is an initial grammatical relationship might be to say that the child knows the implications of calling something a subject for semantic interpretation (loosely this would at least mean that the child knew something like the equivalence "subject = performer of the action of the sentence".)

(2) A functional relationship is defined formally in terms of lexical and major categories

⁸ Just how specification of functional relationships contributes to understanding has been little investigated; one of the few concrete but programmatic suggestions that have been made is that by Katz and Fodor (1963) in which they distinguish several types of projection rule, the one operating dependent upon the functional relationship between the constituents.

and the relationship of direct domination. For the child to have an initial definition of a functional relationship, then, might be for him to know something about the constituent structure of base P markers. For example, if the child possessed the definitions "object of a predicate phrase = the noun phrase directly dominated by the predicate phrase", and "main verb of a predicate phrase = the verb directly dominated by the predicate phrase", he would then know that the base constituent structure was either:



And, in general, full knowledge of the functional relationships would define the base structure rules apart from the order of constituents, the latter information being language specific.

(3) As McNeill notes, the assumption of an initial characterization of functional relations explains the acquisition of hierarchical structure by the child. This is a crucial characteristic of language learning and one not easy to explain just because hierarchical structure is not marked in the speech sample the child is exposed to; perhaps units are indicated to some degree by suprasegmental features, but hierarchical structure is not.

(4) To support his hypothesis that functional relations are part of the initial equipment a child brings to language acquisition, McNeill classified two and three word sentences (produced by a child described by Brown) according to their categorial patterns. As the child supposedly had three categories at this time (nouns, verbs, and pivots) there would be $2^3 = 8$ possible two word sentence patterns and $3^3 = 27$ possible three word sentence patterns. Not all possible sentence patterns, however, would be consistent with the presumed initial functional relationships (when one interprets the categorization of noun phrase modifier as a noun or pivot immediately dominated by NP, there being no class of determiners at this stage.) In fact there are only five possible two word sentence patterns and eleven possible three word sentence patterns that are consistent with the putative functional relationships.⁹ Most of these are exemplified in the data, while, according to McNeill, none of the patterns that are

⁹McNeill claims that the figures are four and eight respectively, but he appears to have overlooked the patterns where modifiers can follow head nouns: N + P (head noun, modifier), N + P + V (object (head noun, modifier), main verb), V + N + P (main verb, object (head noun, modifier)), and N + N + P (predicate, subject (head noun, modifier)).

inconsistent with the functional relationships are exemplified. This observation is of some interest, for it suggests that (i) the child may be speaking base structures directly, since all of his sentences are potentially consistent with the universal functional relationships,¹⁰ and (ii) the child does not at this stage acquire the transformed sentence patterns that he encounters. (McNeill's example of the latter is the pattern V + V + N with an adult model "Come and eat your Pabulum" (McNeill 1966d:48).)

However, there appears to be a problem in that some of the decisions regarding the part of speech classification done by the investigator are made just so that the sentences will fall into a general pattern. This point came up in the discussion following McNeill's paper (1966d) with reference to the sentence "Want do" where both "want" and "do" can occur as pivot words. Fodor

¹⁰However, it does not follow from this observation that the child is speaking only base structures, for many transformed sentences will have a surface categorial sequence consistent with a possible deep structure categorial sequence.

commented that "do" had to be classified as both a pivot and an open word just because two pivots cannot occur in the same sentence. But of course it then would become somewhat circular to draw inferences based on the absence of sentences such as pivot + pivot.¹¹

The sentence patterns that are not consistent with the universal functional relationships are, for two word sentences, any pattern not having a noun. For three word sentences it is any pattern having two or more pivots, or two or more verbs, or the patterns N + V + P and P + V + N. Perhaps not many of these would have adult models, and yet some would, and it is the absence of these that McNeill would explain by his assumption of initially defined functional relationships. What the child is presumed to do is to analyze and form sentences on the basis of his initial functional definitions.

(5) It should be noted that there is a relationship between McNeill's two proposals, the first one on categorial development and the second one dealing with functional relationships. Both of them presume an

¹¹This particular comment referred to Braine; to what extent Brown's analysis, upon which McNeill's conclusions are based, shares this alleged investigatorial bias I do not know.

initial characterization of lexical and major categories. Consequently it seems a little strange that a common reaction to McNeill's paper was to accept his hypothesis about initial functional relationships and yet to reject any initial substantive specification of lexical and major categories (cf. discussion following McNeill 1966d). This does not appear to be consistent, although perhaps what was meant was only a rejection of the specific sequential or hierarchical aspects of the category hypothesis, not any formal or semantic characterization itself.

4.3.3 Hypothesis Testing

The generative framework for language acquisition, as developed by Chomsky, McNeill and others, assumes that the child brings to the task of language acquisition a complex initial characterization of formal and substantive linguistic universals. This initial structure so heavily constrains what constitutes a potential grammar of a natural language that language acquisition is described, somewhat confusingly, as the largely independent creation of a grammar which is subsequently evaluated by reference to the language data, rather than the induction of a grammar given that data. This way of talking about language acquisition is intended to emphasize the degree to

which the resulting grammar is a product of the initial structure rather than the language data, although there is also an emphasis on the activeness of evaluation as opposed to the passive establishment of associative networks. At any rate, in the event that there is more than one grammar consistent with the data, the child will evaluate the alternatives in terms of an initial simplicity metric (which need not, however, accord at all with our intuitions about simplicity) and select the most highly valued grammar.

The difference between the generative and traditional psychological frameworks is reflected in what are taken to be reasonable research goals. If generative grammarians are right in the assertion that there are very many formal and substantive universals in natural language, then speculation about the process of language acquisition which minimizes or eliminates the role of such universals will probably be beside the point. Linguistic analysis of different languages is more likely to lead to insight into linguistic universals, and derivatively into the mechanism for language acquisition, than attempting to describe a learning or induction mechanism that will, given linguistic data, produce a grammar that is not minimally consistent with the alleged linguistic universals.

It is too early to be confident about the outcome of this debate. It may eventually be settled when linguistic theory has advanced far enough for there to be a reasonable consensus on what linguistic universals are. As this comes about, learning theorists will either have to revise their theories to incorporate these universals, or else they will need to show that the universals are only accidental in the sense of being automatic consequences of more basic acquisition principles. Certainly it would be wrong to reject an acquisition model with a large initial component on a priori grounds; it seems that the only legitimate criticism of such a model would be that the initial component was based on alleged universals that were not in fact universal or else only accidentally so.

Braine (1966b) has raised an objection to the general framework of hypothesis testing on the grounds that it appears to require the child to have available data about nonsentences, which Braine would suppose to be unimportant. He argues that if the child has only instances of grammatical sentences, then should not the child end by selecting a very general grammar that provided correct analyses for the data he has and which, because of their generality, are more highly valued (being less ad-hoc)

than what one would presume to be the "correct" grammar? This very general grammar would fail to the extent to which it identified ungrammatical sentences as belonging to the language, but, according to this argument, this type of failure should not be discernible in the absence of negative instances.

Braine's objection would be largely overcome if the child's initial structure was sufficiently detailed so that the class of possible grammars excluded any unreasonably promiscuous (overgeneral) grammars. And this is in fact where generative theoreticians suppose the research problem to lie, in the specification of the narrowest class of possible grammars which the child would have to consider (cf. Chomsky 1965:35). Furthermore, although nonsentences may not be important in language acquisition, there may well be a type of negative data available to the child; namely the signal that a tentative analysis he has made of a sentence is inconsistent with what he knows about the world. This signal indicates that a tentative analysis is (supposedly) incorrect, which information could be used to avoid overgenerality in his grammar (see Chapter V).

Braine also objects to the notion of hypothesis testing on the grounds that it would appear to require

storage of large amounts of data, as a hypothesis would not in general be confirmed or disconfirmed by a single example. For each hypothesis, therefore, something like a value reflecting the number of positive and negative instances would be required. There would also be the problem of deciding when sufficient evidence had been obtained to accept or reject a hypothesis. While all of these observations appear correct, it seems that they would also apply to Braine's (1966b) model of a sentence scanning device and a series of intermediate stores. On the basis of the little description available, it appears that Braine's intermediate memory stores serve exactly the same function as status indicators for a hypothesis; which memory store a property is entered in reflects the extent to which it has been confirmed. Even in a very simple associative theory one presumably needs a mechanism for registering the strength of s-r bonds to reflect the amount of evidence for those bonds.

4.3.4 Summary

McNeill has suggested several components of what he considers to be the child's initial mechanism for language acquisition. The notion of a categorial

hierarchy, involving a set of substantive universals, seems to have certain counterfactual implications and may need revision; while the hypothesis of an initial characterization of functional relationships has some supporting evidence. Although the case for an initial formal characterization of the lexical and major classes, upon which the two substantive universals of McNeill depend, may ultimately be confirmed, this has not been made a part of the model described in Chapter V. That the child has an initial semantic characterization of the functional relationships and of the lexical categories, which he uses in language acquisition as the basis for construction of grammatical hypotheses, is, however, a feature of the model.

4.4 Implications of Unacceptable Input Data

The fact that the speech which a child is exposed to is, to a significant extent, ungrammatical, has often been noted (Bever, Fodor, and Weksel 1965a; Chomsky 1965: footnote 14 p. 200-1; Braine 1966b). However, from this fact different conclusions are drawn. Generative grammarians suppose that it supports a strongly nativistic view of language acquisition, for in spite of the ungrammatical data the child does not fail to discover the correct grammar. Braine (1966b), on the other hand, has argued

that this infirms any hypothesis testing scheme, such a scheme being muddled by the ungrammatical data. The point seems to be this: if one presumes a mechanism that in effect processes all input data uniformly (whether grammatical or ungrammatical), then the sensitivity of the device must be much less than would be the case if the data were entirely grammatical.

Although the possible consequences of ungrammatical input data have been often considered, very little attention has been paid to what is a more serious problem, of which the other is simply a special case. If one assumes an acquisition scheme where the child builds increasingly more complex structures from simpler structures, then it would seem that any grammatical sentence that was more complex than what the child was prepared to deal with would be just as bad for him as an ungrammatical sentence. On this view, what the child wants are instances of data that will either serve to confirm (or infirm) previously acquired constructs or are examples that will bear on the next step in acquisition, but not evidence for constructs which the child will not be ready to acquire until much later. What one should be concerned about, therefore, is not just the problem of

ungrammatical data but also the problem of grammatical data that are too far advanced for the child to be able to use.

The conclusion drawn here from the abundant presence of unacceptable input data (by which is now meant ungrammatical and overly complex grammatical sentences) is that the child probably distinguishes, at least to a considerable extent, between acceptable and unacceptable data prior to the use of such data by his acquisition mechanism. This might sound somewhat mystical, as unacceptable input data will not usually be identified overtly as such when produced; consequently how could the child make this distinction? Two mechanisms are suggested rather tentatively: a signaling mechanism and the mechanism of attempted comprehension.

A signaling mechanism is some fact about the sentence or contextual situation which tells the child, so to speak, not to bother attending to the sentence or trying to understand it. One such signaling mechanism appears to be words which are unfamiliar. On the basis of responses by children 1 1/2 to 2 1/2 years old to the presentation of differently structured imperative sentences (C. Smith 1966), it was

inferred that children "tend not to listen to adult speech beginning with unfamiliar words." And the same type of signal given by unfamiliar words is sometimes given by shifting the tone of one's voice from that usually used in speaking to a small child to one's normal tone (for those adults who indeed make a tonal distinction); the meaning of this signal is something like, "this speech is no longer directed to you".

Even if the child attends to all the words of a sentence, however, the mechanism of attempted comprehension might indicate unacceptability. Suppose that the grammatical knowledge of the child is not adequate to support a correct analysis of the sentence even though some or all of the words in the sentence are understood. Two situations can be distinguished: (1) where the sentence he has heard consists of an uninterpretable sequence of categories; and (2) where the child's grammatical competence enables him to assign an interpretation to the sentence which is, however, wrong.

The case referred to in (1) is that where the child simply does not have any grammatical rules or hypotheses which would enable him to assign an interpretation

to a sentence, the words of which, however, he understands by themselves. An example might be McNeill's sentence, "Come and eat your Pabulum". If the child were to ignore the conjunction and possessive pronoun because of not understanding them, he would have a verb + verb + noun sentence where all three words were understood. But since, by hypothesis, he has no basis for understanding this sentence (no rules that describe verb + verb + noun sequences), the sentence would be unacceptable to him.

The case referred to in (2) is that where the child is able to assign an interpretation to the sentence on the basis of his grammatical knowledge, but where this interpretation, when produced, is known to be incorrect. An example might be the sentence "The dinner was eaten". If the child ignored the article and auxiliary the sentence would be a noun + verb sequence where both words were understood. Furthermore, since this is a common pattern (subject + intransitive verb), the child has (by hypothesis) rules for interpreting this sentence which, among other things, identify dinner as the subject and eat as the verb. But this semantic interpretation does not fit with what the child knows (extralinguistically) about the inanimacy of dinners and the animate subjects required to

eat. And thus, although there is a possible interpretation of the sentence, it is not a correct interpretation, and it is known not to be correct by the child on the basis of his knowledge of the world (see 5.2.8). The sentence is therefore unacceptable. Note, by the way, that the notion of correct interpretation is defined relative to the child's knowledge and not reality. And since his knowledge is limited, misunderstandings (sentence interpretations considered acceptable by the child which are, however, factually incorrect) will be frequent.

In both of the cited examples of (1) and (2) it was assumed that the child was able to ignore parts of a sentence which he did not understand (see 5.2.4.) This illustrates the important point that something less than the full sentence may be comprehended. This sentence reduction facility of the child is absolutely crucial in language learning, for if he were to reject as unacceptable any sentence not completely understood, language acquisition would never take place. Strictly speaking, therefore, both of the examples of (1) and (2) would be classified as unacceptable by the child because of words which were not understood. The point of the examples, however, was to show that the reduced sentences which the child was presumed to

consider ("Come eat Pablum", and "Dinner eaten") would also be considered unacceptable. It might be, of course, that a further reduction would make one or both of the sentences acceptable (e.g. the reduction of "Come and eat your Pablum" to "eat Pablum" might result in an acceptable analysis).

Acceptable sentences, therefore, are just those that the child is able to "correctly" interpret. When a sentence is correctly interpreted this will serve to confirm that part of his grammatical competence which formed the basis for his interpretation.

Although this explains how a child could confirm grammatical constructs, it does not explain how the child could make grammatical progress if he discards (regards as unacceptable) any sentence which he is not already able to interpret. To make grammatical progress requires, in this view, that the child acquire independently some hypothetical grammatical constructs which he would then test against the input data. These constructs have, it seems, two possible sources: (1) they may be initial possessions of the child; or (2) they may be suggested by the child's reaction to a sentence which he understands imperfectly. Slobin (discussion following McNeill 1966d: 89) expressed (2) as: "It seems to me more reasonable to

suppose that it is language that plays a role in drawing the child's attention to the possibility of dividing nouns on the basis of animation; or verbs on the basis of duration, or determinacy, or validity; or pronouns on the basis of social status, and the like." What the relative roles of (1) or (2) as sources for hypotheses might be is not known, although a reasonable conjecture would be that in early stages (1) is more important, simply because so very much of the adult sentence will not usually be understood during the child's early stages. When his language competence has developed far enough so that most of a sentence he hears is understood, then perhaps he will notice features of a sentence which are not explainable by his knowledge of the language, and this might stimulate him to theorize about the function(s) of the unknown features.

Note that the view outlined above of differential response by the child to acceptable or unacceptable input data necessarily carries with it the notion of hypothesis testing, although it leaves open the question of the origin of hypotheses.

Within some domains a person can successfully employ a learning strategy in which there is no differential response to grammatical or ungrammatical data of the type

just described; all data are processed uniformly. This was shown by an experiment of Braine's (1966b) involving the learning of an artificial language by presentation of data which were ungrammatical about 8% of the time. These ungrammatical sentences were randomly mixed with formally indistinguishable grammatical sentences, and the resulting collection presented aurally to subjects. The learning of the language (not the presented set of grammatical and ungrammatical sentences) by this group was not significantly different from that of a control group which was exposed only to grammatical sentences. Since the language was meaningless and the ungrammatical sentences were not formally marked, the subjects could not have responded differentially to the two types of data in the way suggested previously; nonetheless learning still took place. Braine's reasonable explanation was that the subjects were not retaining unsystematic patterns (those outside the language). There appear to be no implications of this result for the differential response hypothesis, however.

The hypothesis outlined for the differential response by the child to acceptable and unacceptable input data has been implemented in the model to be described in

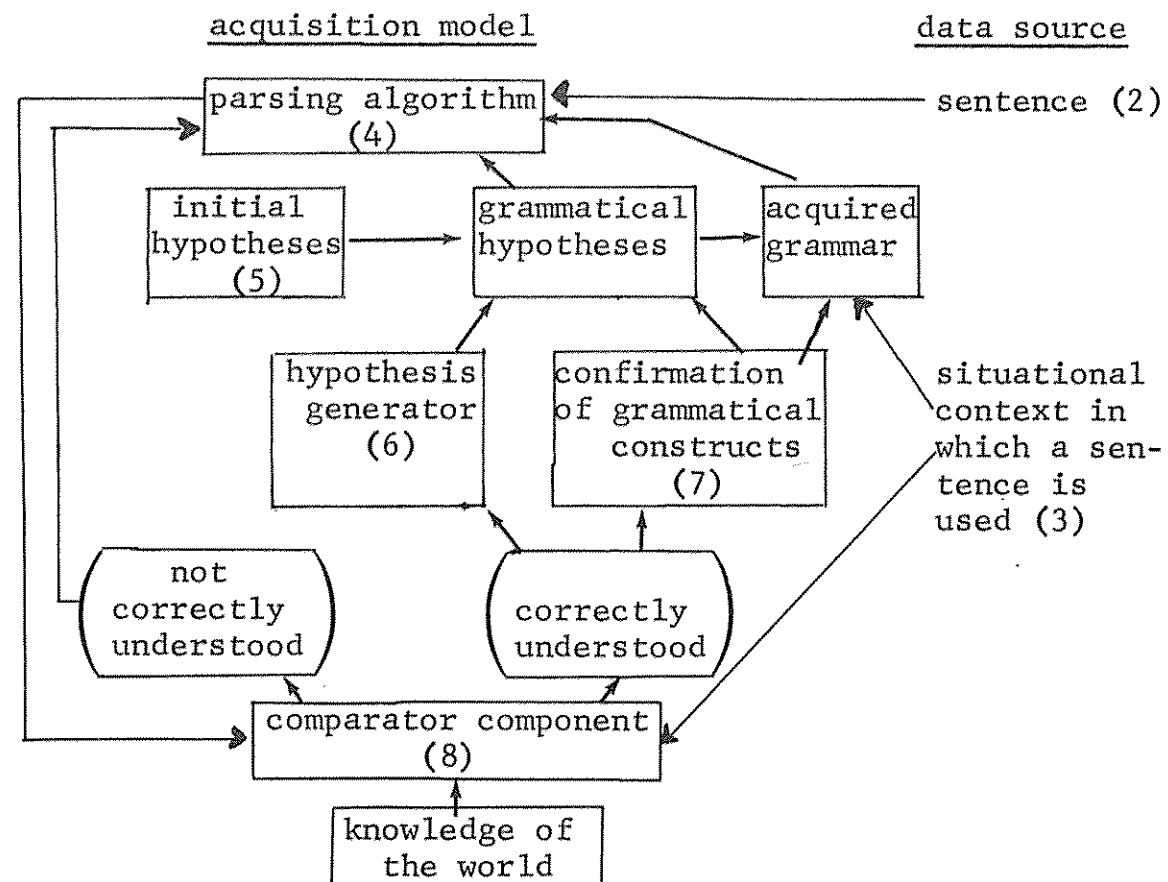
Chapter V. (No signaling mechanisms were included, however, only the mechanism of attempted comprehension.) An experiment was run in which about 20% of the sentences in the input data were ungrammatical, and there was no significant effect on language acquisition (see 5.3).

CHAPTER V

A MODEL OF LANGUAGE ACQUISITION

In this chapter a model for language acquisition will be described in some detail. The model has been programmed and run on a computer; from time to time reference will be made to a trace of the program's operation contained in Appendix C. First a flow diagram of the process itself is outlined, and in following sections the various components are described.

5.1 Flow Diagram



(The numbers in parentheses refer to the subsection of this chapter in which the component is discussed.)

According to this model syntactic acquisition proceeds by processing sentences (2), one after the other, trying to understand each in turn. To understand a sentence means to produce a correct analysis of that sentence. The component which models the process of understanding is the parsing algorithm (4). This component takes a sentence it has been presented with and the grammar of the model, along with any current grammatical hypotheses the model may have generated, and attempts to produce an analysis of the sentence. An analysis consists of a labeled bracketing of the input sentence along with an identification of the functional relations between parts of the sentence.

To determine whether an analysis of an input sentence is correct is the function of the comparator component (8). This component determines whether a sentence analysis is consistent with what the child knows (on the basis of his general knowledge of the world and his particular knowledge of the situational context (3) in which the sentence was spoken). If no analysis of a sentence is consistent in this sense with

what the child knows, the sentence is simply discarded and not used at all. Only if the sentence is correctly understood is there any effect on the grammatical competence of the model.

When a sentence is correctly understood two things happen: first, all of the grammatical constructs used by the parsing algorithm (4) to produce the correct sentence analysis are incrementally confirmed (including any grammatical hypotheses that were used) (7); and secondly, the sentence itself may serve as the stimulus for the generation of some new hypotheses (6). After these two operations the model goes on to obtain another sentence and continue its processing.

In addition to the time scale in which sentences are presented one after the other, there is a larger developmental time scale in which the model passes through acquisition stages. (In this model there are three distinct stages.) Each stage is characterized by the generation of a different set of grammatical hypotheses about the language (5). During a particular stage these hypotheses are tested against the sentences the model experiences, and those hypotheses which are accurate characterizations of the language will be confirmed in the usual manner, while

those which are not simply atrophy and are discarded.

The principal mechanism for acquiring a grammatical construct is for a hypothesis to be sufficiently well confirmed to become a part of the grammar. The only exception to this mechanism is that word meanings are taken to be acquired through the use of a word in one or more contexts (verbal and nonverbal) (3). That is, the grammar of the model may be changed directly by the incorporation of new words into the lexicon.

5.2 The Process

In this section various components and how they operate will be described, but to avoid becoming unduly repetitious later on, a very general qualification of what follows is made. Syntactic acquisition is clearly a multifaceted process, and all theorizing about that process will be in greater or lesser respects incomplete. Furthermore, in a formal study such as this it is rather difficult to conceal what is not known by careful selection of words. It is not possible to avoid making decisions about many aspects of language acquisition, which decisions, as formal approximations to some process or situation, can be faulted in various ways; that is the price one pays for being formal.

In what follows the reference of what is being described may shift from time to time from the computer model to the child and back. The only significance of this is that when referring to children the psychological nature of an assertion is emphasized.

5.2.1 General Features

It may help in understanding the functioning of the model if one of its general characteristics is first noted; that is that there is no direct provision for disconfirming anything (grammatical hypotheses or constructs of the acquired grammar). The presumption that presented nonsentences have no particular significance for syntactic acquisition was argued previously (see 2.2.1); here it is further assumed that negative data of any sort have only an indirect effect on the child's language competence. The negative signal from the comparator component (the only negative data in the system), while crucial to the functioning of the model, results only in an absence of confirmation for certain grammatical constructs and hypotheses. There is no active disconfirmation at all.

What then is the mechanism for infirming a hypothesis or grammatical construct? Hypotheses in fact

start off unconfirmed, and if after a period of time a hypothesis has not been confirmed, it becomes increasingly less viable in the sense that it is considered less frequently as a possible hypothesis; furthermore if it is considered and receives the partial confirmation of an example, that confirmation will count for less than if it had been partially confirmed earlier. The principle is simply that hypotheses will atrophy in the absence of confirmation.

Unless the child's language competence is always correct (although incomplete) at each stage, there must be a mechanism for change. This mechanism is taken to be passive in the sense that a construct will eventually drop from the grammar unless it is continually confirmed (by being used in an acceptable analysis of a sentence). A general characteristic of this model, then, is that every grammatical construct (rules, hypotheses, lexical entries, and features) is subject to the same process of confirmation by successful use and ultimate disconfirmation by the passage of time without such use.

However, there are problems in applying this notion of confirmation. Rare constructs seem to be acquired (and not forgotten) in spite of their being infrequent, while

fairly common constructs can be readily replaced (cf. the examples of overgeneralization cited in 2.2.2; but cf. also the discussion in 3.1). In this model the degree of confirmation of a construct is a function of the number of successful uses of the construct weighted by its relative recency (see 5.2.7).

5.2.2 Data Source - Sentences

This model learns by listening to simple declarative and simple imperative sentences. This of course is a far more homogeneous verbal environment than the child would normally experience; see 5.3 for an experimental extension where some ungrammatical input data were added. The sentences themselves were randomly generated using the simple phrase structure grammar in Table 2 of Appendix A. Since there is no meaningfulness constraint nor any concern for morphological features, the sentences produced are frequently rather strange (chairs read, boys are broken, toys want things, and so on). Any distinctions not built into the grammar (e.g. the animate-inanimate noun distinction) could not of course be acquired by the model.

The data source might have been made somewhat more "natural" by prior selection of a set of sentences

to serve as input, but this should not have changed the results significantly, and the convenience of being able to make changes in the grammar and have the computer randomly generate sentences would have been lost. In Appendix C there is a trace of the model's operation. On the left there is a column described as "spoken sentence", and it is the sentences in this column that the model was given as input data.

5.2.3 Data Source - Situational Contexts

Within this model situational contexts are presumed to have three functions, one of which will be discussed in section 5.2.8 dealing with the comparator component. A second function is related to the acquisition of word meanings. In this model it is assumed that a usual means of acquiring words during the early stages of language learning is by accumulative experience of the word in context.¹ This assumption is not an explanation of the process nor is it intended to be; it is not in fact apparent what needs to be explained, for what has a

¹ See Warner and Kaplan (1950) for a method for investigating word meaning acquisition.

child acquired when he has learned the meaning of a word? Has he learned a set of semantic features, or the position of that word in a complex associative net, or something else?

The third function of situational contexts has to do with the semantic properties of a word. As was mentioned in 4.3 (and as will be described in 5.2.5), categories are presumed to be initially characterized by the child. The characterization of a category is the enumeration of the semantic property (or properties) common to its members. When a child encounters a word in context he classifies that word according to the semantic distinctions he possesses at that time. To illustrate, suppose the child has at a certain stage two initial categorial characterizations, one that defines a "thing" class, and another that defines an "action" class. When a word is encountered in a particular situation the child classifies that word as a member of one class or the other if it fits that class reasonably well in terms of the semantic properties of the word.

The actual way in which this categorial assignment is accomplished will be discussed in 5.2.5 where the operation of the model in the different stages of acquisition is examined.

5.2.4 Parsing Algorithm

How children use their knowledge of the language in understanding a sentence spoken to them is a topic much beyond this research. This model simply has a parsing algorithm that will, given a set of context-free grammar rules with functional relationships identified (some of these rules are part of the model's acquired competence while others are unconfirmed grammatical hypotheses) and a sentence, produce partial or complete analyses. The output from the algorithm is a labeled bracketing along with an identification of the functional relationships between components of the sentence. (For a description of the algorithm see Appendix B.)

Sentence analyses are presented to the comparator component one at a time. The sequence in which they are presented reflects the degree to which an analysis "accounts for" the sentence and how highly valued the analysis is. Here "accounting for" means the proportion of words actually used in the sentence parsing to the total number of words in the input sentence. Possible analyses of the sentence "Mommy will eat that chair" might include (i) "Mommy eat" (which accounts for

2/5 of the sentence), or (ii) "Mommy eat chair" (which accounts for 3/5 of the sentence.) An analysis will be highly valued to the degree in which the categorial assignments of the words are highly valued (well confirmed) and the degree to which the rules used in the analysis are highly valued (well confirmed). For the specific formula used in computing analysis values see Appendix B.

The psychological claims that this component of the model is intended to make are the following:

(1) The process of understanding depends upon the grammatical competence of the child along with any hypotheses he may have about the language.

(2) The child will attempt to account for as much of the sentence as possible in his analysis; and his analysis will be as habitual as possible in the sense that he will tend to use well confirmed grammatical constructs (both rules and categorial assignments) in preference to less well confirmed constructs.

(3) The child is able to skip over parts of a sentence which are not understood and produce a structural description of the remainder of the sentence. He not only skips over words which are not known but also words (and constitutes) which, although understood when they occur by themselves,

cannot be combined into a sentence analysis. This ability was demonstrated in a repetition experiment (C. Smith 1966), in which the children, presented a sentence for repetition, often dropped out complex parts of it. Examples cited include:

(i) (presentation) Mommy could have lost her purse.

(repetition) Mommy lost her purse.

(ii) (presentation) The boy who was running fell down.

(repetition) The boy fell down.

How this skipping is accomplished by the parsing algorithm is described in Appendix B.

5.2.5 Initial Hypotheses

The particular initial hypotheses assumed by the model along with how they are used are described in this section.

The model has initial hypotheses about lexical categories and functional relations. To have a hypothesis about a lexical category is to have a semantic definition of that category (a definition of the property or properties common to all members of the category).

Hypotheses about lexical categories are used by the child to classify words that he might encounter in his verbal environment. The child is presumed to have the capacity to recognize whether a word (with certain semantic properties) fits any of his hypothesized categories; when and if it does, that particular categorial assignment of the word is incrementally confirmed.

To have a hypothesis about a functional relationship is (i) to have a semantic definition of that functional relationship (i.e. what it means for X to bear the relationship R to Y), and (ii) to know what other functional relationships are required to occur when this relationship is evidenced (e.g. the functional relationship "subject of a sentence" requires that there be something else in the sentence for which the subject is the agent).

These hypotheses will be used in the interpretation of sentences. To illustrate this, suppose that at a particular stage in language acquisition the child has two functional relationships which are called "the reference of the sentence" and "the modifier of the sentence". Suppose further that one of the defining characteristics of the relationship "modifier of the sentence" is that there be a unit that functions as the "reference of the sentence".

With this much equipment the child will attempt to interpret sentences that he hears as comprising one of three possible patterns: (i) either the sentence consists of just a reference, or (ii) it consists of a reference followed by a modifier, or (iii) it consists of a modifier followed by a reference. A priori all three sentence analyses are equally likely for the child; there is no way for him to know that in a particular language one or more of the patterns occurs not at all or only infrequently.

Notice that in this example the three possible patterns the child is prepared to deal with are patterns in terms of functional relations; they are not patterns of categorial sequences. As far as the child is concerned any category could function as the "reference of the sentence" and any category could function as the "modifier of the sentence" (when the sentence has a modifier) including categories that might themselves be constitutes. For example, if the child has acquired a rule $X \rightarrow \text{thing} + \text{thing}$ (where "thing" is the name of one of the lexical categories initially defined for the child), then it is perfectly possible, a priori, for an X category to function as the "reference of the sentence" or "modifier of the sentence". The general principle is that a priori any

category (lexical or major) can fulfill a particular grammatical function.

What the child is prepared to do, then, is to determine which patterns (in terms of both functional relations and categories) are evidenced in the language.

As mentioned previously, the model passes through distinct stages in acquisition. What distinguishes one stage from another is the evolution of a new set of initial hypotheses. The specific assumptions about initial hypotheses will be described as they are defined for each stage; their use will be illustrated by the trace of the program.

5.2.5.1 Stage 1

At this stage the child is assumed to possess a single holophrastic linguistic concept with the approximate meaning of "concrete reference". No distinction is made at this stage between category and function; the child simply has one category (a category of concrete things), and that category always functions as the concrete reference. Although the category and function are at this stage coextensive, they are given different names in anticipation of a time when they will not be coextensive (beginning with stage 2). The category is

called "thing" and the functional relation is called "concrete reference" (abbreviated C-REF).

With this initial characterization the child is prepared to begin learning from spoken sentences. He has, at this stage, a single hypothesis: that a sentence will consist of a single concrete referent (a "thing" word functioning as the "concrete reference of the sentence").

However this hypothesis by itself will not enable the child to understand (in part) a sentence, for so far he does not have any vocabulary. Somehow he must learn the meaning of words. In 5.2.3 the acquisition of word meanings was credited to the experiencing of words in context. The words typically acquired first are content words (nouns, verbs, and adjectives). Brown and Bellugi (1964:138-40) explain the favored role of these content words in terms of their referential properties, their high information content, and the heavy stresses they usually carry.

This fact of content word acquisition is modeled (but not explained) by the random selection from each input sentence of one of the content words. If that word is a new word for the model, it is added to the lexicon. This is not to imply that in reality word meanings are typically acquired as a result of the word being used in a single sentence, but

changing the program so that words were acquired only after n examples ($n > 1$) would only have slowed down the program without otherwise affecting the results. Neither is it claimed that all and only all the content words in adult speech are early acquisitions through use in context; this is only a first approximation.

At any rate, the content word randomly singled out in this way is also assigned to a category by the model. In general, as outlined in 5.2.3, the model inspects the categorial definitions it possesses to determine which category, if any, a word fits into. In stage 1 there is only one category (the "thing" category), but this class is sufficiently broad so that any content word selected by the model is presumed to satisfy the requirements for membership.

The program trace in Appendix C illustrates these points. The sentences presented to the model are listed under the heading "spoken sentence"; the first sentence the model hears is "That chair will read." The word "chair" (one of two content words in the sentence) is randomly selected and entered into the model's lexicon with the categorial assignment "thing". Then the model attempts to interpret the sentence on the basis of its

acquired grammar (which right now is empty, as nothing has been acquired) and any grammatical hypotheses it has (in fact there is only one, that a sentence consists of a "thing" word functioning as the "concrete reference of the sentence"). And there is a possible interpretation, namely that "chair" functions as the sentence's "concrete reference". As this interpretation is not inconsistent with anything the child knows extralinguistically (see 5.2.8), it is accepted. On the right is printed under the heading "comprehended sentence" that part of the presented sentence which was understood by the model; in this case the single word "chair". The sentence analysis is printed to the right of this, and it indicates that the word "chair" with the categorial assignment "thing" functioned as the sentence's "concrete reference". Furthermore this was a new word for the model; it had not encountered "chair" before. (Later on the notation "word acquired contextually ____" might mean that the word, although it had been encountered before, had not been used in some time and so had been forgotten; note the discussion of 5.2.1.)

The second sentence the model hears is "Mom see." The content word "see" is randomly selected and entered

into the lexicon with the categorial assignment "thing", and then the sentence is interpreted using the same grammatical hypothesis as before, namely that a sentence consists of a "thing" word functioning as the sentence's "concrete reference".

Sentence three is "This man can sleep." The content word "sleep" is randomly selected and entered into the lexicon with the categorial assignment "thing", and then the sentence is interpreted using the grammatical hypothesis already used in the interpretation of the first two sentences.

Stage 1 continues as more sentences are presented. After every twenty input sentences the model's dictionary and its acquired grammar rules are printed so that the reader can follow the progress of the model. Following the dictionary (which will be discussed in 5.2.6) there is the heading "grammar", beneath which are the acquired rules of grammar. At this stage the model has learned one rule; that a sentence can consist of a "thing" word that functions as the

sentence's "concrete reference". What was formerly just a hypothesis about the language has become a grammatical construct by virtue of the confirming examples of sentences 1 through 20.

Sentence thirty marks the end of the first stage. The conclusion of this stage is not contingent in the model upon anything, although in actual fact the evolution of initial structure would probably be contingent upon acquiring a certain amount of experience or competence. The number thirty was chosen, however, to ensure that the model had acquired a fair vocabulary and a well-confirmed rule by the end of stage 1.

5.2.5.2 Stage 2

This stage marks the actual beginning of syntactic development. Although not necessarily simultaneous in real life, several new hypotheses appear at the beginning of stage 2. Whereas before the model had a single categorial concept (the "thing" category) there are now two such concepts; the model distinguishes between a class of "things" and a class of "actions". What takes place is that the previously very broad class of "things" splits into two classes as the child becomes aware of the distinction between things and actions.

The semantic definitions of these two categories are reflected in two parts of the model. One of these is the process of word meaning acquisition described in 5.2.3. Whereas in stage 1 any randomly selected content word was classified as a member of the "thing" category, some content words will no longer fit that category (in terms of semantic properties), but instead will now fit the action category. The formalization of this in the model is that if a selected content word is an adult verb it is presumed to fit the "action" class, while if it is an adult noun or adjective it is presumed to fit the "thing" class.

The second part of the model that reflects the semantic definitions of these categories has to do with the comparator component (described in 5.2.8). Although adult verbs will hereafter be classified as members of the model's "action" class, there may be some verbs that were classified as "thing" words during stage 1; these verbs would still be members of this stage 2 "thing" category. One might hypothesize that when this stage 2 categorial distinction is made the child works through his lexicon and purges any now incorrect categorial assignments; but the mechanism proposed here is a more passive one that

depends upon the process of interpretation. Any putative analysis that the model now produces in which an adult verb is categorized as a "thing" word is taken to be unacceptable (by the comparator component; see 5.2.8). This simply means that the categorial assignment of verbs to the "thing" class will no longer ever be confirmed, with the result that eventually (by the mechanism of disconfirmation with the passage of time) no adult verbs will belong to this category. This is shown by the program trace; after 20 sentences all verbs are members of the "thing" class (this being stage 1); after 60 sentences most verbs have overlapping membership in both the "thing" and the "action" classes; and after 80 sentences almost all verbs belong only to the "action" class. (See 5.2.6 for a description of the lexicon.)

Stage 2 is also marked by a new hypothesis about functional relations. Whereas before the model had a single functional relation (the "concrete reference of a sentence") there is now an additional relation called "modifier of the sentence" (and abbreviated MODIF). There may in fact be several different types of modification subsumed under this label (possession, attribution,

quantification, etc.) which in a more adequate model would be treated independently. What has happened is that the child's previously unitary linguistic concept of concrete reference has fractionated. A concrete reference is no longer viewed as an indivisible whole, but is rather viewed as a reference that can be modified in one way or another (see the discussion in 3.2.4).

Part of the semantic characterization of the functional relation "sentence modifier" is that there must be a referent (the sentence's "concrete reference") to be modified. Consequently, there are now two new hypotheses available to the model: that a sentence consists of a "sentence modifier" followed by a "sentence reference" or that a sentence consists of a "sentence reference" followed by a "sentence modifier". And there is no a priori reason to favor one of these hypotheses over the other.

Note again that these hypotheses are stated in terms of functional relations; they are not patterns of categorial sequence. As far as the child is concerned any category (either a "thing" or an "action") can function as the "sentence modifier" or the "sentence reference". As there are two possible orders in terms of grammatical function (either the "sentence reference" followed by the

"sentence modifier" or the reverse) and since there are four possible permutations of categorial sequence (two "thing" words, two "action" words, or a "thing" and an "action" word in either order), there are $2 \times 4 = 8$ sentence types possible. These may be viewed as distinct hypotheses about the language, some to be confirmed by the language data.

The previously acquired grammatical construct will continue to play a role during stage 2. It may be that some sentences will be analyzed using that construct as sentences with a single word "concrete reference". This interpretation will only take place, however, if it is impossible to analyze a sentence in terms of the two relations "concrete reference" and "modifier", for as noted in 5.2.4, the model is motivated to account for as much of a presented sentence as possible.

The previously acquired grammatical construct has another influence on stage 2. Recall that at this stage the child is prepared to recognize eight possible sentence patterns (nine if the single word sentence pattern of stage 1 is included). Some of these patterns will be favored by the child, however, in the sense that if he can interpret the presented sentence using a favored

pattern he will. The reason a pattern might be favored is that it is consistent with the child's previous grammatical knowledge. To illustrate: the child's previous grammatical knowledge at the beginning of stage 2 is that sentences consist of a "thing" word that functions as the sentence's "concrete reference". That is to say that the child will have a strong bias toward interpreting "thing" words as fulfilling the function of the sentence's "concrete reference" and vice versa. And some of the aforementioned eight hypothetical sentence patterns are consistent with this bias (an example would be the pattern describing a sentence as a "thing" + "action" where the "thing" functions as the sentence's "concrete reference" and the "action" functions as the "sentence modifier"), while others are not (an example would be the pattern describing a sentence as a "thing" + "action" where the "thing" functions as the sentence modifier and the "action" functions as the sentence's "concrete reference").

The program trace illustrates these points. Sentence 31 is "That boy can read." The word "read" is the content word randomly selected from the sentence. Although this word already appears in the lexicon, it has never been assigned to the category "action", and so this

categorial identification is entered in the lexicon for this word. (The analysis directly below will say that this word was acquired contextually, which is not exactly accurate as the word had already been acquired; what is acquired is the categorial assignment of this word to the "action" class.)

The model then tries to interpret the sentence which probably consists of, in terms of categorial assignments, an undefined word followed by a "thing" word followed by an undefined word followed by a word that is either a "thing" or an "action" ("read" would still be a member of the "thing" category on the basis of its earlier classification; cf. sentences 5, 18, and 30). There are four possible analyses of this sentence that are consistent with the child's stage 2 hypotheses:

- (i) "Boy" is a "thing" word functioning as the sentence's "concrete reference" followed by "read", a "thing" word functioning as the "sentence modifier".
- (ii) "Boy" is a "thing" word functioning as the "sentence modifier" followed by "read", a "thing" word functioning as the sentence's "concrete reference".

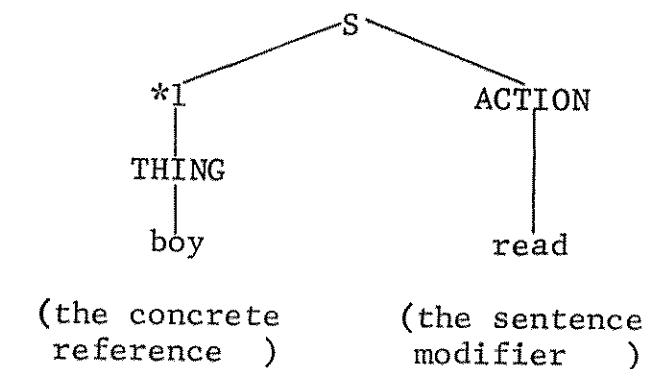
- (iii) "Boy" is a "thing" word functioning as the sentence's "concrete reference" followed by "read", an "action" word functioning as the "sentence modifier".
- (iv) "Boy" is a "thing" word functioning as the "sentence modifier" followed by "read", an "action" word functioning as the sentence's "concrete reference".

Of these four analyses, numbers (i), (ii), and (iii) will be favored over (iv), as they are consistent with the model's present grammatical structure where the sentence's "concrete reference" is a "thing" word.

The program trace does not indicate whether there were any analyses attempted before (iii) was successfully accomplished. There may have been, for both (i) and (ii), if produced and evaluated by the comparator component, would have been rejected as inconsistent just because the word "read" does not fit any longer into the "thing" category (see 5.2.8). Analysis (iv) probably was not attempted as it is not a favored analysis, whereas (iii) is; (iv) would be considered only if all possible favored analyses had failed. Even if it had been attempted it also would have been rejected as inconsistent, for

"read" cannot function as the sentence's "concrete reference" (see 5.2.8).

The accepted analysis is printed to the right of the sentence and corresponds to the tree structure:



The *1 node is simply an uninteresting artifact of the parsing algorithm.² The general principle is that no claim is made by the presence of any *m node (m a number) that dominates only a single constituent.

²Recall that before there was a rule $S = \text{thing}$ (the concrete reference). Now there are a number of hypotheses about possible rules, the one that was actually confirmed by the last sentence being $S = \text{thing} + \text{action}$ (where the thing functions as the sentence's concrete reference and the action as the modifier). This rule, however, includes as one of its components the former rule, and so we actually have the rules $S = \text{thing}$, and $S' = S + \text{action}$. And rather than introduce different symbols for sentences (S' , S'' , etc.) the program does the following: whenever there exists a rule of the form $S = X$ and there is a hypothesis created of the form $S' = \dots S \dots$, then replace the $S = X$ rule by $*m = X$ (where *m is a newly defined symbol); make the rule hypothesis $S = \dots *m \dots$, and define a new rule $S = *m$.

Apparently in sentence 32, "The boy must see", the word "boy" is the content word randomly selected, and as the word already has an appropriate lexical entry, no action is taken. But as a result, in this sentence only the word "boy" is understood by the model. The word "see", although acquired back at sentence 2, has not been used since and so has been forgotten. Consequently the sentence can only be analyzed in terms of the stage 1 single word grammatical construct.

In the next eight sentences (through sentence 40) the hypothesis successfully used in sentence 31 is used three times more and has now gained a fair degree of confirmation. Following sentence 40 the dictionary is printed and then the grammar. As shown in the grammar, the model now has two types of sentences, the early:

S = thing (functioning as a concrete reference)
and the newer rule:

S = thing + action (where the thing word functions as the concrete reference and the action word functions as the modifier). (Note: when an asterisk appears as the label of a functional relationship it is serving simply as a placeholder. The rule:

S = *1 ACTION * MODIF

means that a sentence can consist of an *1 constituent (functional relationship unspecified) followed by an action constituent (which fulfills the function of a sentence modifier).)

Sentence 48 illustrates the first verification of the hypothesis that a sentence might be an "action" word followed by a "thing" word, with the former functioning as the "sentence modifier" and the latter functioning as the sentence's "concrete reference". This hypothesis is confirmed again with sentences 53, 56, and 57, which is sufficient confirmation for the rule to become part of the grammar that is printed following sentence 60.

Sentence 46 illustrates the first verification of the hypothesis that a sentence might be two "thing" words with the first functioning as the "sentence modifier" and the second functioning as the "concrete reference of the sentence". This hypothesis is confirmed again with sentence 49; however this is not yet sufficient confirmation of the hypothesis, and so it does not appear in the grammar following sentence 60. With the confirmation of sentences 66, 71, 76, and 80, however, it becomes a theorem of the grammar as shown in the printout after sentence 80.

5.2.5.3 Stage 3

There is one new functional relation introduced at this stage, and no new hypotheses about categories. The relation "subject of the sentence" has been added to the two functional relations of stage 2 (the "modifier of the sentence" and the "concrete reference of the sentence"). The semantic content of this relation is taken to be quite similar to what the relation of the same name means in an adult grammar (see 5.2.8). What this relation implies in terms of other functional relations that need to occur when this relationship is evidenced is, however, not obvious. Is it the case that a sentence with a subject must have something like a predicate? Clearly something is required; a sentence cannot consist of a subject alone. At any rate in this model it was assumed that when a subject occurred there would also have to be something that functioned as the "sentence reference" and something that functioned as the "sentence modifier". This decision can be faulted in that, among other things, it precludes the possibility of the model understanding sentences with a subject and an intransitive verb. (Such a sentence would only have one part in addition to the subject whereas two parts are required.) There are

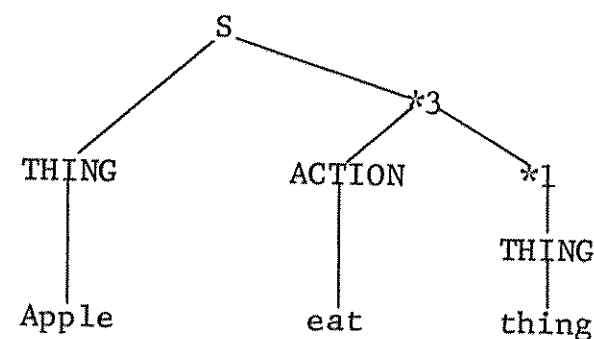
various ways in which one might improve this characterization of sentence subject although none have been tried here; the whole area is one that deserves careful study.

The model now has the new hypothesis that a sentence will consist of three parts; one part that functions as the "sentence subject", one part that functions as the "sentence modifier", and one part that functions as the "sentence reference". However, the hypothesis says nothing about the order of these three parts, and so a priori there are $3! = 6$ possible patterns in terms of functional relationships. Furthermore, any category can satisfy one or more of these functions. At the present time there are five categories; the two lexical categories ("thing" and "action") and the three categories where two words were combined (in stage 2) to form a sentence constituent. The task of the child, therefore, is to determine which of these hypotheses (about the order of grammatical functions and about the categorial assignment of units that fulfill a particular function) are true characterizations of the language.

Even as the grammatical construct acquired in stage 1 had an effect on the operation of the model in stage 2, so the grammatical constructs acquired in stage 2

will have an effect on the operation of the model in stage 3. In particular, any hypothesis that is consistent (in the sense explicated in 5.2.5.2) with these grammatical constructs will be favored. This is intended to reflect the psychological fact that the child will attempt to be as habitual as possible in the sense of using his previously acquired constructs in the analysis of presented sentences.

Stage 3 begins with sentence 81 in the program trace; however, for one reason or another it is not until sentence 89 ("This apple eat the thing") that the model uses the new relation. This sentence the model interprets as a thing + action + thing sentence, where the first element functions as the "subject", the second as a "modifier", and the third as a "concrete reference". However, the sentence is not analyzed into three parts but rather into a hierarchical structure because the last two words themselves form a constitute that has been well confirmed in stage 2. The sentence analysis is:



This example, and others similar to it that follow, illustrates what is taken to be a basic mechanism for the origin of hierarchical sentence structure. A rule for combining elements becomes confirmed through its successful use, and the constitute thus acquires a psychological unity. In later stages the child will attempt to understand new sentences in terms of previously confirmed units, and when he does this by combining one or more of these units with one or more other elements he will have a hierarchical analysis.

The pattern of sentence 89 is confirmed by sentences 90, 91, 93, and 100, which is sufficient for the rule to appear as a construct in the grammar following sentence 100. The rule:

S = THING *3 SUBJCT *

is the precursor to the eventual adult

S = NP + VP.

The model thus accounts for the origin of the subject-predicate analysis of English sentences by assuming that the initial structure of the child's language acquisition mechanism is such that he will acquire a constitute that will eventually evolve into the adult predicate before the functional relation of sentence subject is available to him.

The eventual predicate phrase acquires a psychological unity because it is, at stage 2, an acceptable sentence analysis. Later (in stage 3) as the model's grammatical competence is extended by the addition of the functional relation "sentence subject", the psychologically unified predicate phrase analysis is combined with another element functioning as subject, and a hierarchical subject-predicate analysis has been accomplished.

Nothing further of note happens during this stage. The subject-predicate analysis becomes the favored one after 180 sentences, as shown by its being the first rewrite of S in the grammar, which lists the rules by symbol in order of degree of confirmation.³

5.2.5.4 Summary and a problem

Three stages in syntactic acquisition characterized, in terms of initial structure, by the development from a single functional/categorial concept to three relations and two categories, have been

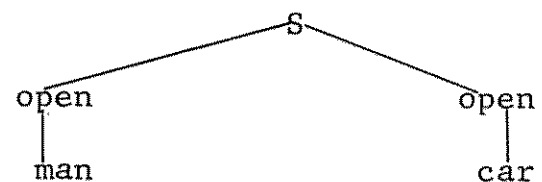
³ Appendix C traces the learning of the model only through presented sentence 121, as the subsequent operation is lengthy but not particularly interesting.

distinguished. Progress from stage to stage occurs through the formation of hypotheses about sentence structure based on the initial structure available as well as any grammatical constructs previously acquired. It is shown that there is a natural explanation for the development of a subject-predicate sentence analysis by the formation of a constituent that is a precursor to the predicate prior to the availability of the subject relation. The origin of hierarchical structure is shown to follow from the combination of a psychologically unified constitute with another element.

One aspect of syntactic acquisition that is not explained, however, is the evolution of a mechanism for characterizing functional relationships in terms of deep structure P markers (Chomsky 1965:68ff). In this model functional relationships are defined by each rule of the grammar; that is, a rule, in addition to defining a categorial sequence, also defines the functional relations that exist between the constituents. Apparently, however, in adult grammars only a single functional relationship holds between a category and the category that immediately

dominates it in the deep structure (Chomsky 1965:72)⁴, and if this is so, then a generalization is missed if the grammatical formalism does not impose this upon the class of grammars that can be acquired.

It is not clear, however, that in early stages of language acquisition the generalization actually does hold. As noted in 3.2.2, an early sentence pattern is a two word sentence consisting of two open-class words. But this creates a problem. The structural description could not be:



if there was a characterization of functional relationships independently from a rule, for two different relations are defined by the configuration (open, S) (the symbol "open" directly dominated by S). It might be, however, that on

⁴That the formalism for expressing functional relations in Chomsky (1965) will prove to be adequate is an open question under investigation (Barbara Partee, personal communication).

other grounds one could justify an extra node that came between the S and one of the open symbols.

Although the rule form of this model is similar to that of tagmemics in that in both cases functional relationships are defined by grammar rules, the similarity does not extend beyond early stages of language acquisition. The difference is that the model is assigning functions to components of what are presumed base rules and will not continue to do so with the development of transformational rules, whereas in tagmemics all patterns (including ones that are transformationally derived in transformational theory) are described by tagmemic (function:form composite) patterns; there is no distinction between deep and surface structures.

5.2.6 Hypothesis Generator

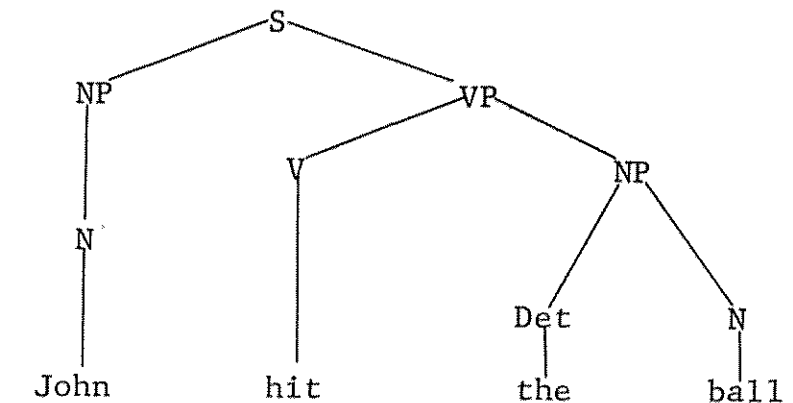
The operation of this component is contingent upon the sentences the model receives as input. The initial structure of the component constrains it to single out certain possibly important properties of sentences, which may be considered language hypotheses. Data that bear on these properties will be accumulated by the component that confirms various grammatical constructs, and a hypothesis, if eventually confirmed,

will become a construct of the grammar.

It is in the functioning of this component that the two grammatical operations described by Braine (contextual generalization and acquisition of positional contingencies) are included. Clearly, both of these operations are conceivable only if the domains in which they are possible are fairly closely limited. If there was, for example, no initial constraint on what could count as a context to be associated with an element, the operation would never get started, for in general there would be huge numbers of possible contexts (e.g. a relevant context for a word might be the preceding word, or the following word, or the preceding word and the following word, or the two preceding words, or the second preceding word removed, and so on). Similarly, the domain in which positional contingencies might hold must be closely limited, for in the absence of such delimitation there would also be an extreme number of possibilities.

The hypothesis incorporated rather tentatively in this model is that a word, categorial label, or higher nonterminal node, immediately preceding or following a particular word, might be associated with that word, either because there is a positional contingency or a

subcategorization contingency. To illustrate, consider the sentence "John hit the ball." If the phrase marker was



then the following elements would be associated with "John":

preceding context = sentence initial position

following context = VP, V, "hit"

The elements associated with "hit" would be:

preceding context = NP, N, "John"

following context = NP, Det, "the", and so on.

All of these relations may be viewed as hypotheses about language structure, and of course most of them are non-significant (the relation of "hit" preceding an NP is a correct subcategorization contingency).

All of these associations would be stored as hypothetical features of lexical entries, and if they are not significantly confirmed, they will drop out. When and if they are confirmed, however, they will become grammatical constructs.

The content of this component of the model, then, is that positional contingencies will be learned if the contingency is between two neighboring elements, and acquisition of contextual invariances of a word will take place (i) when the context is a category that dominates (not necessarily directly) the immediately preceding or following element which, however, does not dominate the word in question; or (ii) when the context is sentence initial or sentence final position. Braine has suggested other possible contextual invariances, such as statement of position relative to intonation markers or closed class morphemes which, however, have not been included in the model.

The trace in Appendix C illustrates the operation of this component. Each lexical entry in the dictionary (printed after every twenty sentences) may have five things associated with it: (i) a list of elements that often immediately precede this word; (ii) a list of categories that often dominate immediately preceding words; (iii) a list of elements that often immediately follow this word; (iv) a list of categories that often dominate immediately following words; and (v) the category

(or categories) to which this word belongs (see 5.2.5).

To illustrate: the first word in the dictionary following sentence 20 is "apple", and it has three features; a tentative contingency between "a" and "apple" and between "apple" and "will", and the categorial feature "thing". These are the only features that have received a fair amount of confirmation. (Note: when the symbol // occurs in the dictionary it means sentence boundary.)

After 120 sentences "apple" has acquired seven features. The words that often precede "apple" are "a" and "this", and "apple" often occurs in sentence final position. It also often follows and precedes a member of the category "action", and often precedes an *3 phrase (a predicate phrase). And it is (still) a member of the "thing" category.

To summarize, three points should be noted:

(1) There must be a large initial constraint on the types of sentence properties to which the child is responsive. (2) The one important claim made by these specific suggestions (apart from which all of the suggestions are very tentative) is that the child may associate with a word the abstract property of neighboring categories; his ability to make associations is not

limited only to possible relations between physically occurring words. (3) Although in early stages of language acquisition this component functions only to recognize invariant associations of one type or another, it is likely that in later stages, as the linguistic knowledge of the child increases, a more active hypothesis generation and testing mechanism comes into use. This possibility, discussed in 4.4, has not been reflected in this model.

To what extent any of the data accumulated by this component can provide an explanation for the acquisition of subcategorization features is not known. It may be that the distinction, say between transitive and intransitive verbs, is a generalization acquired through many examples of verbs preceding noun phrases or in sentence final position, but it would not be surprising if there were another quite different explanation perhaps involving certain semantic implications of the distinction.

Contingent hypotheses generated by this component are not taken to play a central role in early syntactic acquisition, and it is in part because of this that no attempt has been made to solve the difficult

problem of determining when one of these hypotheses is sufficiently well confirmed to constitute a grammatical construct. At the present time all contingent hypotheses are simply confirmed (or infirmed by the passage of time without successful use); criteria for successful confirmation have not been established. This means that any positional contingencies reflected in the lexicon are not transformed into grammatical rules (e.g. if the element X was usually followed by the element Y, it would be reasonable to combine these two elements into a constitute whose distribution could then be learned). Subcategorization features, on the other hand, are not transformed into rules at all; a hypothesis about subcategorization is stated in precisely the form in which the eventual grammatical construct will be stated when and if the hypothesis is confirmed. This is one of the convenient consequences of the use of lexical features and a lexical substitution rule in the most recent characterization of transformational grammar (Chomsky 1965).

5.2.7 Confirmation of Grammatical Constructs and Hypotheses

The most important feature of the confirmation component is simply that it is not used at all unless

a sentence is understood. But when it does function it will confirm all grammatical constructs and hypotheses that were used in the accepted sentence analysis or, in the case of the contingent hypotheses, are consistent with that analysis. The general principle is that a confirmation counts in proportion to its recency. This is accomplished in the model by defining a function called a "unit of confirmation" that increases (linearly) with the number of presented sentences. When a construct is to be confirmed its value is incremented by this unit. There are a number of ad hoc decisions that were made as to when constructs and hypotheses should be printed or dropped, but on the whole they are rather uninteresting. The only point worth mentioning is that in addition to storing and computing values for all constructs, certain other values are computed so that it is possible to determine relative frequencies. For example, what is of interest in the feature + __NP for a particular transitive verb is not just its value but its value relative to the number of occurrences of this verb. Consequently, the model will compute a value which represents the maximum valuation possible on such constructs, and from this the relative frequency of the construct may be determined.

5.2.8 Comparator Component

The process of taking a hypothesized sentence analysis and trying to integrate it with what one knows extralinguistically has already been discussed (see 4.2.1 and 4.4.) When this attempted integration fails the child has not understood the sentence; when it is successful he believes that he has understood the sentence. It is this process of attempted integration that is called comparison, for in a loose manner of speaking, the child is matching his tentative semantic interpretation of a sentence against the rest of his knowledge, and when they are inconsistent he will reject his tentative analysis.

The degree to which the child uses his "comparator component" is presumed to be in inverse relationship to the degree to which he has acquired grammatical competence. That is to say that initially the process of comparison is very important indeed, but as acquisition of a grammatical system proceeds, the child will tend more and more not to reject a hypothesized analysis as much as to disbelieve what is said or not to understand it at all. Clearly, if an adult hears a sentence which conflicts with his world view, say "The dinner ate Lonnie", he does not reject his subject, verb, object analysis of the sentence but rather

questions what was said. Of course what is going on here is the presentation of a single counter-example to an extremely well confirmed linguistic system.

Perhaps what is being claimed can be explained in terms of cognitive dissonance theory. When a person derives a semantic interpretation for a sentence he has heard, he is faced with reconciling his interpretation with the rest of his knowledge. When the sentence interpretation and that knowledge are essentially inconsistent in one way or another, the child just beginning to learn a language will find it easiest to reject his semantic interpretation and the linguistic analysis on which it is based, whereas the child or adult who has acquired a well confirmed linguistic system will find it easiest to question the validity of the sentence.

Since formalization of either the child's knowledge of the world or how he might use this knowledge in trying to verify a semantic interpretation is out of the question, it was necessary to devise a shortcut. Since the parsing algorithm produces a putative structural description for a sentence, it is assumed that unless this structural description is "valid", it will not support a correct semantic interpretation, and consequently,

the comparator component will reject this analysis. What does "valid" mean here? It means that the putative structural description must be consistent with, in a sense to be explicated, the correct structural description (the one produced by the process that generated the sentence in the first place). For example, if the sentence "Daddy ate his dinner" were generated, a correct structural description would identify "Daddy" as the subject, "his dinner" as the object, and so on. A putative structural description that distinguished subject and object would be incorrect unless it made those identifications. On the other hand, if the putative structural description was from an earlier stage (stage 1) where only a concrete reference was identified, then the structural description would be incorrect only if "his" was selected as the concrete reference--"Daddy", "ate", and "dinner" all being acceptable concrete references.

The operation of the comparator component can thus be described in two ways: (1) at the psychological level it is the matching of a hypothesized sentence interpretation against what the child knows from his knowledge of the world and his knowledge of the particular situational context in which the sentence occurred;

(2) at the computational level the component is matching a structural description produced by the parsing algorithm against the correct structural description produced when the original sentence was generated. It is important to distinguish these two modes of description, for otherwise one might suppose that, because the comparator component has as part of its input the correct structural description, the model presumes a signaling of constituent structure in the sentence itself or in the situational context in which the sentence was spoken. This is not meant at all; note that the only output from the comparator component is a correct-incorrect signal. The correct structural description produced when the sentence was originally generated is never used by any of the processes that modify the grammatical competence of the model.

The specific relations that must hold between a putative sentence analysis and the correct structural description are stated negatively; if any of these are violated the comparator component produces a negative signal.

(1) Stage 1: No constraints.

(2) Stage 2: The main verb of the correct

analysis may not function as the concrete reference of a sentence nor may it be considered a member of the "thing" category.

(3) Stage 3: In addition to the condition of stage 2, the head of the phrase which functions as subject in the correct analysis must be the same word that functions as subject in the putative analysis.

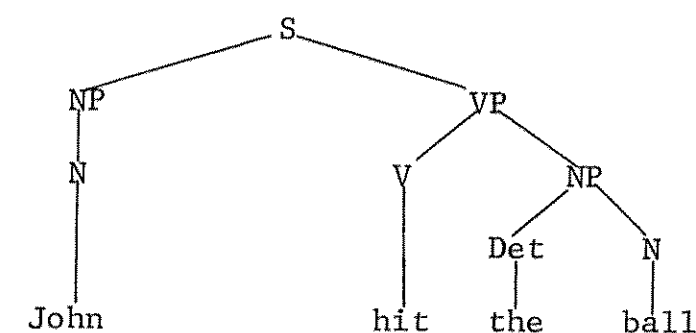
The condition of stage 2 that a main verb of the correct analysis may not function as the concrete reference of a sentence illustrates the conceptual development of a functional relation. Whereas before any content word could function as the sentence's concrete reference, now only nouns and adjectives may, and so the meaning of "the sentence's concrete reference" has changed. The other condition of stage 2, that the main verb of the correct analysis may not be considered a member of the "thing" category, illustrates in the same way the development of a categorial refinement. The condition of stage 3 that what the child identifies as the sentence subject must indeed be the sentence subject in the adult analysis is equivalent to claiming that the semantic implications of calling something a subject have been

mastered by the child. This quantum leap in the model is doubtless a gradual process in the child; presumably a more accurate model would be one in which more than three stages were distinguished.

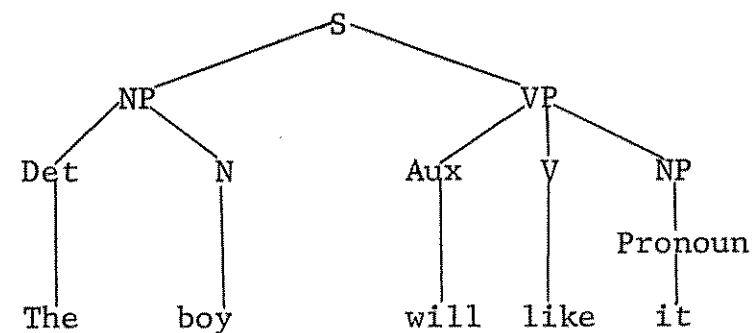
One counterfactual consequent of the comparator component is that in every sentence the specific constraints enumerated above must hold. To take stage 3 as an example, what is being claimed is that the child's knowledge will serve to disallow any sentence analysis in which there is an incorrect identification of the sentence subject. But there are undoubtedly numerous occasions when this would not be true: If the child were to hear the sentence "The ball hit the bat" in a situational context that did not illustrate the sentence, he presumably would not be able to disallow a tentative sentence analysis in which "the bat" functioned as the subject and "the ball" functioned as the object, for there is nothing semantically anomalous about bats hitting balls. Perhaps this situation could be approximated in the model by disabling the comparator component a certain percentage of the time; at any rate, unless this percentage were large there would be no significant effect on the operation of the model.

5.3 An Experiment with Ungrammatical Input Data

To determine what effects the introduction of ungrammatical sentences would have on the model an experiment was performed. Ungrammatical sentences were created by randomly selecting a subtree from the phrase marker of a grammatical sentence and conjoining it with a second randomly selected subtree from the phrase marker of another grammatical sentence. For example, given the tree structure:



there are eight nonterminal nodes, any one of which can be selected as a subtree (the subtree will be the sentence itself in the event (probability = 1/8) that the node labeled S is selected). Suppose the node labeled VP was chosen. Now consider the second sentence



with nine possible subtrees. Suppose the node labeled Aux is chosen. The resulting sentence will be "Hit the ball will." The purpose in creating ungrammatical sentences in this way, rather than say simply permuting words in the dictionary, is that typically ungrammatical speech is structured up to the point of restarting or interruption, and these ungrammatical sentences will have just this property of being, in part, grammatical.

In the experiment the comparator component was disabled whenever an ungrammatical sentence was produced, mainly because the specific constraints built into the comparator component reflect the actual grammar used in sentence generation, and thus would not apply meaningfully to the ungrammatical phrase markers. This means, therefore, that no sentence analysis is rejected for ungrammatical sentences, a condition doubtless counterfactual, but at any rate it will only impair language acquisition rather than help.

An experiment was run providing the acquisition model with 80% grammatical and 20% ungrammatical sentences. (The actual percentage was slightly less than this just because occasionally two conjoined subtrees formed an acceptable sentence.) Table 3 of Appendix A lists the ungrammatical sentences produced. There was only a small effect on grammar acquisition; it took slightly longer to acquire the rules of the grammar than it took with 100% grammatical input, but they were all acquired eventually (except for an S = thing + thing rule, for which there were apparently not enough models); in particular, the basic subject-predicate analysis of the sentence was learned. Furthermore the lexical entries and features that were learned were converging on those acquired with perfectly grammatical input.

What happened in general was that the model just did not interpret these ungrammatical sentences; there were no hypotheses for which these were relevant models. Sometimes the sentence would be interpreted, but any incorrect confirmation of grammatical constructs or hypotheses would have no lasting effect just because there was no system to the ungrammatical sentences.

5.4 Summary

What is taken to be the logic of early syntactic acquisition has been outlined in this chapter. Even if all that has been suggested should turn out to be true it would be just a beginning, for there must be very much more substance to the hypothesis generator and the initial hypothesis components than has been indicated.

Some problems remain: The question of how and when the child's grammar changes (if indeed it does) from one with rule-related functional relationships to a more abstract rule-independent characterization has already been discussed. The problem of what the functional relation "sentence subject" entails in terms of other functional relations was also mentioned. If the functional relation "sentence predicate" does not develop concurrently with the relation "sentence subject" (as indeed it does not in this model; in fact there is no such relation in the model), then where and when does the relation develop? Furthermore, when the constitute formed in stage 2 is combined in stage 3 with an element that functions as subject, there must be some sort of shift in the functional relations of components of this constitute, for the relations

are no longer between a category and a sentence but a category and a predicate phrase. There is very much to be studied here.

CHAPTER VI

SUMMARY

In this research an attempt has been made to single out some of the central processes, or conditions on these processes, of early syntactic acquisition. In the previous chapters the following points have been argued:

1. Syntactic acquisition depends crucially upon the process of comprehension by the child; speech production is taken to be of secondary consequence at most.

2. The mechanism for syntactic acquisition is taken to be the testing of hypotheses about the language by reference to language data. These hypotheses are both initial (either innate or acquired extra-linguistically) and contingent (derived from the language data themselves).

3. Semantics is taken to play a central role in syntactic acquisition. The differential response hypothesis (by which the mechanism for syntactic acquisition uses only "correctly understood" sentences) depends upon the child's interpretation of a sentence he has heard and his recognition of semantic anomaly, which in turn depend upon the semantic properties of both functional relationships and lexical categories.

4. The child need only be exposed to meaningful sentences from the language to acquire syntactic competence. In particular, language instruction (in the form of expansions, corrections, or ungrammatical utterances identified as such) is not essential.

The degree to which various aspects of this model are consistent with one or another learning theory or theory of behavior has been noted occasionally. It was suggested that a component of the mechanism could be interpreted within a mediational framework, or the theory of cognitive dissonance (and presumably the hypothesis testing mechanism would be entirely consistent with an information processing theory of psychology such as that of Miller, Galanter, and Pribram (1960)), but the particular framework in which this proposal fits has been of only peripheral concern.

The kind of psychological theorizing done by Miller, Galanter, and Pribram has not always been well received because of their supposed lack of experimental verification. What they have done principally is to describe a psychological framework, and then to show, in a frequently very programmatic manner, how various types of behavior can be accommodated within that framework.

Verification of their proposal is indirect and consists of showing that there are problems that appear to be explainable by that theory, and that alternate explanations of the problems are not well formed. The epistemological validity of this indirect type of verification is unimpeachable.

This has been the course followed here.

Relevant experimental data have been reported, but very many of the features of this model have little direct confirmation. For example, one might ask for the evidence in support of the differential response hypothesis (probably the most controversial assertion in this model, as it entails almost everything else). The principal evidence is indirect and consists of showing that with certain assumptions and approximations early syntactic acquisition can be explained; furthermore, there are certain deficiencies or problems with other models. To the extent to which this is true, the hypotheses of the model--such as, in particular, the differential response hypothesis--are supported.

This is not to say, however, that indirect verification is all that is possible. There are several research areas that deserve investigation, most of which have already been mentioned:

1. There is a great need for observational and experimental/manipulative studies, where the goal is to ascertain what functional relations are part of the child's language competence at different stages of development.

2. A language comprehension study might determine whether the rather strange phenomenon of losing a correct grammatical structure for a period (see the example in 2.2.2.2) is an accurate statement of a change in language competence or simply an artifact of the difference between speech and comprehension (see 3.1).

3. The research area under most active investigation and ultimately likely to be of most consequence is the purely linguistic study of grammatical universals in natural languages. When a strong case can be made for a particular formal or substantive universal then that fact will have to be accounted for by a theory of language acquisition (see 4.3.3). This seems to be exactly what has happened with respect to the hypothesized formal universal that a grammar have both base rules and transformational rules; all of the people reviewed in Chapter IV were concerned with accounting for the acquisition of transformations.

It is certain, however, that along with direct psycholinguistic investigation there will continue to be a need for the type of indirect verification afforded by formal modeling. One of the not incidental virtues of such an approach is the relative ease with which a model may be reformulated and tested; a whole set of different proposals about substantive universals (e.g. which functional relationships develop when, etc.) could be examined with little difficulty. In this sense, the formal system, in addition to being an instantiation of a particular set of hypotheses, may be viewed as a laboratory for conducting experiments. As knowledge of syntactic acquisition grows and theories become more complex, the problem of determining the content and implications of a theory will increase, and towards this end the use of formal models is expected to be helpful.

APPENDIX A

TABLE 1

SOME EARLY TWO WORD SENTENCES

- | | |
|------------------|--|
| 1. Pretty boat. | (Gregory)* A pivot + open construction. "Pretty" is a modifier (adjectival) of the sentence reference "boat". |
| 2. See boy. | (Gregory) A pivot + open construction. If "boy" is the sentence object rather than sentence subject, "see" might be the sentence reference and "boy" a modifier of "see", which would place the emphasis on the act of seeing. |
| 3. My milk. | (Gregory) A pivot + open construction. "My" is the modifier (possessive) of the sentence reference "milk". |
| 4. Do it. | (Gregory) An open + pivot construction. The sentence reference is "do", and "it" is perhaps a pro-form modifier. |
| 5. Byebye plane. | (Gregory) A pivot + open construc- |

 *The names Gregory and Andrew refer to subjects of Braine's (1963b) study.

tion. This construction might be a comment-topic pattern where the topic is "plane" and the comment is "byebye", or perhaps it is a vocative modeled after "Byebye, Daddy".

6. All broke.

(Andrew) A pivot + open construction. The emphasis might be on "all" with the sense of completion or totality with "broke" a modifier (stative) of "all". Alternatively the pattern may be similar to the adult relation where "all" would be the modifier (quantifier) of the sentence reference "broke".

7. No bed.

(Andrew) A pivot + open construction. "Bed" (in this context presumably meaning the state of being put to bed) is the sentence reference with "no" a modifier (negative).

8. Baby highchair.

(Brown and Bellugi 1964) An open + open construction. "Baby" is

the sentence reference and "highchair" a modifier (locative).

(The mother's expansion was "Baby is in the highchair.")

9. Eve lunch.

(Brown and Bellugi) An open + open construction. The apparent meaning is "Eve is eating lunch." Which of the words is the sentence reference and which the modifier is a guess.

10. A coat.

(Brown and Bellugi) A pivot + open construction. Coat is the sentence reference and "a" some sort of modifier (perhaps if "a" is semantically empty it should be regarded as a noun marker.)

TABLE 2

INSTRUCTOR'S GRAMMAR

| | | | | |
|--------|---|--------|-----|----|
| S | = | NP | AUX | VP |
| S | = | NP | VP | |
| S | = | IMP | | |
| IMP | = | VI | | |
| IMP | = | VT | NP | |
| NP | = | PROPER | | |
| NP | = | DET | N | |
| NP | = | DEM | N | |
| NP | = | PERSPR | | |
| DET | = | THE | | |
| DET | = | A | | |
| DEM | = | THIS | | |
| DEM | = | THAT | | |
| PROPER | = | DAD | | |
| PROPER | = | MOM | | |
| PERSPR | = | HE | | |
| PERSPR | = | IT | | |
| N | = | TOY | | |
| N | = | CHAIR | | |
| N | = | APPLE | | |
| N | = | BOY | | |
| N | = | THING | | |
| N | = | MAN | | |
| VP | = | VI | | |
| VP | = | VT | NP | |
| VP | = | BE | ADJ | |
| AUX | = | CAN | | |
| AUX | = | MUST | | |
| AUX | = | WILL | | |
| ADJ | = | GONE | | |
| ADJ | = | GREEN | | |
| ADJ | = | TIRED | | |
| ADJ | = | BROKEN | | |
| ADJ | = | HOT | | |
| VI | = | SEE | | |
| VI | = | READ | | |
| VI | = | EAT | | |
| VI | = | SLEEP | | |
| VT | = | WANT | | |
| VT | = | EAT | | |
| VT | = | READ | | |
| VT | = | FIND | | |

TABLE 3

AGRAMMATICAL SENTENCES

1. *chair dad
8. *mom see
12. *want see
16. *the apple dad
20. *the read
30. *thing see
34. *see mom
37. *he see
41. *dad be tired it
44. *eat he he
50. *want mom
57. *toy want
63. *find dad a toy
80. *that chair will
81. *see see
82. *that apple
95. *man sleep
97. *see must
102. *this man be tired the
111. *want a toy sleep
113. *want dad thing

- 117. *a chair must
- 118. *dad chair

APPENDIX B

COMPUTATIONAL DISCUSSION

The particular machine configuration on which the model described in Chapter V was executed was an IBM 7044 (32K memory). The programs were written in FORTRAN IV, although some functions and character manipulation routines were coded in MAP. SLIP, a list processing system created by Weizenbaum (1963), was modified to run on the 7044, and has been used extensively throughout the model.

As this report is not intended to be a program description, the computational processes and programs used will not be discussed except for the parsing algorithm, which might be of some theoretical interest to those who work on string analysis procedures.

One of the requirements on a parsing algorithm entailed by the model was that it be able to parse a sentence only partially and produce an analysis without necessarily accounting for every word in that sentence. Clearly children do this, and it was necessary that the model be able to do so as well. If the problem were simply

that the child ignored words in a sentence which he did not understand (words for which he had no lexical entry), then an existing parsing algorithm could be easily modified to delete any undefined words from a string before parsing. But the problem is more complex than that. Consider the sentence, "The boy, who went to town yesterday, is swimming." A child who had not mastered relative clauses might well understand this sentence as "The boy is swimming" in spite of the fact that the meaning of all or nearly all of the words in the relative clause was known. Somehow he is able to skip over meaningful words to produce a sentence interpretation. This is just the problem in its most general terms: to determine if a string has one or more analyses when arbitrary substrings are ignored.

The other novel entailment on the parsing algorithm by the model is that analyses be presented to the comparator component in sequence according to their degree of confirmation. What this means in terms of the algorithm to be described is that there must be a method for computing the value of a sentence analysis given the values of the specific categorial assignments of the elements in the string and the values of the rules (or hypothesized rules) of the grammar. Once values are

assigned to sentence analyses, the analyses may easily be ordered for sequential presentation to the comparator component.

The algorithm used is an extension of one first described by Kay (1967). The scheme works from bottom to top and from right to left (although it could just as well work from left to right). It differs from the Cocke logic (Kay 1966) which is basically just bottom to top, there being no significant right-to-left effect. The principal advantage of the Kay algorithm, if one is concerned only with context-free grammars, is that it can easily handle non-binary rules, whereas the generalization of the Cocke algorithm to handle n-ary rules appears to be quite a bit more complex.

There are two ways in which the algorithm has been extended: (1) to allow for the simultaneous computation of node values during parsing; and (2) to allow for parsing a string with an incomplete grammar. These extensions will be discussed in turn without, however, describing the original algorithm, for which one should see Kay (1967).

B1 Node Values

The aim is to compute for every S analysis of a sentence a single number that reflects the degree to which

the grammatical constructs upon which the analysis is based are well confirmed. There are two factors that influence this computation: (1) the value associated with the categorial assignment of a word; and (2) the value associated with the rules that combine constituents into a constitute. If a particular categorial assignment of a word is well supported, the value associated with that assignment will be high. And if the combination of some number of constituents to form a higher level constitute is well supported, the value associated with the composition rule will be high.

The way in which node values are computed is fairly arbitrary; other computations might work just as well. The formula for computing the value V_x of a node X that directly dominates constituents y_1, y_2, \dots, y_n with respective values v_1, v_2, \dots, v_n that are combined with a rule $X \rightarrow y_1 + y_2 + \dots + y_n$ with associated value V_r is:

$$V_x = V_r \cdot \sum_{i=1}^n v_i$$

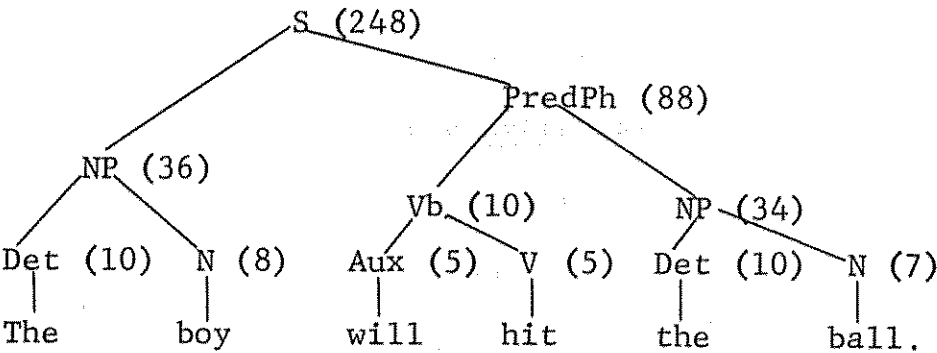
To illustrate, suppose the categorial assignments in the aforementioned sentence were valued as follows:

| <u>Word</u> | <u>Category</u> | <u>Value</u> |
|-------------|-----------------|--------------|
| the | Det | 10 |
| boy | N | 8 |
| will | Aux | 5 |
| hit | V | 5 |
| the | Det | 10 |
| ball | N | 7 |

And there were rules as follows:

| <u>Rule</u> | <u>Value</u> |
|------------------------------|--------------|
| $NP \rightarrow Det + N$ | 2 |
| $Vb \rightarrow Aux + V$ | 1 |
| $PredPh \rightarrow Vb + NP$ | 2 |
| $S \rightarrow NP + PredPH$ | 2 |

Then the values associated with each node in the analysis would be:



B2 Parsing With an Incomplete Grammar

A common characteristic of parsing algorithms is that successful analysis of a string means taking into account every word in that string. Unless all words in the string can be combined into a sentence by rules of the grammar, the analysis fails. The aim in this scheme is to allow the parsing algorithm to skip over a word or words where necessary in the analysis. Before describing the computational procedure itself, the term "incomplete sentence analysis" is defined.

Definition: A sequence $\phi' = \mu_1, \mu_2, \dots, \mu_m$ is a subsequence of $\phi = \alpha_1, \alpha_2, \dots, \alpha_n$ if there is an identity mapping from the members of the sequence ϕ' into the members of the sequence ϕ such that $\mu_i \rightarrow \alpha_j$ (and $\mu_i = \alpha_j$) and $\mu_k \rightarrow \alpha_p$ (and $\mu_k = \alpha_p$) and $i < k$ implies $j < p$.

Definition: We will say that there is an incomplete $\# S \#$ analysis of a string of ϕ if there is a ϕ' which is a subsequence of ϕ and there is a $\# S \#$ analysis, in the usual sense, of ϕ' .

With each incomplete $\# S \#$ analysis we will associate a number (its deletion coefficient) equal to $n - m$ (the number of words deleted in the full string).

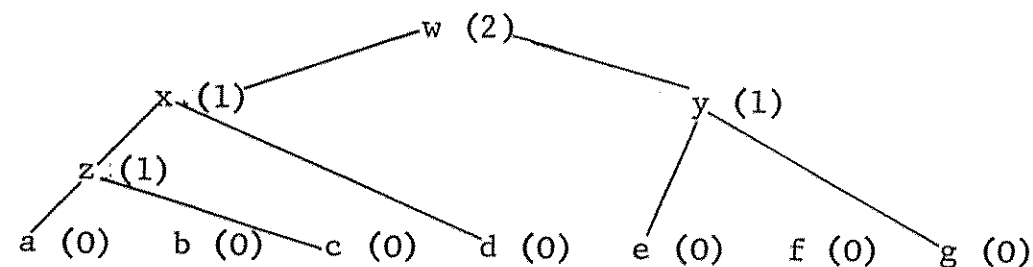
In general we will be interested in the incomplete $\# S \#$ analyses with small deletion coefficients, as we want to "account for" as much of the input string as possible. The scheme outlined here will produce all $\# S \#$ analyses of a string first, and then, if directed to continue, will go on to produce all incomplete $\# S \#$ analyses of a string with deletion coefficient of 1, and then those with a deletion coefficient of 2, and so on. This is accomplished by associating with each constitute (i) the deletion coefficient of that constitute, and (ii) the beginning and endpoint of that constitute. Taking the latter first, the beginning and endpoint of a word α_i in the sequence $\alpha_1, \alpha_2, \dots, \alpha_1, \dots, \alpha_n$ are both equal to i . When a constitute $X = C_1 + C_2 + \dots + C_n$ is formed its beginning point (b_X) equals the beginning point of C_1 (b_1) and its endpoint (e_X) equals the endpoint of C_n (e_n).

The deletion coefficient of a terminal symbol is zero, while the deletion coefficient of a constitute is computed as follows: Given a sequence of symbols $\phi = A_1, A_2, \dots, A_n$ (where the symbols may be terminal or non-terminal) and a composition rule $X \rightarrow C_1 + C_2 + \dots + C_m$

such that C_1, C_2, \dots, C_m is a subsequence of \emptyset , and where the subsequence has associated deletion coefficients d_1, d_2, \dots, d_m , beginning points b_1, b_2, \dots, b_m , and end points e_1, e_2, \dots, e_m , then the deletion coefficient of the constitute X (d_X) is defined:

$$d_X = \sum_{i=1}^m d_i + \sum_{i=1}^{m-1} (b_{i+1} - e_i - 1)$$

That is, the deletion coefficient of a constitute is equal to (i) the sum of the deletion coefficients of its constituents plus (ii) the length of any gaps that occur between its constituents. The following example illustrates the computation of deletion coefficients (shown in parentheses):



The algorithm first parses a string as completely as possible by forming all constituents with deletion coefficients of zero; during the next cycle it will form all constitutes with deletion coefficients of one, perhaps using the results obtained by the prior cycle; and so on. The algorithm was designed for use in an interactive

environment so that as analyses are obtained (in sequence by increasingly large deletion coefficients) they can immediately be evaluated, and the parsing process terminated as soon as a correct analysis found. This logic minimizes the computing time necessary to obtain a sentence parsing. Even so, however, the algorithm appears to be inherently impractical for long strings where the first acceptable analysis has a large deletion coefficient. In cases such as this there is apparently no way of using an exhaustive procedure; shortcuts and heuristics are necessary.

APPENDIX C

A PROGRAM TRACE

SPOKEN SENTENCE

COMPREHENDED SENTENCE

| | | | |
|------------------------------|-------|--|----------------|
| 1. THAT CHAIR WILL READ | CHAIR | C-REF. CHAIR WORD ACQUIRED CONTEXTUALLY | CHAIR THING |
| 2. MOM SEE | SEE | C-REF. SEE WORD ACQUIRED CONTEXTUALLY | SEE THING |
| 3. THIS MAN CAN SLEEP | SLEEP | C-REF. SLEEP WORD ACQUIRED CONTEXTUALLY | SLEEP THING |
| 4. WANT THAT THING | WANT | C-REF. WANT WORD ACQUIRED CONTEXTUALLY | WANT THING |
| 5. A BOY MUST READ DAD | READ | C-REF. READ WORD ACQUIRED CONTEXTUALLY | READ THING |
| 6. THAT BOY WILL BE TIRED | BOY | C-REF. BOY WORD ACQUIRED CONTEXTUALLY | BOY THING |
| 7. EAT HE | EAT | C-REF. EAT WORD ACQUIRED CONTEXTUALLY | EAT THING |
| 8. IT WANT DAD | WANT | C-REF. WANT WORD ACQUIRED CONTEXTUALLY | WANT THING |
| 9. FIND THIS BOY | BOY | C-REF. BOY WORD ACQUIRED CONTEXTUALLY | BOY THING |
| 10. FIND IT | FIND | C-REF. FIND WORD ACQUIRED CONTEXTUALLY | FIND THING |
| 11. DAD SEE | DAD | C-REF. DAD WORD ACQUIRED CONTEXTUALLY | DAD THING |
| 12. A APPLE WILL BE GONE | APPLE | C-REF. APPLE WORD ACQUIRED CONTEXTUALLY | APPLE THING |
| 13. SLEEP | SLEEP | C-REF. SLEEP WORD ACQUIRED CONTEXTUALLY | SLEEP THING |
| 14. THAT MAN WANT HE | WANT | C-REF. WANT WORD ACQUIRED CONTEXTUALLY | WANT THING |
| 15. WANT THE TOY | WANT | C-REF. WANT WORD ACQUIRED CONTEXTUALLY | WANT THING |
| 16. A BOY CAN BE BROKEN | BOY | C-REF. BOY WORD ACQUIRED CONTEXTUALLY | BOY THING |
| 17. THIS TOY MUST WANT A TOY | TOY | C-REF. TOY WORD ACQUIRED CONTEXTUALLY | TOY THING |
| 18. READ HE | READ | C-REF. READ WORD ACQUIRED CONTEXTUALLY | READ THING |

19. HE RE HOT
HOT
C-REF. HOT THING
WORD ACQUIRED CONTEXTUALLY HOT
EAT
C-REF. EAT THING
DICTIONARY
WORD FEATURES
.....
APPLE A ... WILL
THING
BOY A ...
THAT ...
THIS ...
... CAN
... WILL
... //
THING
DAD // ... SEE
THING
EAT // ... HE
... //
THING
FIND // ... IT
... THING
HOT BE ... //
... THING
READ // ... HE
... THING
SLEEP // ... //
... THING

GRAMMAR
.....

S = THING C-REF.

- 21. A MAN EAT MOM
- 22. HE READ THE TOY
- 23. A TOY SLEEP
- 24. THIS BOY CAN READ A CHAIR
- 25. THE CHAIR BE HOT
- 26. DAD MUST BE HOT
- 27. MOM WILL READ
- 28. THIS BOY MUST BE BROKEN
- 29. THAT BOY BE HOT
- 30. READ

FUNCTIONS NOW UNDERSTOOD BY THE CHILD ARE 'CONCRETE REFERENCE' AND 'MODIFIER'

- 31. THAT BOY CAN READ
BOY READ
C-REF. BOY #1
MODIF BOY THING
WORD ACQUIRED CONTEXTUALLY READ ACTION
- 32. THE BOY MUST SEE
BOY
C-REF. BOY #1
BOY THING
- 33. IT WILL BE GONE
GONE
C-REF. GONE #1
WORD ACQUIRED CONTEXTUALLY GONE THING
- 34. SLEEP

35. A APPLE EAT THE BOY
APPLE EAT
C-REF. APPLE #1
MODIF EAT THING
WORD ACQUIRED CONTEXTUALLY EAT ACTION

36. THAT MAN WILL SLEEP
MAN SLEEP
C-REF. MAN #1
MODIF MAN THING
SLEEP ACTION

37. IT SEC
38. FIND THAT MAN
MAN
C-REF. MAN #1
MODIF MAN THING

39. DAD FIND THIS BOY
CAC
C-REF. DAD #1
MODIF DAD THING

40. A APPLE EAT THAT TOY
APPLE EAT
C-REF. APPLE #1
MODIF APPLE THING
EAT ACTION

DICTIONARY

WORD FEATURES
.....

APPLE A ...
... EAT
... ACTION
THING

BOY THAT ...
THIS ...
... CAN
... MUST
THING

DAD // ...
... FIND
... MUST
THING

FAT APPLE ...
... THAT
... #1 ...
THING ...
ACTION
THING

GONE RE ...
... //
THING

HOT RE ...
... //
THING

MAN THAT ...
... WILL
... //
... ACTION
THING

MOM // ...
... WILL
THING

READ CAN ...
// ...
... HE
... //
#1 ...
THING ...
ACTION
THING

SLEEP WILL ...
// ...
... //
#1 ...
THING ...
ACTION
THING

TOY A ...
THE ...
... SLEEP
... //
THING

WANT MAN ...
// ...
... HE
... THE
THING

GRAMMAR
.....

#1 = THING C-REF.
S = #1 ACTION * MODIF
S = #1 *

41. IT WILL EAT
42. SLEEP
43. THE CHAIR CAN BE BROKEN

44. WANT DAD

45. IT READ
46. DAD BE TIRED

47. HE MUST SLEEP
48. WANT THAT BOY

49. THE THING WILL EAT THIS MAN

50. HE BE GREEN

51. DAD WILL READ DAD

52. DAD READ

53. READ MOM

54. THE THING MUST EAT

CHAIR CHAIR #1 THING
C-REF. CHAIR #1 THING
DAD DAD #1 THING
DAD TIRED
DAD TIRED #1 THING
MODIF DAD DAD THING
C-REF. TIRED #1 THING
WORD ACQUIRED CONTEXTUALLY TIRED
WANT BOY WANT BOY #1 ACTION
C-REF. BOY #1 THING
MODIF THING #1 THING
C-REF. MAN #1 THING
WORD ACQUIRED CONTEXTUALLY THING
GREEN GREEN #1 THING
C-REF. GREEN THING
WORD ACQUIRED CONTEXTUALLY GREEN
DAD READ DAD #1 THING
MODIF DAD READ ACTION
DAD READ DAD #1 THING
C-REF. DAD THING
MODIF DAD READ ACTION
READ MOM READ MOM #1 ACTION
C-REF. MOM #1 THING
THING EAT THING #1 THING
MODIF EAT ACTION

55. WANT IT
56. A MAN READ THIS CHAIR
57. THIS TOY MUST WANT THAT TOY
58. A BOY BE GONE
59. THIS CHAIR SLEEP
60. HE READ

READ CHAIR READ CHAIR #1 ACTION
C-REF. CHAIR #1 THING
WANT TOY WANT TOY #1 ACTION
C-REF. TOY #1 THING
BOY BOY #1 THING
CHAIR SLEEP CHAIR #1 THING
MODIF CHAIR SLEEP ACTION

DICTIONARY

WORD FEATURES
.....
APPLE A ... EAT
... EAT
... ACTION
THING
BOY THAT ...
THING
CHAIR THE ...
THIS ...
... CAN
... SLEEP
... //
ACTION ...
... ACTION
THING
DAD // ...
... ACTION
THING
EAT APPLE ...
#1 ...
THING ...
ACTION
THING

GONE BE ...
... //
THING

GREEN BE ...
... //
THING

HOT BE ...
... //
THING

MAN THAT ...
THIS ...
... WILL
... //
THING ...
... ACTION
THING

MOM READ ...
// ...
... WILL
... //
ACTION ...
THING

READ DAD ...
MAN ...
WILL ...
// ...
... DAD
... MOM
... THIS
... //
#1 ...
THING ...
... #1
... THING
ACTION
THING

SLEEP CHAIR ...
WILL ...
... //
#1 ...
THING ...
ACTION
THING

GRAMMAR
.....

#1 = THING
S = #1 ACTION
S = #1
S = ACTION #1

C-REF.
* MODIF
*
MODIF *

61. HE CAN READ DAD

62. EAT THE TOY

63. DAD MUST BE TIRED

64. FIND MOM

| | |
|-------|---|
| THING | THE MUST ... WILL ... ACTION ... #1 ... THING THING |
| TIRED | BE // THING ... THING |
| TOY | A ... THAT // ACTION ... THING |
| WANT | MUST ... // THAT ... #1 ... THING ACTION THING |

| | | |
|--------|----------|--------------|
| MODIF | READ | ACTION |
| | DAD | #1 |
| C-REF. | DAD | THING |
| MODIF | EAT | ACTION |
| | TOY | #1 |
| C-REF. | TOY | THING |
| C-REF. | DAD | #1 |
| | DAD | THING |
| MODIF | FIND | ACTION |
| | MOM | #1 |
| C-REF. | MOM | THING |
| WORD | ACQUIRED | CONTEXTUALLY |
| | FIND | |

| | | | | |
|----------------------------|--------------|--------|------------|-------------|
| 65. SLEEP | TOY MAN | MODIF | TOY MAN | THING #1 |
| 66. A TOY WANT THE MAN | | C-REF. | MAN | THING |
| 67. THIS MAN EAT | MAN EAT | | MAN | #1 |
| | | C-REF. | MAN | THING |
| | | MODIF | EAT | ACTION |
| 68. DAD WILL EAT A BOY | EAT BOY | | EAT | ACTION |
| | | MODIF | BOY | #1 |
| | | C-REF. | BOY | THING |
| 69. READ THE BOY | READ BOY | | READ | ACTION |
| | | MODIF | BOY | #1 |
| | | C-REF. | BOY | THING |
| 70. DAD CAN WANT THAT TOY | WANT TOY | | WANT | ACTION |
| | | MODIF | TOY | #1 |
| | | C-REF. | TOY | THING |
| 71. THIS TOY MUST BE GREEN | TOY GREEN | | TOY | THING |
| | | MODIF | GREEN | #1 |
| | | C-REF. | GREEN | THING |
| 72. DAD MUST READ HE | DAD READ | | DAD | #1 |
| | | C-REF. | DAD | THING |
| | | MODIF | READ | ACTION |
| 73. A TOY CAN WANT DAD | WANT DAD | | WANT | ACTION |
| | | MODIF | DAD | #1 |
| | | C-REF. | DAD | THING |
| 74. FIND HE | GREEN | | GREEN | #1 |
| 75. IT BE GREEN | | C-REF. | GREEN | THING |
| 76. A CHAIR BE BROKEN | CHAIR BROKEN | | CHAIR | THING |
| | | MODIF | BROKEN | #1 |
| | | C-REF. | BROKEN | THING |
| 77. WANT THAT TOY | WANT TOY | | WANT | ACTION |
| | | MODIF | TOY | #1 |
| | | C-REF. | TOY | THING |
| 78. A TOY SLEEP | TOY SLEEP | | TOY | #1 |
| | | C-REF. | TOY | THING |
| | | MODIF | SLEEP | ACTION |

| | | | | |
|-----------------------|----------|-----------|------------|-------------|
| 79. THAT TOY WILL SEE | TOY SEE | | TOY TOY | #1 THING |
| | | C-REF. | SEE | ACTION |
| 80. MOM MUST BE GONE | MOM GONE | | MOM | THING |
| | | MODIF | GONE | #1 |
| | | C-REF. | GONE | THING |
| DICTIONARY | | | | |
| WORD FEATURES | | | | |
| | | | | |
| | APPLE | A ... | ... | EAT |
| | | ... | ... | ACTION |
| | | ... | ... | THING |
| | BOY | ... | ... | ... |
| | | ... | ... | ... |
| | | ... | ... | ... |
| | BROKEN | BE ... | ... | ... |
| | | ... | ... | ... |
| | | ... | ... | ... |
| | CHAIR | A ... | ... | ... |
| | | THIS ... | ... | ... |
| | | ... | ... | ... |
| | | ... | ... | ... |
| | | ... | ... | ... |
| | DAD | // ... | ... | ... |
| | | ... | ... | ... |
| | | ... | ... | ... |
| | EAT | #1 ... | ... | ... |
| | | THING ... | ... | ... |
| | | ... | ... | ... |
| | | ... | ... | ... |
| | | ... | ... | ... |

GONE BE ...
... //
THING ...
THING

GREEN BE ...
... //
THING ...
THING

HOT BE ...
... //
THING

MAN THAT ...
THIS ...
... //
THING ...
... ACTION
THING

MOM // ...
... MUST
... //
ACTION ...
... #1
... THING
THING

READ // ...
... DAD
... //
#1 ...
THING ...
... #1
... THING
ACTION
THING

SEE WILL ...
... //
#1 ...
THING ...
ACTION

SLEEP CHAIR ...
TOY ...
WILL ...
... //
#1 ...
THING ...
ACTION
THING

GRAMMAR
.....

#1 = T-ING C-REF.
S = ACTION #1 MODIF *
S = #1 ACTION * MODIF
S = #1 *
S = T-ING #1 MODIF *

THE FUNCTIONAL RELATIONSHIP 'SUBJECT OF A SENTENCE' IS NOW AVAILABLE TO THE CHILD

81. SEE
82. THAT THING RE TIRED

THING TIRED
MODIF THING TIRED #4
C-REF. TIRED #1
THING

83. FIND HE
84. A CHAIR WILL RE GREEN

CHAIR GREEN
MODIF CHAIR GREEN #4
C-REF. GREEN #1
THING

85. DAD WANT HE

DAD WANT #2
DAD #1
C-REF. DAD THING
MODIF WANT ACTION

THING THE ...
... MUST
... WILL
... ACTION
... #1
... THING
THING

TIRED RE ...
... //
THING ...
THING

TOY THAT ...
... //
ACTION ...
THING

WANT CAN ...
// ...
... THAT
... #1
... THING
ACTION

86. WANT THE MAN
WANT MAN
MODIF MAN #3 ACTION
C-REF. MAN #1 THING

87. FIND THE BOY
FIND BOY #3 ACTION
BOY #1 THING
C-REF. BOY THING

88. MOM MUST EAT
MOM EAT #2
MOM #1 THING
C-REF. MOM THING ACTION
MODIF EAT ACTION

89. THIS APPLE EAT THE THING
APPLE EAT THING
EAT THING #3
MODIF EAT ACTION
C-REF. THING #1 THING

90. MOM CAN WANT THE TOY
MOM WANT TOY #3 THING
MODIF WANT TOY #3 ACTION
C-REF. TOY #1 THING

91. A THING EAT DAD
THING EAT DAD
SUBJECT THING THING
MODIF EAT DAD #3
C-REF. EAT DAD #1 THING

92. THAT CHAIR MUST BE TIRED
CHAIR TIRED #4
MODIF CHAIR THING
C-REF. TIRED #1 THING

93. MOM MUST WANT THIS TOY
MOM WANT TOY #3 THING
MODIF WANT TOY #3 ACTION
C-REF. TOY #1 THING

94. READ THIS TOY
READ TOY #3
MODIF READ ACTION
C-REF. TOY #1 THING

95. FIND THIS THING
FIND THING #3
MODIF FIND THING #1 ACTION
C-REF. THING THING

96. A CHAIR MUST BE TIRED
CHAIR TIRED #4
MODIF CHAIR THING
C-REF. TIRED #1 THING

97. FIND IT
98. IT CAN BE BROKEN
BROKEN #1
C-REF. BROKEN THING

99. READ
100. MOM FIND DAD
MOM FIND DAD
SUBJECT MOM THING
MODIF FIND DAD #3
C-REF. DAD #1 ACTION
DAD THING

DICTIONARY

WORD FEATURES
.....

APPLE A ...
THIS ...
... EAT
... ACTION
... #3
THING

BOY ... //
ACTION ...
THING

BROKEN BE ...
... //
THING ...
THING

CHAIR A ...
... MUST
... #1
... THING
THING

104. READ
105. SLEEP
106. READ
107. EAT
108. THIS MAN MUST FIND THAT APPLE

109. THIS APPLE FIND THAT MAN

110. WANT THAT MAN

111. IT WILL SLEEP
112. IT SLEEP
113. MOM BE BROKEN

114. WANT DAD

115. EAT HE
116. FIND HE
117. EAT THIS APPLE

118. A THING FIND THE BOY

119. WANT A MAN

| | | | |
|----------------|---------|---------------|--------|
| MAN FIND APPLE | SUBJECT | MAN | THING |
| | MODIF | FIND APPLE #3 | ACTION |
| | C-REF. | APPLE #1 | THING |
| APPLE FIND MAN | SUBJECT | APPLE | THING |
| | MODIF | FIND MAN #3 | ACTION |
| | C-REF. | MAN #1 | THING |
| WANT MAN | MODIF | WANT MAN #3 | ACTION |
| | C-REF. | MAN #1 | THING |
| MOM BROKEN | MODIF | MOM BROKEN #4 | THING |
| | C-REF. | BROKEN #1 | THING |
| WANT DAD | MODIF | WANT DAD #3 | ACTION |
| | C-REF. | DAD #1 | THING |
| EAT APPLE | MODIF | EAT APPLE #3 | ACTION |
| | C-REF. | APPLE #1 | THING |
| THING FIND BOY | C-REF. | THING #1 | THING |
| | SUBJECT | FIND | ACTION |
| | MODIF | BOY | THING |
| WANT MAN | MODIF | WANT MAN #3 | ACTION |
| | C-REF. | MAN #1 | THING |

120. WANT IT

| DICTIONARY | |
|------------|--|
| WORD | FEATURES |
| | |
| APPLE | A ... THIS // ACTION ACTION ... #3 THING |
| BOY | ... // ACTION ... THING |
| BROKEN | BE // THING ... THING |
| CHAIR | A MUST ... #1 ... THING THING |
| DAD | // ... THING |
| EAT | #1 ... THING #1 ... THING ACTION |
| FIND | // THAT THING #1 ... THING ACTION |

GONE BE ...
... //
THING ...
THING

GREEN BE ...
... //
THING ...
THING

HOT BE ...
... //
THING

MAN ... //
ACTION ...
THING

MOM // ...
... MUST
... ACTION
... #3
THING

READ // ...
#1 ...
THING ...
... #1
... THING
ACTION
THING

THING A ...
... ACTION
THING

TIRED BE ...
... //
THING ...
THING

TOY ... //
ACTION ...
THING

WANT ... THAT
... #1
... THING
ACTION

GRAMMAR
.....

#1 = THING
#2 = #1 ACTION
#3 = ACTION #1
#4 = TING #1
S = #3
S = THING #3
S = #4
S = #1
S = #2

C-REF.
* MODIF
MODIF #
MODIF #
* SUBJECT #
*
*
*

121. A CHAIR READ DAD

CHAIR READ DAD

| | | |
|---------|----------|--------|
| SUBJECT | CHAIR | THING |
| MODIF | READ DAD | #3 |
| C-REF. | READ | ACTION |
| | DAD | #1 |
| | DAD | THING |

BIBLIOGRAPHY

- Bellugi, U. 1964. The emergence of inflection and negation systems in the speech of two children. Paper presented at the New England Psychological Association Convention.
- , and R. Brown (eds.) 1964. The Acquisition of Language. Monograph of the Society for Research in Child Development. 29.
- Berko, J. 1958. The child's learning of English morphology. WORD 14.150-77.
- Bever, T.G., J.A. Fodor, and W. Weksel, 1965a. On the acquisition of syntax: a critique of "contextual generalization". Psychological Review 72.467-82.
- 1965b. Is linguistics empirical? Psychological Review 72.493-500.
- Black, M. 1962. Models and Metaphors. Ithaca, New York: Cornell University Press.
- Braine, M.D.S. 1963a. On learning the grammatical order of words. Psychological Review 70.323-48.
- 1963b. The ontogeny of English phrase structure: The first phase. Language 39.1-13.
- 1965a. On the basis of phrase structure: A reply to Bever, Fodor, and Weksel. Psychological Review 72.483-92.
- 1965b. The insufficiency of a finite state model for verbal reconstructive memory. Psychonomic Science 2.291-2.
- 1965c. Inferring a grammar from responses: discussion of Gough and Segal's comment. Psychonomic Science 3.241-2.
- 1966a. Learning the position of words relative to a marker element. Journal of Experimental Psychology 72.532-40.

- Braine, M.D.S. 1966b. The acquisition of language in infant and child. Mimeographed. To appear in Carroll Reed (ed.) Learning of Language. The National Council of Teachers of English.
- Brown, R. 1957. Linguistic determinism and the part of speech. The Journal of Abnormal and Social Psychology 55.1-5.
- , and C. Fraser. 1963. The acquisition of syntax. Verbal Behavior and Learning, ed. by C. Cofer and B. Musgrave, 158-209. New York: McGraw Hill.
- , C. Fraser, and U. Bellugi. 1964. Explorations in grammar evaluation. The Acquisition of Language, ed. by U. Bellugi and R. Brown, 79-92. Monograph of the Society for Research in Child Development. 29.
- , and U. Bellugi. 1964. Three processes in the child's acquisition of syntax. Harvard Educational Review 34.133-51. Reprinted in Lenneberg (1964).
- Cazden, C.B. 1965. Environmental Assistance to the Child's Acquisition of Grammar. Unpublished thesis, Graduate School of Education, Harvard University.
- 1967. The role of parent speech in the acquisition of grammar. Paper presented at the Project Literacy Conference, Cambridge, Mass.
- Chomsky, N. 1959. Review of Verbal Behavior, by B.F. Skinner. Language 35.26-58.
- 1962. Explanatory models in linguistics. Logic, Methodology and Philosophy of Science, ed. by E. Nagel, P. Suppes, and A. Tarski, 528-50. Stanford, California: Stanford University Press.
- 1965. Aspects of the Theory of Syntax. Cambridge, Mass.: M.I.T. Press.
- Cofer, C.N., and B.S. Musgrave (eds.) 1963. Verbal Behavior and Learning. New York: McGraw Hill.

- Ervin, S.M. 1964. Imitation and structural change in children's language. New Directions in the Study of Language, ed. by E. Lenneberg, 163-89. Cambridge, Mass.: M.I.T. Press.
- Esper, E.A. 1925. A technique for experimental investigation of associative interference in artificial linguistic material. Language Monograph 1.
- Fodor, J.A. 1966. How to learn to talk: some simple ways. The Genesis of Language, ed. by F. Smith and G. Miller, 105-128 (including discussion). Cambridge, Mass.: M.I.T. Press.
- Fraser, C., U. Bellugi, and R. Brown. 1963. Control of grammar in imitation, comprehension, and production. Journal of Verbal Learning and Verbal Behavior 2.121-35.
- Gough, P., and E. Segal, 1965. Comment on "the insufficiency of a finite state model for verbal reconstructive memory". Psychonomic Science 3.155-6.
- Gruber, J.S. 1967. Topicalization in child language. Foundations of Language 3.37-65.
- Harris, Z.S. 1951. Methods in Structural Linguistics. Chicago: University of Chicago Press.
- Hunt, E. 1965. Selection and reception conditions in grammar and concept learning. Journal of Verbal Learning and Verbal Behavior 4.211-5.
- Hunt, J. McV. 1961. Intelligence and Experience. New York: Ronald Press.
- Jenkins, J.J. 1965. Mediation theory and grammatical behavior. Directions in Psycholinguistics, ed. by S. Rosenberg, 66-96. New York: MacMillan.
- , and D.S. Palermo. 1964. Mediation processes and the acquisition of linguistic structure. The Acquisition of Language, ed. by U. Bellugi and R. Brown, 141-76 (including discussion). Monograph of the Society for Research in Child Development. 29.

- Johnson, N.F. 1962. The Cue Value of Sentence Frames for the Acquisition of Speech Categories. Unpublished doctoral dissertation, University of Minnesota.
- Katz, J.J. and J.A. Fodor. 1963. The structure of a semantic theory. Language 39.170-210.
- Kay, M. 1966. The Tabular Parser: A Parsing Program for Phrase Structure and Dependency. RM-4933. Santa Monica, Calif.: RAND Corp.
- 1967. Parsing with Transformational Grammar. (Forthcoming). Santa Monica, Calif.: RAND Corp.
- Lenneberg, E.H. 1962. Understanding language without ability to speak: a case report. Journal of Abnormal and Social Psychology 65.419-25.
- (ed.) 1964. New Directions in the Study of Language. Cambridge, Mass.: M.I.T. Press.
- McNeill, D. 1966a. Some universals of language acquisition. Studies in Language and Language Behavior II.
- 1966b. The creation of language. Studies in Language and Language Behavior III.
- 1966c. On theories of language acquisition. Studies in Language and Language Behavior III.
- 1966d. Developmental psycholinguistics. The Genesis of Language, ed. by F. Smith and G. Miller, 15-103 (including discussion). Cambridge, Mass.: M.I.T. Press.
- Menyuk, P. 1963. A preliminary evaluation of grammatical capacity in children. Journal of Verbal Learning and Verbal Behavior 2.429-39.
- Miller, G. 1967. Project grammarama. The Psychology of Communication, G. Miller. New York: Basic Books.

- Miller, G., E. Galanter, and K. Pribram. 1960. Plans and the Structure of Behavior. New York: Holt, Rinehart and Winston.
- , and M. Stein. 1963. Grammarama I. Preliminary Studies and Analysis of Protocols. Scientific Report CS-2. Center for Cognitive Studies, Harvard University.
- Miller, W., and S. Ervin. 1964. The development of grammar in child language. The Acquisition of Language, ed. by U. Bellugi and R. Brown, 9-42 (including discussion). Monograph of the Society for Research in Child Development. 29.
- Skinner, B.F. 1957. Verbal Behavior. New York: Appleton-Century-Crofts.
- Slobin, D.I. 1964. Preliminary outline of pilot studies of comprehension and speech in pre-school children. Project Literacy Reports 2.31-6. Cornell University.
- 1966a. The acquisition of Russian as a native language. The Genesis of Language, ed. by F. Smith and G. Miller, 129-152 (including discussion). Cambridge, Mass.: M.I.T. Press.
- 1966b. Abstracts of Soviet studies of child language. The Genesis of Language, ed. by F. Smith and G. Miller, 361-86. Cambridge, Mass.: M.I.T. Press.
- Smith, C.S. 1966. Two studies of the syntactic knowledge of young children. Paper presented at the Linguistic Colloquium, M.I.T.
- Smith, F., and G. Miller (eds.) 1966. The Genesis of Language. Cambridge, Mass.: M.I.T. Press.
- Smith, K.H. 1963. Recall of Paired Verbal Units under Various Conditions of Organization. Unpublished doctoral dissertation, University of Minnesota.
- Suci, G. 1967. The validity of pause as an index of units in language. Journal of Verbal Learning and Verbal Behavior 6.26-32.

Warner, H., and E. Kaplan. 1950. Development of word meaning through verbal context: an experimental study. Journal of Psychology 29.251-7.

Weir, R.H. 1962. Language in the Crib. Hague: Mouton.

Weissberg, P. 1963. Social and nonsocial conditioning of infant vocalizations. Child Development 34.377-88.

Weizenbaum, J. 1963. Symmetric list processor. Communications of the ACM 6.524-36.

Weksel, W. 1965. Review of The Acquisition of Language, U. Bellugi and R. Brown (eds.) Language 41.692-709.

Whitman, J., and W.R. Garner. 1963. Concept learning as a function of form of internal structure. Journal of Verbal Learning and Verbal Behavior 2.195-202.