DOCUMENT RESUME

ED 107 155 FL 006 932

AUTHOR	Wexler, Kenneth; And Others
TITLE	Learning: Theoretic Foundations of Linguistic Universals. Social Science Working Paper No. 60.
INSTITUTION	California Univ., Irvine. School of Social Sciences.
SPONS AGENCY	Office of Naval Research, Washington, D.C. Personnel and Training Research Programs Office.
PUB DATE	1 Nov 74
NOTE	73p.
EDRS PRICE	MF-$0.76 HC-$3.32 PLUS POSTAGE
DESCRIPTORS	Descriptive Linguistics; *Language Development; Language Patterns; *Language Universals; *Learning Processes; Learning Theories; *Linguistic Theory; Mathematical Linguistics; *Psycholinguistics; Semantics; Structural Analysis; Syntax; Transformation Generative Grammar
IDENTIFIERS	Freezing Principle; Invariance Principle

ABSTRACT

Some aspects of a theory of grammar are presented which derive from a formal theory of language acquisition. One aspect of the theory is a universal constraint on analyzability known as the Freezing Principle, which supplants a variety of constraints proposed in the literature. A second aspect of the theory is the Invariance Principle, a constraint on the relationship between semantic and syntactic structure that makes verifiable predictions of syntactic universals. The relationship between the notion of "explanatory adequacy" of a theory of grammar and the learnability of a class of transformational grammars is discussed. (Author)

ED107155

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report No. 2

2. GOVT ACCESSION NO.:

3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): Learning-Theoretic Foundations of Linguistic Universals

5. TYPE OF REPORT & PERIOD COVERED: Semi-Annual Technical Report (1 Oct 1973 - 31 Mar 1974)

6. PERFORMING ORG. REPORT NUMBER: Soc.Sci.Working Paper #60

7. AUTHOR(s): Kenneth Wexler, Peter Culicover, Henry Hamburger

8. CONTRACT OR GRANT NUMBER(s): N00014-69-A-0200-9006

9. PERFORMING ORGANIZATION NAME AND ADDRESS: School of Social Sciences, University of California, Irvine, CA 92664

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61153N; RR 042-06; RR 042-06-01; NR 154-362

11. CONTROLLING OFFICE NAME AND ADDRESS: Personnel and Training Research Programs, Office of Naval Research (Code 458), Arlington, VA 22217

12. REPORT DATE: November 1, 1974

13. NUMBER OF PAGES: 62

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Submitted to Theoretical Linguistics

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

linguistic universals, syntax, freezing principle, semantics, language acquisition, learning theory, mathematical linguistics

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Some aspects of a theory of grammar are presented which derive from a formal theory of language acquisition. One aspect of the theory is a universal constraint on analyzability known as the Freezing Principle, which supplants a variety of constraints proposed in the literature. A second aspect of the theory is the Invariance Principle, a constraint on the relationship between semantic and syntactic structure that makes verifiable predictions of

DD FORM 1473, 1 JAN 73 — EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ERIC Full Text Provided by ERIC

syntactic universals. The relationship between the notion of "explanatory adequacy" of a theory of grammar and the learnability of a class of transformational grammars is discussed.

ERIC Full Text Provided by ERIC

LEARNING-THEORETIC FOUNDATIONS OF LINGUISTIC UNIVERSALS

I. Introduction

A. General objectives

We have acheived results in the realm of explanatory adequacy, a subject which, in spite of its recognized centrality to linguistic theory, has been largely neglected. On the other hand, two interacting shorter-range goals have attracted considerably more attention from linguists. These are descriptive adequacy and formal universals. Given that grammars should consist of rules of certain forms, a linguist seeks a descriptively adequate grammar of a particular language, a description of adult competence. On the other hand (s)he may ask what forms rules should be allowed to take. This latter task can be approached by noting which kinds of rules seem to be universally useful for describing natural language. In this way, universal formalism may be advanced.

Suppose that a universal set of rule types and conditions is found which allows grammars to be constructed for many particular languages, and that these grammars provide adequate descriptions and even insightful generalizations about their respective languages. Even then, a puzzle remains: why these particular formal universals? Are they an accident, or do they have some special formal property which makes them particularly appropriate? Chomsky (1965) argues that there _is_ such a property which distinguishes among formal universals and that in particular it has to do with the fact that language must be learned by every child. He writes (page 25):

> To the extent that a linguistic theory succeeds in selecting a descriptively adequate grammar on the basis of primary linguistic data, we can say that it meets the condition of _explanatory adequacy_.

ERIC

We add to this requirement that the selection procedure be psychologically plausible.

Here we shall attempt to be both plausible and detailed in showing that the requirement of "learnability" can force a selection among formal universals. Further, this research has yielded the particularly interesting and unique result that a linguistic principle which was motivated by abstract developments in language acquisition turns out to provide an account of several adult syntactic structures which is descriptively more satisfactory than previous accounts. If validated, this would be an instance of the kind of scientific event in which a theoretical analysis leads to an improved empirical account. Thus it is appropriate and in fact important to proceed in this unified manner. Even if our linguistic analysis should ultimately require modification, we consider it worth explicating our work as one example of how one might go about achieving explanatory adequacy. A more detailed presentation of various parts of the theory with extensive discussion appears in various published and unpublished papers, and a complete presentation will appear in a book which is presently in preparation[1].

B. Fundamental theoretical background

The major goal of linguistic theory is to characterize human language in a way that is consistent with the fact that any child can learn any human language, provided that he is born into a community where that language is spoken. Thus our characterization of language must not call for a potential range or complexity of structures that would necessarily bewilder the child by virtue of being logically impossible to learn. To quote Chomsky, (1965, p. 58)

> It is, for the present, impossible to formulate an assumption about initial, innate structure rich enough to account for the fact that grammatical knowledge is attained on the basis of the evidence available to learner.... The real problem is that of developing a hypothesis about initial structure that is sufficiently rich to account for acquisition of language, yet not so rich as to be inconsistent with the known diversity of language.

This goal has never been approached, and, in fact, linguists have never seriously taken up the question of language learnability. Most of the work by linguists with regard to discovering the formal constraints on the structure of human language has been concerned with the inspection of languages and the subsequent positing of constraints or universals on the basis of such inspection. We will provide examples of such investigations as they relate to our own work in Section II below.

On the other hand, it is also possible to consider the question of linguistic constraints and universals by first establishing the requirements which a plausible learning theory (of language) places on the languages which it can learn. If a plausible learner cannot learn a given type of language, then this constitutes evidence either that the languages which we call "natural" languages are not of this type, or that some refinement is required in our notion of plausible learner.

It is demonstrable (Gold 1967) that if there are no constraints whatsoever on what kinds of grammars could be grammars of natural languages, then no conceivable learning procedure could guess, from data from the language, which one of the conceivable grammars was the grammar corresponding to that language.

In Hamburger and Wexler (1973a,b) and Wexler and Hamburger (1973) a model of a minimally plausible learner is constructed, and the question of the learnability of various types of languages is then investigated. It is shown that even if all human languages possessed the same deep structures,

ERIC

and differed only in the transformations which constituted their grammars, no conceivable learning procedure would be able to guess the correct grammar of any such language given data from that language in the form of grammatical sentences. Furthermore, it was demonstrated that a minimally plausible learning procedure can learn the grammar of a language if (a) the procedure is presented with the semantic interpretation of a sentence when the sentence is presented, and (b) if certain formal constraints are placed on the applicability of transformations. We will describe these results and possible extensions of them more fully in Section II below.

It follows from the work just mentioned that a theory of grammar learning is a theory of grammar in that a precise specification of the learner leads to a specification of the class of things that are learnable. Hence a correct specification of the procedure by which human beings learn the grammars of languages will lead to a specification of the class of possible human languages.

C. Methodology

A fundamental requirement of the theory is that the learning procedure be plausible. It is necessary, therefore, to append to a minimal learning procedure more sophisticated notions of memory, attention, self-correction, external correction, rate of learning, type of input, cognitive capacity, etc. Ideally, the plausible learner, should behave just like the child in an empirically defined language learning environment with respect to all these factors.

A second requirement of the theory is that for the constraints placed on the class of languages by the learning procedure, all the available phenomena from natural language support their adoption as constraints on natural languages. In fact, we wish to show that such constraints regularly produce the deepest and most compelling explanation (to the linguist) of the linguistic data. It is therefore of considerable importance to conduct a systematic investigation of well-known (and new) syntactic phenomena in natural language which might provide evidence in support of or in opposition to the precise constraints arising from the learning theory. Some work of this nature is described in Section III.

A third requirement of the theory is that the constraints arrived at, as well as the specification of the learning theory, be universal, and that all implications which arise from these specifications also be universal. In particular, we assume for the purposes of maintaining a plausible learning procedure that there exists a universal constraint on the relationship between semantic and syntactic structure. Assuming that semantic structure is universal, this leads to a number of predicted universals of syntactic structure. Hence we are also concerned with investigating a variety of the world's languages to determine the plausibility of such putative universals. We discuss this further in Section IV.

Finally, a requirement of the theory is that it make only correct predictions about the actual course of language development in the child. We have not constructed experimental situations in which such predictions are tested. Rather we are concerned w. .h the more primary task of constructing firm and falsifiable predictions, and seek to discover evidence which bears on them in the literature on developmental psycholinguistics. We discuss these questions in more detail in Section V.

II. Learnability Theory

A. Theories of language acquisition

A theory of (first) language acquisition defines a procedure which models the essential characteristics of how the child acquires his language. This procedure must be powerful enough to learn any natural human language, since we start with the fundamental observation that any normal child can learn any natural language, given the proper environment. That this requirement (of learnability) is difficult to attain is evident from the fact that no existing theory of language acquisition comes close to satisfying it.

By far the bulk of work in the study of language acquisition involves the description of the child's linguistic knowledge at various ages. From this work a number of interesting generalizations may be drawn about the child's language. But very little attention has been given to a dynamic theory; that is, a theory of how, given the input that is available to him, the child arrives at an adult's knowledge of language.

A few studies (an important one is Brown and Hanlon 1970) have asked the question: why does a child learn language? That is, what compels a child to change his grammar over time? Although very important, this question is only a part of the problem of the study of language acquisition. Even if we had an unequivocal answer to this question we would still not know what the procedure is which the child uses to construct his grammar. (That is, we would not know how a child learns his language).

When we come to those studies in the language acquisition literature which attempt to sketch a theory, that is those proposals which suggest a procedure, we find a number of proposals, but none of the proposals meet the first requirement stated above; that is, none of the theorists attempt to show that the procedure is strong enough to learn all human languages, given what we know about human language. In fact, the theories are either too vague for the question to be seriously asked, or they are clearly too weak to learn any substantial amount of syntax.

The common methodology which most of these studies of the theory of language acquisition adopt is to take some description of the speech of a child at an early age and to then hypothesize a way in which that speech could have been learned. This is true for example, of McNeill (1966) and Braine (1963). The correct description of children's knowledge of language at a given age is not easy to attain, and this can cause problems. Thus Braine (1963) outlines a theory of how a pivot grammar might be learned, but Bloom (1970) and Brown (1973) show quite clearly that pivot grammars are not appropriate models of children's language.

For the problem of learning transformations we find little help in the literature. Although the construction of an "evaluation procedure" is taken as a central goal of Linguistics, no linguist has offered a procedure and demonstrated that it can converge to a correct grammar. In the field of language acquisition, McNeill (1966) discusses the learning of transformations and offers a hypothesis (namely, that transformations reduce memory load) as to why they are acquired. But he offers no hypothesis about the procedure by which they are acquired, and, therefore, no proof that a given procedure is strong enough to learn language. Fodor (1966) recognizes the difficulty of the problem and suggests one strategy, which he claims might account for one very small part of the procedure wherein base structures are "induced" from surface strings, but no proof of success is given. Slobin (1973) suggests such "operating principles" as "pay attention to the order of words and morphemes", but no more explicit procedures nor outline of a proof

ERIC

of success are proposed. Braine (1971) offers some hints at a "discovery-procedures" model, and applies the model to some simple examples, but the model is certainly not strong enough to have success claimed for it. In most other studies (there are a large number of them--see Ferguson and Slobin 1973, for a bibliography), no hypotheses about learning procedures are suggested.

The field of computer simulation also provides little insight. Kelley (1967) has written a language learning program which deals with only the simplest stages of language acquisition and which makes no mention of transformations nor of the phenomena accounted for by transformations. The only grammatical hypotheses which his learner can make represent contingencies between adjacent elements in phrase-markers--far too weak to account for the learning of transformations. Also, as is common with simulation studies, it is not clear exactly what the program can do.

Klein and Kuppin (1970) have written a program to learn transformational grammar. The program is intended to be more a model of the linguistic field-worker than of the child learning a first language. Again, it is not clear what the program can learn. A few simple examples are given, but the range of the program is undefined. Indeed, the authors call the program "heuristic" because it does not guarantee success. It seems to us that heuristic (in this sense) programs might be acceptable as models of humans in situations where humans may, indeed, fail (say, problem solving, or the discovery of scientific theories, or writing a grammar as a field-worker for some foreign language, which, in fact, is Klein and Kuppin's situation), but the fundamental assumption in the study of language acquisition is that every normal child <u>succeeds</u>. Thus we must have what Klein and Kuppin call "algorithmic" procedures--ones for which success is guaranteed. (Note that Klein and Kuppin's sense of "heuristic" and "algorithmic" is not necessarily the sense in common usage in the field of artificial intelligence.)

Klein and Kuppin make a number of assumptions which would be quite implausible in models of a child learning a first language. First, they assume that the learner receives information about what strings are non-sentences. Although this information may be available to a field-worker, it is probably not available to a child (Brown and Hanlon 1970; Braine 1971; Ervin-Tripp 1971). Second, they assume that the learner can remember and use all data it has ever received. Third, each time the learner hypothesizes a new transformation it tests it extensively.

All these assumed capacities of the learner seem to be unavailable to the child. On the other hand, only obligatory, ordered transformations are allowed, so that the class of grammars is not rich enough to describe all natural languages. Still, there is no reason to believe that Klein and Kuppin's learner can learn an arbitrary grammar of the kind they assume.

Gold (1967) provided a formal definition of language learning and showed that according to this definition most classes of languages (including the finite state languages and thus any super-class of these such as the context-free languages) were not learnable if only instances of grammatical sentences were presented. Many of these language classes are learnable if "negative information", that is, instances of non-sentences, identified as such, are also presented. However, as noted above the evidence is that children do not receive such negative information. Any theory of language learning which depends heavily upon negative information will probably turn out to be incorrect and will very likely not yield insights on formal grammatical universals. With such a powerful input, what constraints actually exist will be unnecessary.

Other studies on grammar learning have been made by Feldman (1967, 1969), Feldman et al. (1969), and Horning (1969). These studies, while interesting in themselves, do not deal with the question of learning systems which linguists argue are necessary for natural language (e.g., transformations).

B. Formal results on learnability

The absence of linguistically relevant results in learnability theory led us to study the learnability of transformational grammars. Since each transformational grammar includes a phrase-structure grammar as a part of it, Gold's results would seem to preclude learnability from information consisting only of sentences. At this point there are two ways to proceed: either restrict the class of grammars or enrich the information. We will discuss each of these possibilities in turn.

The first approach (Wexler and Hamburger 1973) is to try to restrict the class of grammars to achieve learnability from the presentation of grammatical sentences only. We showed that even a very severe restriction on the grammars did not give learnability. Specifically we required that there be a universal context-free base grammar and that each language in the class of languages be defined by a finite set of transformations on this base grammar. If the base is taken as universal, then it may conceivably be regarded as innate, and hence need not be learned. Still remaining to be learned, however, are the particular transformations that appear in the language to be learned. Linguists are in broad agreement (a possible exception is Bach 1965) that most of these at least must be learned. Thus by assuming a universal base, we make the learner's task as easy as we can, without trivializing it. Still we obtained a negative result; that is, we proved that, given sentences as data, no learner could succeed in learning an arbitrary language of this kind.

It is important to stress that the function of making over-strong assumptions when we are obtaining negative results is not to claim that the over-strong assumptions are correct, but to show that even _with_ these over-strong assumptions the class is unlearnable, and thus without them it is also unlearnable. For example, here we made the too-strong assumption of a universal base and showed non-learnability of certain classes of transformational languages. Thus without a universal base such classes are _a fortiori_ unlearnable.

The next step (Hamburger and Wexler 1973a,b) was to enrich the information presentation scheme in an attempt to achieve a positive result. We thus made the assumption that given the situational context of a sentence the learner had the ability to infer an interpretation of the sentence and from the interpretation to infer its deep structure. Now this is a very strong assumption (Chomsky 1965 notes that it is very strong, though not necessarily wrong), and we have already begun to weaken it further. But the important point is that we finally achieved a positive result. That is, if we assume that the information scheme is a sequence of (b,s) pairs where b is a base phrase-marker and s is the corresponding surface _sentence_ (not the surface phrase-marker, since there is no reason to assume that this information is available to the learner in complete detail) a procedure can be constructed which will learn any finite set of transformations which satisfy the assumed constraints.

By "learn" we mean that the procedure will eventually (at some finite time) select a correct set of transformations and will not change its selection after that time. For a sketch of the proof and a discussion of assumptions, see Hamburger and Wexler (1973a). For the complete proof, see Hamburger and Wexler (1973b).

In the event that the reader thinks that with these strong assumptions the proof of learnability is easy and straightforward he should look at the proof of the learnability theorem in Hamburger and Wexler (1973b). As Peters (1972) notes, the power of transformations that have been assumed is far too large. And, in fact, in addition to assumptions made (explicitly or implicitly) in Chomsky (1965) (for example, all recursion in the base takes place through S, and transformations are cyclic), it was necessary to make six special assumptions in order to derive the result. The first, called the Binary Principle, states that no transformation may analyze more deeply than two S's down. It is quite significant that this principle, assumed for the proof of the learnability theorem, was later proposed independently on purely descriptive grounds by Chomsky (1973), who called it the "Subjacency" Condition. We have since found further descriptive evidence for it. We propose that the reason that the Binary Principle exists is that without it natural language would be unlearnable. The fact that the Binary Principle is necessary both for learning and descriptive reasons lends strong support to its status as a formal linguistic universal. (It should be noted that the descriptive arguments are controversial--see Postal (1972) for arguments that transformations must analyze more deeply).

The other assumptions are all motivated by the fact that, even with the Binary Principle, the number of possible structural analyses is unbounded, so that the learning procedure can be led astray. We therefore made some rather brute-force asssumptions about the analyzability of certain nodes after raising and some other operations. (For the explicit definition of these five assumptions see Hamburger and Wexler 1973b).

Even though these five extra assumptions enabled us to show learnability, there was one rather unsatisfying feature of the result. We showed that the average number of data it took for the learner to get to a correct grammar was less than a certain upper bound, but this bound was very high in comparison to the number of sentences a child hears in the few years it takes him to learn his language.

It was therefore extremely compelling for us to discover later that the five assumptions can be replaced by a single constraint called the Freezing Principle (see Section III, Wexler and Culicover 1973, Culicover and Wexler 1973,1974a) which still allows the learnability theorem to be proved and which has the following properties that (compared to the original five assumptions):

1. a) It is more simply and elegantly stated and in more "linguistic" terms.
   b) The proof of the learnability theorem is much more natural and simple.
2. It provides a better description of English, and in fact is more adequate in explaining judgments of grammaticality in English for a crucial class of phenomena than other constraints considered in linguistics to date.
3. The learning procedure is simplified and is more plausible as a model of the child.
4. All transformations can be learned from data of degree 0, 1 or 2; that is, the learner does not have to consider

sentences which contain sentences which contain sentences which contain sentences, or sentences more complex than these. This result permits a drastically reduced bound on expected learning time. (Result 4 only holds with added assumptions, interesting in themselves.)

These results (especially, from the standpoint of learning, the third and fourth) lend strong credence to the Freezing Principle. As a side-light, it is quite interesting to observe that neither the Freezing Principle nor the five assumptions are stronger than each other in terms of generative capacity. That is, each allows derivations that the other does not allow. Thus the crucial questions in language acquisition and linguistic theory do not depend on the grammatical hierarchy and thus bear out the conjecture of Chomsky (1965, p. 62) who wrote:

> It is important to keep the requirements of explanatory adequacy and feasibility in mind when weak and strong generative capacities of theories are studied as mathematical questions. Thus one can construct hierarchies of grammatical theories in terms of weak and strong generative capacity, but it is important to bear in mind that these hierarchies do not necessarily correspond to what is probably the empirically most significant dimension of increasing power of linguistic theory. This dimension is presumably to be defined in terms of the scattering in value of grammars compatible with fixed data. Along this empirically significant dimension, we should like to accept the least "powerful" theory that is empirically adequate. It might conceivably turn out that this theory is extremely powerful (perhaps even universal, that is, equivalent in generative capacity to the theory of Turing machines) along the dimension of weak generative capacity, and even along the dimension of strong generative capacity. It will not necessarily follow that it is very powerful (and hence to be discounted) in the dimension which is ultimately of real empirical significance.

It is further evidence for the Freezing Principle that it turns out to be quite powerful in just this way. As we have written (Wexler and Culicover 1973, p. 21):

> In fact, we aim to show that a version of the Freezing Principle is a fundamental component of the evaluation metric for syntactic descriptions: by assuming the Principle we are forced into rather particular descriptions. Unlike some of current linguistic theory, a theory with the Freezing Principle is not at all neutral with respect to alternative descriptions in general, but makes unequivocal statements as to which of the alternatives is correct in most cases.

The Freezing Principle is thus unique among linguistic constructs in that it is supported both by learning-theoretic and by descriptive linguistic arguments. Such merging of these two kinds of arguments elevates the discussion to the level of "explanatory adequacy" (Chomsky, 1965).

We propose the Freezing Principle as a formal universal of language and claim as evidence for it that (a) it plays a key role in making language learnable in a reasonable amount of time, while at the same time (b) it also provides in our opinion the best available syntactic description for a wide variety of adult linguistic data. By simultaneously satisfying these two criteria, this theory begins to explain *why* adult language has the structure it does, rather than merely describing that structure.

A major controversy in the study of the theory of language acquisition in recent years has been the question of whether formal structural universals had to be innate in the human child or whether only general cognitive learning abilities were required, as argued, for example, in Putnam (1967). It seems to us that our work provides evidence for the formal universal position since, without assuming the existence of formal universals, we cannot show that language is learnable. We did not come to this conclusion *a priori*; rather the study of learnability theory forced it on us. Also, it should be noted that in order to obtain the proof of the learnability theorems we had to construct an explicit procedure which can be taken as a model of some aspects of the child learning language. This procedure contains a number of aspects which might reasonably be called parts of a "general learning strategy". For example, the procedure forms hypotheses based upon the evidence with which it is presented and changes these hypotheses when evidence counter to them is presented. It is conceivable that this kind of learning is operative in many cognitive domains but that the particular formal structure of the objects upon which hypotheses are formed or which constitute data are different in the various domains.[2] At any rate, to our knowledge, no "general learning strategies" theory exists which has been proved to be successful in learning language, or even a significant part of it.

Recall that we require not only that the learning procedure converge to an appropriate grammar, but that it do so in a "reasonable" way, that is, by being in at least approximate accord with the evidence as to how human children learn language. The fact that the procedure is able to learn from degree 0, 1 and 2 data is in accord with this requirement. But there are, of course, other properties of the procedure which must meet the requirement. The procedure works by always hypothesizing a finite set of transformations (the transformational component). If at any time a (b,s) pair is presented which is not correctly handled by the current component, either a) one of the current transformations is rejected from the component or b) one is added. This is, of course, done in a reasonable, not arbitrary, manner. In this way, a correct set of transformations is eventually obtained. This last statement, of course, requires a long and complex proof.

Note that this procedure has two properties which are quite desirable. First, only one transformation at a time is changed. This seems more in accord with what we observe in the child's developing grammar than would the wholesale rejection of transformational components called for by Gold's (1967) methods. Although the grammar changes gradually (rule-by-rule), the language (i.e., the set of sentences) may exhibit discontinuities over time in that the change of one rule may affect a large number of different kinds of sentences. This is exactly as we would expect from studies of children's grammar.

Secondly, the procedure does not have to store the data with which it has been presented. (Such storage is a feature both of Gold's formal studies and of Klein and Kuppin's simulations.) Rather it determines the new transformational component completely on the basis of the current transformational component plus the current datum. This is desirable because it is quite unlikely that the child explicitly remembers all the sentences he has heard. As Braine (1971) notes:

> The human discovery procedure obviously differs in many respects from the kinds of procedures envisaged by Harris (1951), and others.... A more interesting and particularly noteworthy difference, it seems to me, is that the procedure must be able to accept a corpus utterance by utterance, processing and forgetting each utterance before the next is accepted, i.e., two utterances of the corpus should rarely, if ever, be directly compared with each other. Unlike the linguist, the child cannot survey all his corpus at once. Note that this restriction does not mean that two sentences are never compared with each other; it means, rather, that if two sentences are compared, one of them is self-generated from those rules that have already been acquired.

The fact that transformational components are learnable even given these two rather severe restrictions on the procedure lends further support to the theory.

III. Syntax

A. The Freezing Principle

The Freezing Principle enters into a descriptive account of English as a universal constraint on the operation of transformational rules. There is one crucial difference between the Freezing Principle and other constraints on the application of transformations which have been proposed in the literature; namely, the Freezing Principle emerges from a theoretical analysis of the foundations of linguistic theory (i.e., learnability studies), while other constraints are (more or less abstract) generalizations from the data of syntactic description[3]. The Freezing Principle also turns out, we believe, to be more descriptively adequate than other constraints proposed in the literature.

Before stating the Freezing Principle, we state a few of the assumptions of syntactic theory. The theory (in the by now well-known notation[4]) assumes that context-free phrase-structure rules (the base) generate phrase-markers (trees). (These trees are ordered; this assumption will be modified in the next section.) In the derivation of any sentences, let P_0 be the phrase-marker generated by the base, that is, the deep structure of s[5]. Then a transformation changes P_0 to the phrase-marker P_1, another transformation changes P_1 to P_2, and so on, until P_n, the surface structure of s, is reached. The terminal string of P_n is s. P_1, P_2, P_3 are called <u>derived</u> phrase-markers.

For nodes A and B in a phrase-marker we have the notion <u>A dominates B</u>, where the root (i.e., the highest S-node) dominates all other nodes. We mean <u>strictly</u> dominate, so that A does not dominate A. If A dominates B and there is no node C so that A dominates C and C dominates B, then we say A <u>immediately</u> dominates B. The <u>immediate structure</u> of A is the sub-phrase-marker consisting of A, the nodes $A_1 \ldots A_n$ that A immediately dominates, in order, and the connecting branches. The immediate structure of A is a <u>base immediate structure</u> if A $\rightarrow$ $A_1 \ldots A_n$ is a base rule. Otherwise it is non-base. Before formally stating the Freezing Principle we will illustrate its application to some particularly clear and simple data, for which no explanation other than the Freezing Principle has (so far as we know) ever been proposed. In fact these observations have not, as far as we know, ever been made before.[6]

There is a transformation called COMPLEX NP SHIFT which moves a complex NP (i.e., one which immediately dominates an S) to the end of its verb phrase, as illustrated in (1).

(1a) John gave [the poisoned candy which he received in the mail] to the police.

(1b) John gave to the police [the poisoned candy which he received in the mail].

(The brackets indicate the substring which comprises the complex NP in (1).) Ross (1967:51ff) has shown that the rule applies to a structure with constituents ordered as in (1a) to produce a structure with constituents ordered as in (1b).

A surprising fact is that there can be no movement of the object of the <u>to</u>-phrase (henceforth the "indirect object") just in case COMPLEX NP SHIFT has applied first. Compare (2a) and (2b). ("Ø" indicates the underlying location of the moved constituent, which is underlined.)

(2a) <u>Who</u> did John give [the poisoned candy which he received in the mail] to Ø?

(2b) * <u>Who</u> did John give to Ø [the poisoned candy which he received in the mail]?

Similar facts hold for relative clauses.

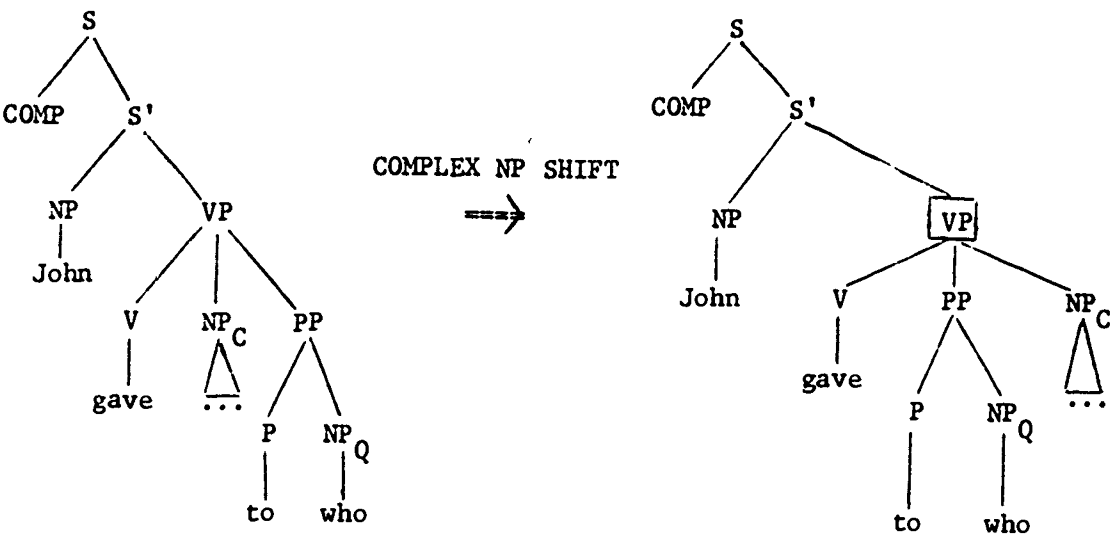
(3a) The police who John gave [the poisoned candy which he received in the mail] to ∅ were astounded by his bad luck.

(3b) * The police who John gave to ∅ [the poisoned candy which he received in the mail] were astounded by his bad luck.

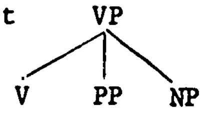
At first sight it might seem as if there might be a number of possible explanations of these facts. In Wexler and Culicover (1973), however, we offer evidence and arguments to rule out possible explanations involving currently available devices of linguistic theory. These include rule ordering, global deviational constraints and perceptual strategies.

The Freezing Principle, however, works perfectly here. The Freezing Principle essentially says that if a structure has been transformed so that it is no longer a base structure (i.e., generable by the phrase-structure rules) then no further transformation may apply to that structure. To see how this applies to these data, note how the transformation of complex NP-SHIFT affects the phrase-marker (4).

(4)



In the derived phrase-marker VP immediately dominates the sequence V PP NP. But



is not a base structure, that is there is no phrase-structure rule in the base component of the form VP → V PP NP. Thus we say that VP is "frozen", which means that no transformation may analyze any node which VP dominates. (To indicate that VP is frozen we place a box around it). In particular no transformation may analyze NP_Q, since it is under VP. Thus WH-FRONTING may not apply, and (2b) and (3b) are ungrammatical.

To give a more formal account of the Freezing Principle we first make the following definition of a <u>frozen</u> node.

<u>Defintion</u>: If the immediate structure of a node in a derived phrase-marker is non-base then that node is <u>frozen</u>.

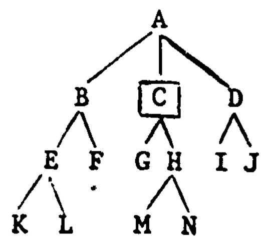
We can then state the

<u>Freezing Principle</u>: if a node X of a phrase-marker is <u>frozen</u>, then no node which X dominates may be analyzed by a transformation.

Note that <u>no</u> node which X dominates may be analyzed, not just the nodes which X immediately dominates. Also note that by this definition, since X does not dominate X, if X is frozen, it may itself be analyzed by a transformation (unless some Y which dominates X is also frozen).

<u>Notation</u>: A box around a node X in a phrase-marker P, i.e. $\boxed{X}$, indicates that X is frozen.

Example:



In this example, C is frozen, i.e., C → G H is not a base rule. Thus the nodes labelled G,H,M, and N may not be analyzed by a transformation.

The Freezing Principle blocks the application of all transformations to parts of a phrase-marker. It does this by freezing certain nodes. If a transformation distorts the structure of a node so that it is no longer a base structure, then no further transformation may apply to elements beneath that node.

This definition captures formally our discussion of the complex NP-SHIFT data. Note in particular that only VP is frozen, so that the subject of the sentence may be questioned or relativized.

(5a) Who gave to the police the poisoned candy which John received in the mail?

(5b) The man who gave to the police the poisoned candy which John received in the mail was his brother.

B. Some empirical justification

We have shown in Wexler and Culicover (1973) and Culicover and Wexler (1973, 1974a) that the Freezing Principle applies to a wide variety of apparently unrelated syntactic domains. These include adverb placement, GAPPING, WH-FRONTING, deletion rules, "seems", DATIVE, and many more. Many of the arguments are rather complex, and require the presentation of considerably more data than this exposition can comfortably accommodate.

We will restrict ourselves here to the development of several of these cases.

The first case illustrates that the Freezing Principle explains phenomena resistant to some of the most successful constraints on the application of transformations proposed to date. It is a well known fact that a constituent of a complement sentence may be questioned and relativized, except when the sentence is a subject complement. Thus,

(6a) It is obvious $_S$[that Sam is going to marry Susan].

(6b) Who is it obvious $_S$[that Sam is going to marry Ø]?

(6c) Susan is the girl who it is obvious $_S$[that Sam is going to marry Ø].

(7a) $_S$[that Sam is going to marry Susan] is obvious.

(7b) *Who is $_S$[that Sam is going to marry Ø] obvious?

(7c) *Susan is the girl who $_S$[that Sam is going to marry Ø] is obvious.

Similar results obtain with the comparative, which Bresnan (1972) argues involves deletion in the <u>than</u>-clause.

(8a) John is dumber than it is conceivable $_S$[that George could ever be Ø].

(8b) *John is dumber than $_S$[that George could ever be Ø] is conceivable.

The usual explanation of these facts is the A-over-A constraint (Chomsky 1964, 1968:43), which requires that an extraction transformation applying to a phrase of type A such as the one illustrated in (6) - (7) must apply to the maximal phrase of that type. Under this analysis the subject complement is immediately dominated by NP, so that the WH-FRONTING rule cannot extract any NP which is contained within the subject complement. This condition does not apply to the extraposed complement sentence, however, and thus (6b) and (6c) are acceptable. It is not clear whether the A-over-A principle could be extended to the deletion case of (8).

Furthermore, and more importantly, Chomsky (1968:46-47) notes that there are a number of cases which require that changes in the A-over-A constraint be made, and cites Ross' evidence (1967) that there are cases which could be handled by the A-over-A constraint only with ad hoc modifications. He concludes that "perhaps this indicates that the approach through the A-over-A principle is incorrect, leaving us for the moment with only a collection of constructions in which extraction is, for some reason, impossible." We believe that there is evidence that the reason is the Freezing Principle.

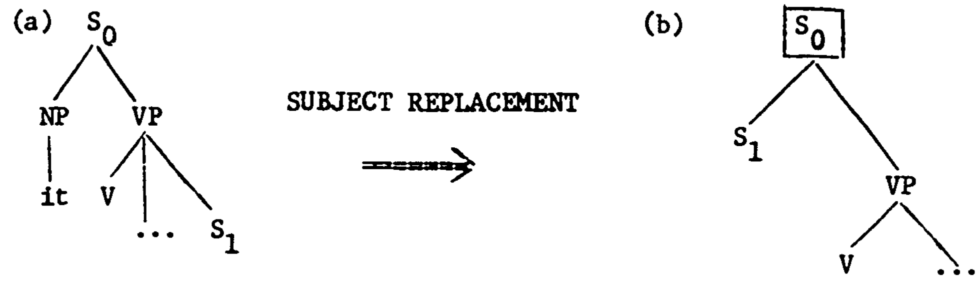
Similarly, Ross (1967:243) proposes the "Sentential Subject Constraint" to account for the failure of WH-FRONTING and other movement rules to apply to a constituent within a sentential subject:

> SSC: "No element dominated by an S may be moved out of that S if that node S is dominated by an NP which itself is immediately dominated by S."

As we will show, this constraint is not sufficiently general to account for the entire range of data subsumed by the Freezing Principle.

To see how the Freezing Principle predicts these data, we make use of Emonds' (1970) analysis, in which (9b) is derived from (9a) by means of a rule of SUBJECT REPLACEMENT.[7]

(9) (a) (b)



Since S_0 now dominates S_1 VP and S $\rightarrow$ S VP is not a base rule, S_0 is frozen. Thus no element of S_1 may be moved and thus (7b) and (7c) are ungrammatical.

So far, looking at just these data, on the purely descriptive level there is no reason to prefer either the Sentential Subject Constraint or the Freezing Principle. But now notice

(10a) It is obvious $_S$[that John is going to need some help].

(10b) *Is $_S$[that John is going to need some help] obvious?

To derive (10b), first apply SUBJECT-REPLACEMENT, freezing S, and then INVERSION. The Freezing Principle predicts that (10b) is ungrammatical, since the structure to which INVERSION applies in (10b) is frozen. The Sentential Subject Constraint, however, does not make this prediction.

Ross (1967:57) accounts for (10b) with the following output condition: "Grammatical sentences containing an internal NP which exhaustively dominates S are unacceptable". Thus Ross' two constraints, which we have called generalizations from the data (as opposed to theoretical propositions), are accounted for nicely by the Freezing Principle. We would say that these data in themselves would force us to prefer the Freezing Principle. But the situation is even more clear-cut, for there are related data which none of Ross' principles account for, but which are predicted by the Freezing

Principle. These are[8]

(11a) How obvious is it $_S$[that John is going to need some help]?

(11b) *How obvious is $_S$[that John is going to need some help]?

(11c) How necessary is it $_S$[for John to leave]?

(11d) *How necessary is $_S$[for John to leave]?

Once again, SUBJECT-REPLACEMENT freezes the entire sentence, so that the adjective phrase may not be moved, according to the Freezing Principle. Since nothing has been moved out of the subject, the Sentential Subject Constraint does not apply, and since the sentential complements in (11b) and (11d) are not internal, Ross' output condition does not apply. Thus not only does the Freezing Principle predict all the data that Ross' two constraints predict, but it predicts data that Ross' constraints cannot predict.

Another case involves the transformation which derives (12b) from (12a) (cf. Chomsky 1970 for discussion).

(12a) John's pictures

(12b) the pictures of John's

Alongside (12b) we observe the construction exemplified by (12c).

(12c) the pictures of John

While (12c) corresponds to a possible base structure, and may in fact be a base generated structure, (12b) is derived by a transformation which clearly causes freezing. Hence the Freezing Principle predicts that it should be possible to question the object of the preposition <u>of</u> in a construction like (12c), but not in a construction like (12b). This prediction is correct, as the examples below show.

(13a) Mary saw the pictures of who's ⟹ *Whose did Mary see the pictures of?

(13b) Mary saw the pictures of who ⟹ Who did Mary see the pictures of?

As a last case consider the dative construction in English. As we show in Culicover and Wexler (1973), after the DATIVE transformation has applied, deriving (14b) from (14a), no other transformation, such as WH-FRONTING, for example, can apply to the indirect object. However, these transformations can apply to the indirect object if DATIVE has not applied.[9]

(14a) John gave a book to Bill.

(14b) John gave Bill a book.

(14c) What did John give to Bill?

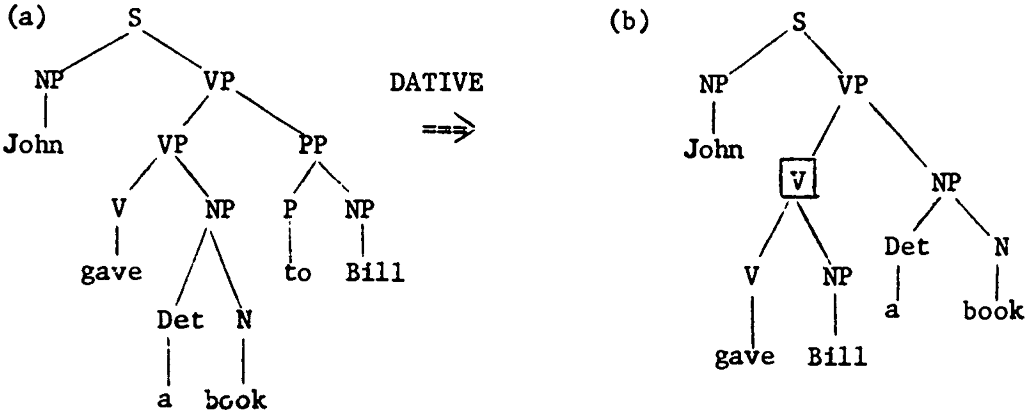
(14d) Who did John give a book to?

(14e) What did John give Bill?

(14f) *Who did John give a book?

These judgments are generally accepted in the literature, but have resisted explanation. Langendoen (1973), in fact, noting that the data cannot be explained by rule ordering, suggests two special ad hoc conditions either of which could explain the data and then writes, "Either way, the solution seems inelegant and ad hoc, and one is led to question the grammaticality judgments which motivated them in the first place". Of course, if it happens too often that the intractability of an analysis requires judgments to be questioned, then the entire empirical basis of linguistics is gone. Thus it is intriguing that the Freezing Principle provides a natural solution to this problem with no change at all in the data. Assume that (14b) is derived from (14a) as in (15).

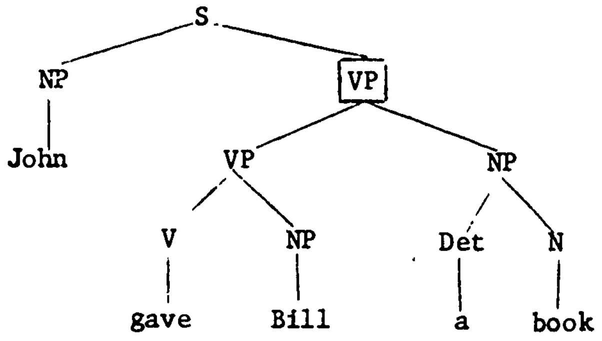
(15)



Since there is no base rule of the form V → V NP, the upper V node in (15b) is frozen, and thus WH-FRONTING cannot move the NP dominated by V and thus (14f) is ungrammatical by the Freezing Principle. But since the NP <u>a book</u> is not frozen, (14e) is grammatical.[10]

But apparently there is some "dialect" variation in these judgments. Hankamer (1973) finds sentences like (14e) ungrammatical, although he otherwise accepts these judgments. That is, after DATIVE, Hankamer cannot question either the direct or indirect object.[11] Note that exactly this pattern of grammaticality judgments will be predicted if the upper V in (15b) is changed to a VP, as in (16).

(16)



Since there exists no rule in the base of the form VP → VP NP, the upper VP in (16) will be frozen and thus, by the Freezing Principle, neither the indirect object nor the direct object may be questioned, thus predicting this second pattern of judgments.

But how is a learner to choose between (15b) and (16)? If (16) were indeed correct (i.e., was being used by the speakers from whom he was learning the language), and if the learner had decided on an analysis of the form (15b), then, if there is no correction of ungrammatical utterances, the learner will never have reason to change his analysis.[12]

In short, the data, together with the language learning procedure, might not determine whether (15b) or (16) is correct. There _might_ be a general constraint which determines that when Chomsky-adjuction takes place, inserting a node between X and Y (with X dominating Y), then the new node is always called Y, as in (15b). If the judgments listed in (14) are correct, then this constraint seems reasonable. If the mentioned "dialect" variations actually exist, then the constraint possibly is _not_ correct, and the learner may be free to choose either X or Y as the name for the new node.[13]

Note the power of the Freezing Principle here. Although it allows both sets of grammaticality judgments, it does _not_ allow a third set, in which one could move the _indirect_ object after DATIVE, but not the _direct_ object, that is, one in which (14e) is ungrammatical and (14f) grammatical. This is because there is no way of stating the transformation so that a node dominating the direct object is frozen, but not a node dominating the indirect object. So there is a formal, precise prediction that this third dialect cannot exist, and so far as we know this pattern does _not_ exist for any native speaker.

C. Rule-ordering

We have also found that there is considerable reason to believe that transformations need not be extrinsically ordered if one assumes that the Freezing Principle is a constraint which is operative in natural language. It should be evident that the goal of dispensing completely with extrinsic ordering would be a desirable one to attain, provided that it is consistent with the empirical evidence.

To consider a particular example, let us return to sentences involving extraposed and non-extraposed sentential complements. It turns out that it is impossible to delete a that-complementizer if the complement appears in subject position.

(17a) It is obvious $\left\{\begin{matrix}\text{that}\\ \emptyset\end{matrix}\right\}$ Mary was here yesterday.

(17b) $\left\{\begin{matrix}\text{That}\\ *\emptyset\end{matrix}\right\}$ Mary was here yesterday is obvious.

In order to block the deletion of that in the sentential complement one might order the rule of THAT-DELETION after SUBJECT REPLACEMENT. Alternatively, if one wished to argue that the rule relating (17a) and (17b) was EXTRAPOSITION, where the underlying constituent order is that of (17b), then one would order THAT-DELETION after EXTRAPOSITION. Presumably the structural description of THAT-DELETION would be stated in either case so that it could not apply when the complement was in subject position.

However, observe that if the Freezing Principle is assumed, then the transformations need not be ordered in the SUBJECT REPLACEMENT analysis. If SUBJECT REPLACEMENT applies first, then THAT-DELETION is blocked by the frozen structure. If THAT-DELETION applies first, then either the resulting structure is frozen, or else the resulting structure fails to meet the structural description of SUBJECT REPLACEMENT, depending on independent requirements of the analysis. On the other hand, it can be seen that such an explanation is impossible in terms of the EXTRAPOSITION analysis. Hence the Freezing Principle, for this body of data at least, permits us to do without extrinsic rule ordering, and in doing so, leads to an unambiguous interpretation of the data.

Another example involves the interaction between DATIVE and COMPLEX NP SHIFT (noted by Ross 1967:53ff). In its most general statement COMPLEX NP SHIFT moves an NP to the end of the VP which dominates it. However, this rule cannot apply after DATIVE has applied.

(18a) I gave a book about spiders to the man in the park.

(18b) I gave to the man in the park a book about spiders.

(19a) I gave the man in the park a book about spiders.

(19b) *I gave a book about spiders the man in the park.

One way to rule out (19b) would be to order COMPLEX NP SHIFT before DATIVE. Application of COMPLEX NP SHIFT would then destroy the environment for the latter application of DATIVE. However, since both DATIVE and COMPLEX NP SHIFT cause freezing at the VP which dominates the two objects, the application of either transformation will block the later application of the other if the Freezing Principle is assumed. Hence it will be unnecessary to state an extrinsic ordering of the two rules.

Finally, consider Emonds' (1970) list of "root" transformations in English.

Directional adverb preposing EX: Away John ran.

Negated constituent preposing EX: Never will anyone do that.

Direct quote preposing EX: "John is a fink," Bill said.

Non-factive complement preposing EX: John is a fink, Bill assumes.

Topicalization EX: Beans I hate.

VP Preposing EX: John said I would like her, and like her I do.

Left dislocation EX: John, he really plays the guitar well.

Comparative substitution EX: Harder to fix would be the faucet.

Participle preposing EX: Standing in the doorway was a witch.

PP substitution EX: In the doorway stood a witch.

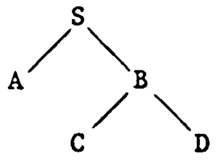
As Emonds points out, only one of these transformations may apply in any derivation. This condition follows as a consequence of the Freezing Principle, if one makes the reasonable assumption that each of these transformations causes freezing at the S-node to which it applies. Observe that in this case it is simply impossible to find an extrinsic ordering of all of the rules mentioned which will account for the fact that only one of them may apply at a given S. Hence not only does the Freezing Principle permit us to do away with a number of cases where extrinsic ordering would otherwise be required, but it accounts for a situation in which rule ordering alone is not adequate to account for the data.

IV. Semantics

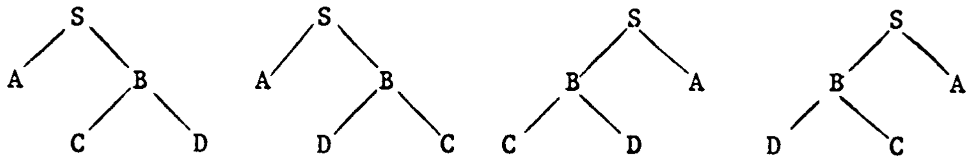
A. The Invariance Principle

The role of semantics in the linguistic system must be analyzed carefully, because, in addition to the necessity of providing an adequate descriptive semantics, we must understand how meaning helps to provide structural information to the language learner. As a first step we assumed the Universal Base Hypothesis, which says that there is one syntactic base for all languages. But, of course, since languages have different syntactic deep structures (e.g., all languages are not SVO), this assumption must be modified. In Wexler and Culicover (1974) we modify this assumption along lines which have been previously suggested. We assume that there is a "semantic" structure, which is hierarchical but not ordered from left to right, and we assume that this structure is related to the syntactic deep structure in a very constrained way: the hierarchical relations in the semantic representation are retained in the syntactic deep structure, although any left-to-right order, given this constraint, is acceptable. This constraint is called the Invariance Principle, because the grammatical relations are assumed to be invariant from semantic to syntactic structure. As an artificial example, suppose the semantic representation has the unordered structure in (20a). Then any four of the ordered deep structures in (20b) are possible, by the Invariance Principle.

(20a)



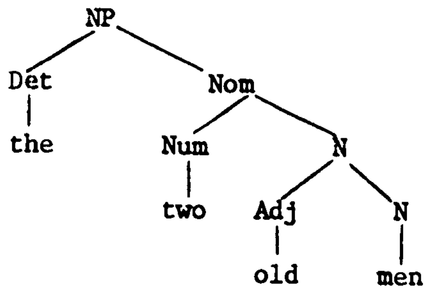
(20b)



We also assume that the "semantic grammar" is universal, but that natural languages differ in which <u>ordered</u> deep structure they have. All of these deep structures are related, however, by the Invariance Principle. This is a very strong assumption, and has the virtue that it allows the deep structures of a language to be learned by a fairly simple learning procedure. But although this is such a strong assumption, there is considerable evidence for it. This evidence is presented in Culicover and Wexler (1974b), where data from 218 languages is considered.

The evidence takes the form of predictions about universals of word order. For example, suppose the universal unordered semantic representation for the Noun Phrase is

(21)



There is evidence that the ordered form of this structure as shown in (21) is correct for English. Then, the Invariance Principle predicts that only eight deep structure orders are possible for the four categories Det, Num, Adj, N ; namely those obtained by permuting each branch of the structure. Thus the possible orders are Det Num Adj N , Num Adj N Det, Det Adj N Num, Adj N Num Det, Det Num N Adj, Num N Adj Det, Det N Adj Num, and N Adj Num Det.

Without constraints, of course, there are 4! = 24 orders of the four categories available. Therefore the prediction that only 8 are possible is a strong prediction. In Culicover and Wexler (1974) we find that, of all the languages for which adequate data is available, there is only one exception to this prediction, that is, only one order of these constituents which is not in the eight predicted ones.[14] All the other languages have an order which is one of the eight predicted ones.

Thus note that the Invariance Principle together with the assumed universal semantic representation makes very strong predictions which can be confirmed. In Culicover and Wexler (1974) we also confirm the predictions for a number of other structures.

All of this evidence is used to support both the Invariance Principle and the assumed universal semantic representation, which is hierarchically structured (i.e., it is like, though in detail different from, an unordered version of traditional context-free deep structures for English). There have been a number of other proposals in the literature for the form of the "semantic base", most of them being more similar to a version of the predicate calculus notation (e.g., Lakoff 1970a) or a case system (e.g., Fillmore 1968). It is important to note that none of these proposals can satisfy the Invariance Principle, and that, so far as we can see, they cannot (without numerous ad hoc assumptions) make the strong predictions about universals of word order in Culicover and Wexler (1974). Thus we have evidence that the traditional structured deep structure is correct.

To take another example, note that the Invariance Principle, together with the assumption that the semantic grammar rewrites S as NP-VP, where the VP is expandable as either V or V-NP, predicts that if the subject of a sentence precedes the V in a transitive sentence then the subject must precede the V in an intransitive sentence. Once again our data completely confirm this prediction, and there is no non-ad hoc way for the predicate calculus formulations to predict these phenomena.

The kind of counter-example to these claims that might occur to the linguist is a so-called "subjectless" language, in which, it has been argued, there is no deep subject-predicate structure. But the existence of these languages has, it seems to us, not been at all demonstrated. In Culicover and Wexler (1974) we analyze Kapampangan, a language which it is claimed is subjectless, and show that an analysis which assumes an underlying subject-predicate division accounts more readily for a number of interesting grammatical phenomena in the language than does a "subjectless" analysis, for example, Mirikitani's (1972).

Thus there is evidence that the Invariance Principle is correct. It is also true that, given the constraints imposed by the Invariance Principle, the (ordered) deep structure rules are quite easily learnable (Wexler and Culicover 1974), which, of course, is a goal of the analysis.

B. Semantic adequacy

There is one other very important kind of analysis which must be made to justify the system, and this is to provide evidence that the semantic structures which the Freezing Principle and Invariance Principle force us to assume are in fact descriptively adequate.

Application of the Freezing Principle places very strong restrictions on what the deep structure configuration of a sentence may be given the appropriate kinds of information about what the transformational mapping between the deep structure and the surface structure must account for. Hence the assumption that hierarchical arrangements in deep structures and semantic structures are preserved by the mapping between them (the Invariance Principle) together with the predictions about deep structures made by the Freezing Principle serve to make quite explicit predictions about the nature of semantic structures. It is necessary to show that the theory sketched out above is in fact explanatorily adequate, in that it leads directly to a descriptively adequate semantic account. In other words, we wish to show that the semantic structures which we arrive at are the correct ones in terms of the interpretations assigned to them by the semantic component. Our results in this area are somewhat tentative, so we must restrict our remarks here to a discussion of the direction in which such an investigation might lead.

1. The extensionality of the subject.

Let us say, following a traditional terminology of modern logic, that the extension of an expression is its reference, where the extension of a sentence is either truth or falsity depending on whether the sentence is true or false. Let us also say that the intension of an expression is a function defined in the semantic component which assigns to each expression its extension if it has one.

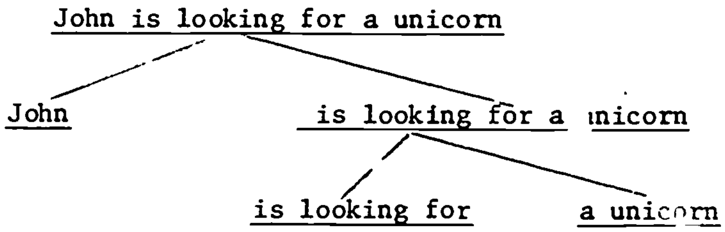
An opaque context is one in which a sub-expression of an expression need not have an extension in order for the entire expression to have an extension. One such example is (22).

(22) John is looking for a unicorn.

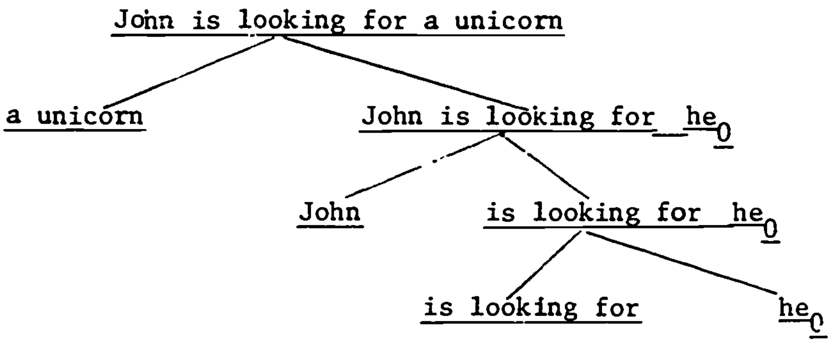
(22) may be true or false even if there is no such thing as a unicorn. There is a second reading, of course, in which a unicorn must exist.

Montague (1973) represents this ambiguity of an expression such as (22) in the following way. In the syntactic derivation of the sentence the direct object of the verb is looking for may be either the intension of a unicorn, which we may represent here informally as a unicorn', or the object of the verb may be a variable expression he_i , whose intension may be represented informally as he_i' . In the latter case the surface structure of the sentence is derived by replacing the expression he_i by the expression a unicorn. Thus the sentence is syntactically as well as semantically ambiguous, by virtue of the fact that it has two derivations. (In fact it has several more which do not lead to further semantic ambiguity.) Associated with the two derivations are different rules of semantic interpretation, so that the semantic structure associated with the sentence is different depending on the syntactic rules which participate in the derivation. The two syntactic derivations are given informally as (23a) and (23b) respectively, while the corresponding semantic representations are given informally as (24a) and (24b) respectively.

(23a)



(23b)



(24a) John' (is looking for' (a unicorn'))

(24b) $\exists$x (unicorn'(x) & (John' (is looking for' (x)))) .

In essence, the device of introducing a noun phrase in the syntactic derivation outside of the context of the verb is looking for permits Montague to maintain in principle the semantic ambiguity by keeping the translation into the semantic representation of a unicorn within the context of the verb in the first case, and outside of it in the second case.

In fact, however, most verbs do not possess this property of permitting their direct object to be intensional. In a case where there is a non-

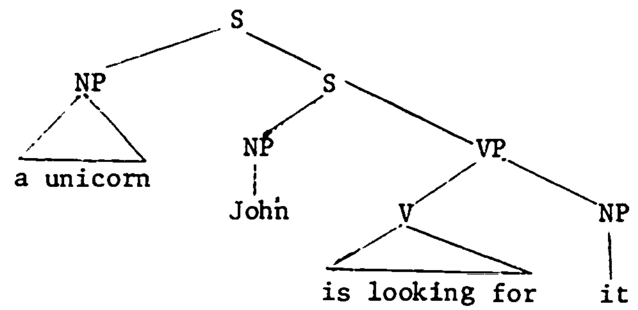
ERIC

intensional verb, such as hit, or saw, Montague applies a meaning postulate which "maps" the semantic representation of the form (24a) into the semantic representation of the form (24b). This rule is inapplicable just in case the verb is one like is looking for.

It is clear that this is not a logically necessary analysis of the data. It is certainly possible to imagine an alternative formulation, in which there is only one syntactic derivation of the simple sentence, and in which there is a semantic rule which obligatorily derives semantic representations such as (24b) from those like (24a), except when the verb is of the type is looking for, in which case the rule applies optionally.

Application of the Invariance Principle leads us to favor the second alternative. There is no syntactic evidence to suggest that a possible deep structural analysis of (22) is that given in (25) below.

(25)



If this is the correct analysis of the syntactic data, as we believe it is, the Invariance Principle will not in itself lead us to two semantic representations for a sentence such as (22). It is worth asking, therefore, whether there is any evidence that the second alternative formulation of the ambiguity of (22) is in principle the correct one.

It is important to point out that in Montague's analysis the first level of semantic representation is one in which all noun phrases are translated into their corresponding intensional expressions. As Montague correctly points out, there are no verbs such that the subjects of such verbs may not be further translated into extensional expressions. We have already seen that there are verbs whose objects may not be so translated, however. Consequently Montague is forced to state two rules, one of which extensionalizes the direct objects of non-intensional verbs (such as hit, see, etc.) and the other of which extensionalizes the subjects of all verbs. This formulation, as can be seen, is ad hoc in that it provides no explanation for why it should be that subjects are always extensional but objects are not.

Furthermore, Montague uses a device of reducing the primary semantic representations to representations of the form of the predicate calculus with a function (argument, argument,...) structure. Hence he finds it necessary to then state rules of extensionality for expressions with one argument, another for expressions with two arguments, and he would have presumably had to state one rule for expressions with three arguments, another rule for expressions with four arguments, and so on, had he extended his analysis to more complex types of expressions. The crucial infelicity of such an approach is that it fails to explain why it should be that the subject is always extensional regardless of the form of the expression. While it is certainly possible to express this fact within Montague's framework, it does not follow as a necessary consequence of the analysis.

A notable characteristic of Montague's approach to the translation of expressions with syntactic structure into semantic representations is that the basic structure of the expression is preserved in the primary semantic representation. The mapping in his framework therefore conforms to the Invariance Principle. Furthermore, the syntactic structure is one which displays the subject/predicate split, and this split is therefore preserved in the primary semantic representation. It is only at a secondary level that Montague reduces the semantic representation to an expression which closely conforms to the type of representation traditionally employed in the predicate calculus. It seems to us, however, that it is not logically necessary to perform this reduction of structure in a semantic component whose goal is to provide a precise characterization of the notion of truth. That such a reduction may even be wrong is shown by the fact that it destroys the structure which might otherwise serve to contribute to a precise and general characterization of opaque contexts.

A first approximation to a solution of the problem would be the following: First, formulate an hypothesis about what constitutes an opaque context in terms of the structure in which the element which creates this context participates. Second, state a semantic rule which is sensitive to the presence of an opaque context which will account for the ambiguity of an expression which contains one at the semantic level. Third, show that this definition is extendible to a wide variety of expressions, and that it can be used as a diagnostic for semantic structure. Fourth, show that the semantic structures arrived at in this way are appropriately related by the Invariance Principle to the syntactic structures arrived at by independent application of the Freezing Principle to the transformational component.

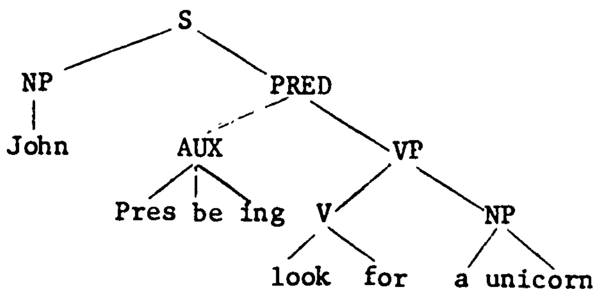
2. Definition of an opaque context.

Let us return to example (22).

(22) John is looking for a unicorn.

We assume that the syntactic structure of (22), and hence its semantic structure exclusive of constituent order, is as in (26).

(26)



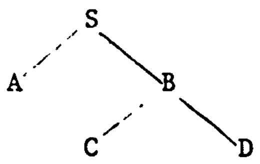
Let us refer to expressions such as *look for* as *opacity causing elements*, or OCE's. What properties of the structure will permit us to distinguish between the subexpressions which are within the context of an OCE, and those which are not?

The property which we would like to suggest is that of *in construction with*. Klima (1964) defines *in construction with* as follows (p. 297), rephrased slightly:

> *Definition*: A constituent *A* is in construction with a constituent *B* if *A* is dominated by the first branching node which dominates B, and B does not dominate A.

For the sake of clarity we will say that if A is in construction with B, then B *governs* A. To illustrate, in (27) below *A* governs *B*, *C* and *D* and B is governed only by *A*. *C* and *D* govern one another.

(27)



Returning to (26), now, we find that *governs* serves to distinguish between the NP *John* and the NP *a unicorn* in terms of their structural relationship with the OCE *look for*. The former, which is outside of the opaque context, is not governed by the V *look for*, while the latter, which is inside of the opaque context, is governed by *look for*. On the basis of this observation we may formulate the following definition of what constitutes an opaque context.

Definition: an expression *E* is in an opaque context with respect to an opacity causing element *O* if *O* governs *E*.

It turns out that if a constituent *A* is governed by a constituent *B* then every constituent which *A* dominates is also governed by *B*. If the definition of an opaque context given above is correct, then, we would expect that any constituent of a constituent in an opaque context is also within an opaque context. This prediction is verified by examples such as the following:

(28a) John is looking for a unicorn with two horns.

(28b) John is looking for a unicorn with two horns that have blue and green polka dots on them.

(28c) John is looking for a unicorn that can ride a bicycle.

As can be seen, not only is it the case that the unicorns defined in the examples in (28) need not exist in order for the expressions to be true, but neither do horns, horns with blue and green polka dots on them, blue and green polka dots, or bicycles have to exist in order for these expressions to be true. Since it is well-known that prepositional phrases and relative clauses such as those found in the examples in (28) are constituents of the NP's which they modify, these observations serve to verify to some extent the prediction made by the definition of opaque context which we have formulated above.

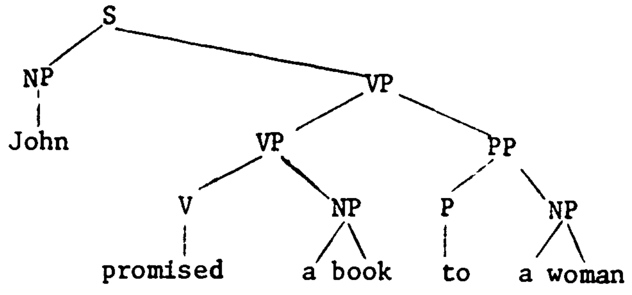
One further example will show how syntactic and semantic evidence converge to require the same analysis. In Section III we showed how the Freezing Principle explains many previously anomalous facts about the DATIVE transformation. In order to explain these facts, a structure had to be taken as basic which included the prepositional phrase, and the other structure had to be derived from that. Thus (29b) must be derived from (29a), and not vice-versa, in order for the Freezing Principle to correctly predict the phenomena.

(29a) John promised . book to a woman.

(29b) John promised a woman a book.

The structure underlying (29a) is (30).

(30)



But the semantic evidence supports this analysis also. Since promise is an OCE, we predict that a referent need not exist for an NP which it governs. Thus, assuming structure (30), the referent, a book need not exist. On the other hand, since promise does not govern a woman in (30), the referent of a woman must exist. These predictions are correct. In other words, (29) is two-ways ambiguous, the ambiguity depending on whether or not a certain book had been promised.[15]

But if (29b) were taken as basic, then these predictions would not be made. Presumably both NP's (a book, a woman) would be in construction with promise (in a "double object" construction) and the prediction would be that (29a,b) were four ways ambiguous, which is not the case.

Thus syntactic and semantic evidence, of very different kinds, converge on one analysis and lend credence to the joint assumptions.

V. Language Acquisition Data

As we noted at the beginning of this article, the empirical basis for the justification of our theory lies, for the moment, in linguistic data, rather than in the data of child speech. Our approach is different from the one usually adopted in the study of language acquisition, which is to study the language of children who have not yet acquired adult competence. The two approaches should be seen as complementary. Ultimately, of course, we hope that a more direct empirical justification could be found for our theory in data concerning child language. At the moment, however, we must be content with a situation not unheard of in science, in which indirect justification is all that is available.

Let us, however, consider ways in which our theory might make contact with empirical data concerning child language. Logically, there seem to be two ways in which this can happen. First, it might be possible to test empirically various of the assumptions of the theory. Secondly, the theory might make predictions about the course of language acquisition which could be tested.

With respect to testing the assumptions of the theory, some of this has already been accomplished. For example, we assume that the child is not corrected for ungrammatical sentences, and, as we mentioned earlier, this seems to be an empirical result (Brown and Hanlon 1970). Other assumptions have not been tested so directly. For example, we assume that the child hears sentences in situations which are clear enough for him to be able to interpret the meaning without understanding the sentence. Although so far as we know this assumption has not been directly tested, it is certainly consistent with empirical results (e.g., Ervin-Tripp 1971, Snow 1972) which show that children are spoken to simply (the assumption being that, all other things being equal, the meaning of simple sentences is easier to determine from the situation). The fact that our theory (with the Freezing Principle) allows transformations to be learned from relatively simple sentences is also consistent with the simplicity of input to the child.

The second way in which the data of language acquisition might be relevant to our studies is that our theory might make testable predictions about the course of language acquisition. For example, the combination of our assumptions about language and the learning procedure might make predictions about which transformations developed first. This is a very subtle question however. The problem is that there are so many ways of changing parameters (e.g., the order of input, the weighing of hypotheses, the pragmatic importance of various transformations) that there may be no unique or small collection of possible orders of development predicted by the theory. And, with respect to transformations it may be that the order of development differs from child to child. Another important difficulty with respect to making these kinds of predictions is that performance considerations (e.g., problems of short-term memory and the actual sentence generation mechanism used by the child, what Watt (1970) calls the development of the "abstract performative grammar") might have large effects on children's utterances, as might aspects of cognitive development.

However, more subtle kinds of predictions might be made. For example, it is a well-known observation (Bellugi-Klima 1968) that children sometimes learn a transformation and use it correctly when no other transformation is involved in the sentence, but when another transformation is needed, both cannot be used together, and one is not applied. An example is INVERSION and WH-FRONTING. Thus a child might say "Is your name Bill?" thus demonstrating INVERSION, but also say "what your name is?" thus not using INVERSION when WH-FRONTING is necessary. The suggested explanation of these observations is that there is a performance limitation on the child; namely he can use only one transformation at a time. However, it may be that the Freezing Principle can play a role in the explanation of these phenomena. The child's grammar may be such that one of these transformations causes freezing and blocks the other one. Thus both transformations cannot apply together. This, of course, is not true of the adult grammar, but the child must learn the appropriate statement of each transformation. There is considerable room for error, even if he assigns the surface string correctly in some cases.

We wish to emphasize that the above suggestion is only speculative, and that very much analysis of the child's grammar would have to be undertaken to make it a reasonable hypothesis. In particular, one would have to find ways to tease it apart from the performance "one transformation at a time" hypothesis. It is only mentioned to indicate the possibility of the interaction of the syntactic portions of the theory with the data of language acquisition.

Another example of how the theory can be used to make predictions about the data of child language acquisition is provided by the problem of word order in early child language. There is some difficulty in finding relevant data because it is possible that the development of the base grammar (i.e., the order of the elements that define grammatical relations) is very fast, at least for the major categories. Thus one would have to observe the child quite early in his linguistic development, right from the start of the two-word stage, in order to capture data relevant to the predictions. In fact, it is entirely consistent with the theory for the child to make no production errors at all with respect to the order of the deep structure constitutents, since the procedure which learns this order is quite simple and straight-forward. In contrast with the procedure which learns transformations, this procedure converges very quickly, and it is quite conceivable that convergence has taken place before the child starts to actually produce two-word utterances. So we require very subtle ways of finding those few errors that do occur.

The base grammar that children develop will, of course, depend on the base of the language that they are learning. But since many of the sentences the child hears involve transformations, there is no reason to suppose that necessarily all children learning a given language will pass through exactly the same stages. In particular, a given learner might, at some stage, posit an incorrect base grammar. However, if the learner is obeying the constraints that we have proposed, namely the Invariance Principle, then we can formally predict that there are certain patterns that should never be observed. In particular, all the universals which we have predicted for the base grammar of any language (Culicover and Wexler 1974b) should hold for a given stage of one language learner.

For example, we predict that no language would have (as deep structures) VSO order for transitive sentences and SV order for intransitives. Thus we predict that, since he is forming his grammars under the constraint of the Invariance Principle, no child will simultaneously have these orders for deep structures. (It is possible that at one time a child has SVO and SV and at a later time VSO and VS).

One can test this prediction by looking at reports of children's utterances. Kernan (1969, 1970) has found that, in the two-word stage, a Samoan child has VS and VO orders (Kernan actually uses a case grammar description, but for these purposes this can be modified). Thus in three word sentences we would predict either VSO or VOS. In fact, the one three word utterance the child makes is VOS. Thus the prediction is confirmed.[16]

Another more interesting case is Gia I in Bloom (1970). Gia at this (early two-word) stage made (according to Bloom's criteria) 3 utterances with a subject and a verb. They were "girl write" (in response to the question "What's the little girl doing?") and two instances of "Mommy back". The fact that in these intransitive verb cases the subject comes first

ERIC

(i.e., they are SV) predicts, according to the Invariance Principle, that the subject will come before the verb in transitive sentences. The only other utterances with verbs that Gia makes at this stage are 3 utterances of the form OV (object verb), for example "balloon throw". Thus we know that O comes before V. Now, in N plus N constructions (presumably the V has been left out), Gia always puts the S before the O, that is SO. Thus since O comes before V and S comes before O, we know that S comes before V in transitive sentences, which is the prediction made from the Invariance Principle given the data that SV was the order in intransitives. Thus Gia's order is consistent with the predictions made by the Invariance Principle.[17]

VI. Summary

In Section II we considered the nature of the constraints which notions of learnability impose on the class of possible human languages, and on the nature of the human language learning mechanism. Section III dealt with some linguistic evidence to support the universals of syntax which emerge from the learnability studies, namely the Freezing Principle and the Binary Principle. In Section IV we discussed some theoretical and empirical work in semantics.

The significance of semantic considerations rests on two crucial aspects of the theory: first, our theory of language acquisition utilizes semantics as a crucial component of information for the language learner; second, any theory of syntax must provide structures which are consistent with a theory of semantic interpretation.

In Section IV it was also shown how the Universal Base Hypothesis may be replaced by a less restrictive hypothesis called the Invariance

Principle, which relates syntactic and semantic structures. Given the Invariance Principle the base component of the grammar is learned by a simple learning procedure. In addition, we discussed briefly the notion that the Invariance Principle and the Freezing Principle taken together make a number of very strong, and we believe correct, predictions concerning universals of constituent order in human language.

In Section V we considered how various kinds of techniques used in developmental psycholinguistics may be used to find empirical evidence relevant to the learning theory. We also discussed several examples which may prove to be fruitful upon further close examination.

Thus, the work reported on here represents research towards the following objectives:

1. the specification of a theory of grammar of human language, insofar as it is characterizable in terms of formal linguistic structural universals;
2. the precise specification of a psychologically plausible theory of the language learner;
3. the formal demonstration that the device specified in 2 above learns the grammar of any possible language specified by 1 above;
4. the demonstration that the linguistic representations and constraints arrived at in 1 above and the procedure specified in 2 above, are empirically correct.

ERIC

Given the fundamental correctness of the assumptions and arguments summarized in this paper we would hope that the successful completion of the work will simultaneously yield a theory of grammar, a theory of language acquisition, a proof of their mutual compatibility, and further empirical support for the entire theoretical apparatus and the interactions between its components.

ERIC

FOOTNOTES

*This work was partially supported by the Office of Naval Research and the Advanced Research Projects Agency, ONR Contract N 00014-69-A-0200-6006.

[1]The published work consists of Hamburger and Wexler (1973a), and Wexler and Hamburger (1973). Hamburger and Wexler (1973b) will appear in print shortly. The unpublished work consists of Culicover and Wexler (1973a,b; 1974), and Wexler and Culicover (1973,1974). The book in preparation is Wexler, Culicover and Hamburger (in preparation).

[2]It is even conceivable, but highly speculative, that some formal universals of language, for example, the Freezing Principle, are special cases of a principle that applies in all cognitive domains, and that the function of the principle in all these domains is the same--namely, it makes the domains learnable. We know of no evidence for or against this conjecture, which nevertheless suggests directions for research in other fields. It is possible however, that the nature of linguistic structure may be sufficiently different from that of other cognitive domains to make the search for something like the Freezing Principle a difficult one.

[3]An exception to this statement is Chomsky (1955, Ch. VIII especially), in which the original constraints on transformations are proposed on the basis of logical analysis (although not on the basis of formal learnability considerations).

[4]As presented, for example, in Chomsky (1970).

[5]We are ignoring here the stages in the derivation prior to the completion of lexical insertion. P_0 is assumed to be the base phrase marker with all lexical items inserted in this discussion.

[6]Much of the following account is taken directly from Wexler and Culicover (1973).

[7]Higgins (1973) argues against Emonds' analysis, but we feel that there is considerable value in trying to maintain Emonds' analysis in light of the applicability of the Freezing Principle as shown in this discussion. Many of the difficulties that Higgins points out can be dealt with within the framework of the SUBJECT REPLACEMENT analysis. Also, many of his arguments do not apply to the Freezing Principle analysis.

[8]Higgins (1973) notes this data.

[9]One serious problem with this analysis which we have discovered thus far is that the PASSIVE transformation may apply to the output of the DATIVE transformation, giving sentences like

(i) Mary was given a book by John.

In Culicover and Wexler (1973a) we suggest an explanation for this fact; however, we do not find the explanation particularly satisfactory, and the problem remains.

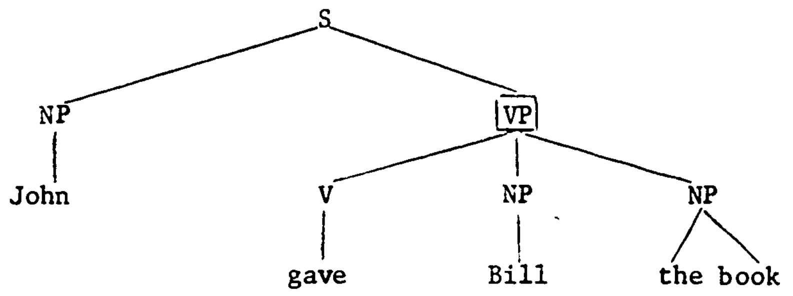
[10]We believe, in fact, that (15a) is the correct structure. The structure used in (4) is given for expository purposes only. In either case, none of the arguments are affected.

[11]He writes, "Ben Shapiro (personal communication) has found that some people, like me, reject any sentence involving chopping either the direct object or the indirect object; others accept some sentences in which the direct object has been chopped, but reject sentences in which the indirect object has been chopped."

[12]We note here in passing that this possibility might provide a mechanism in the child's learning procedure which will predict that over time sentences of a certain kind will change from being ungrammatical to being grammatical. Historical change, of course, provides a rich source of phenomena to which this theory might be applicable, the point of view being that much change is caused by the language learning mechanism, particularly when more than one analysis is compatible with the data available to the child and with the language learning procedure. It seems possible that the theory can make precise predictions about what changes will take place.

[13]Thus this discussion does not make the usual assumption that in Chomsky-adjunction the label of the new node is identical to that of the node which it dominates. If we wished to maintain this assumption, however, then there is an alternative account of Hankamer's judgments. Suppose that the learner hypothesized that the output of DATIVE was (i).

(i)



If there is no base rule of the form VP → V NP NP then VP in (i) will be frozen. Hence neither NP which this VP dominates will be moveable by WH FRONTING.

The issue thus reduces to the question of whether only one type of adjunction is possible, with a possible ambiguity in the labelling of the newly created node, or whether there are two kinds of adjunction possible. While we have no reason to prefer one over the other at this point, it may well be that some of the learnability theorems can only be proved in the case of one alternative and not the other.

[14]Thus we would argue that this must be a transformationally derived order, as is suggested by Venneman (1973).

[15]Tom Wasow has informed us of an observation of Richard Oehrle's concerning pairs of sentences like the following.

(i) John bought a cemetary plot for his grandchildren.

(ii) John bought his grandchildren a cemetary plot.

According to Oehrle, (ii) must have the interpretation that John's grandchildren exist, while in the case of (i) John need not have any grandchildren yet. Given that this is in fact the state of affairs, it follows first that <u>for</u> causes opacity, and second that both (i) and (ii) are possible underlying structures, i.e., there is no transformation of FOR-DATIVE. However, from the second conclusion it follows that the transformation of DATIVE in the case of verbs like <u>give</u> does not cause freezing since it derives a possible base structure. Hence it may be necessary to account for the ungrammaticality of <u>*Who did you buy a book</u> by some other device than the Freezing Principle. This reformulation of the analysis of DATIVE would permit us to avoid the problem with the PASSIVE transformation raised in footnote 9 above.

On the other hand, it seems to us that (i) can be analyzed as (iii).

(iii) John bought $_{NP}$[a cemetary plot for his grandchildren].

If this is the case, then one might make the argument that the <u>for</u> which undergoes FOR-DATIVE is not an opacity causing element, while the <u>for</u> which appears in the NP in (iii) is. The difference between the two <u>for</u>'s is clear: the first is benefactive, while the second is purposive. The following examples make the distinction apparent.

(iva) John bought a box for storing his toys.

(ivb) John bought a box for his mother.

(va) *John bought storing his toys a box.

(vb) John bought his mother a box.

Example (ivb) has two interpretations. Either John bought a box to give to his mother (benefactive) or he bought a box for his mother to use (purposive). The benefactive for, since it implies immediate transfer of possession to the benefactee, requires the existence of the benefactee. The purposive for, since it implies the use of the item by someone at an unspecified time in the future, does not carry with it this requirement.

[16]These data also show that the child probably has not yet completely learned the deep structure order, since Samoan (according to Schwartz, 1972) is VSO. Note that our theory does not explain why there is a two word stage. This may very well have to do with a memory limitation, as has been suggested in the literature, or it may be a result of a child following a certain testing strategy for discovering the order of deep structure categories. (To our knowledge this latter hypothesis has not been mentioned in the literature.) It may be that the child can get more useful information about this order if he attempts to test the relative order of two categories at once, rather than three or more categories, from the outset of learning. To understand this question precisely, of course, would require considerably more analysis.

[17]Note that the only deep structure order consistent with these data and the Invariance Principle is SOV, so we might hypothesize that this is the order which Gia has established at this stage. That is, she has two

REFERENCES

Bach, E. (1965), On some recurrent types of transformations. Georgetown University Monograph Series on Language and Linguistics, 18, 3-18.

Bellugi, U. (1968), Linguistic mechanisms underlying child speech. In E.M. Zale (Ed.), Proceedings of the Conference on Language Behavior. Appleton-Century-Crofts, New York.

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

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

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

Bresnan, J. (1972), Theory of Complementation in English Syntax, Ph.D. dissertation, M.I.T., Cambridge, Massachusetts.

Brown, R. (1973), A First Language: The Early Stages. Harvard, Cambridge.

Brown, R., C. Cazden, and U. Bellugi (1969), The child's grammar from I to III. In J.P. Hill (Ed.), Minnesota Symposia on Child Psychology, Vol. II, University of Minnesota Press, Minneapolis.

Brown, R. and C. Hanlon (1970), Derivational complexity and order of acquisition in child speech. In J.R. Hayes (Ed.), Cognition and the Development of Language. Wiley, New York.

Cazden, C.B. (1965), Environmental Assistance to the Child's Acquisition of Grammar. Ph.D. dissertation, Harvard University.

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

Chomsky, N. (1970), Remarks on Nominalization. In R. Jacobs and P. Rosenbaum (Eds.), Readings in English Transformational Grammar. Ginn, Waltham, Mass.

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

Culicover, P. and K. Wexler (1973a), An application of the Freezing Principle to the Dative in English. Social Sciences Working Paper #39, University of California, Irvine.

Culicover, P. and K. Wexler (1973b), Three further applications of the Freezing Principle in English. Social Science Working Paper #48, University of California, Irvine.

REFERENCES

Bach, E. (1965), On some recurrent types of transformations. Georgetown University Monograph Series on Language and Linguistics, 18, 3-18.

Bellugi, U. (1968), Linguistic mechanisms underlying child speech. In E.M. Zale (Ed.), Proceedings of the Conference on Language Behavior. Appleton-Century-Crofts, New York.

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

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

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

Bresnan, J. (1972), Theory of Complementation in English Syntax, Ph.D. dissertation, M.I.T., Cambridge, Massachusetts.

Brown, R. (1973), A First Language: The Early Stages. Harvard, Cambridge.

Brown, R., C. Cazden, and U. Bellugi (1969), The child's grammar from 1 to III. In J.P. Hill (Ed.), Minnesota Symposia on Child Psychology, Vol. II, University of Minnesota Press, Minneapolis.

Brown, R. and C. Hanlon (1970), Derivational complexity and order of acquisition in child speech. In J.R. Hayes (Ed.), Cognition and the Development of Language. Wiley, New York.

Cazden, C.B. (1965), Environmental Assistance to the Child's Acquisition of Grammar. Ph.D. dissertation, Harvard University.

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

Chomsky, N. (1970), Remarks on Nominalization. In R. Jacobs and P. Rosenbaum (Eds.), Readings in English Transformational Grammar. Ginn, Waltham, Mass.

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

Culicover, P. and K. Wexler (1973a), An application of the Freezing Principle to the Dative in English. Social Sciences Working Paper #39, University of California, Irvine.

Culicover, P. and K. Wexler (1973b), Three further applications of the Freezing Principle in English. Social Science Working Paper #48, University of California, Irvine.

Culicover, P. and K. Wexler (1974), The Invariance Principle and universals of grammar. Social Sciences Working Paper #55, University of California, Irvine.

Ervin-Tripp, S. (1971), An overview of theories of grammatical development. In D. Slobin (ed.), The Ontogenesis of Grammar. Academic Press, New York.

Feldman, J.A. (1967), First thoughts on grammatical inference. Artificial Intelligence Memorandum No. 55, Department of Computer Science, Stanford University.

Feldman, J.A. (1969), Some decidability results on grammatical complexity and inference. Artificial Intelligence Memorandum, Department of Computer Science, Stanford University.

Feldman, J.A., J. Gips, J.J. Horning, and S. Reder (1969), Grammatical complexity and inference. Technical Report No. CS-125, Department of Computer Science, Stanford University.

Ferguson, C.A. and D.I. Slobin (Eds.) (1973), Studies of Child Language Development. Holt, Rinehart and Winston, New York.

Fillmore, C. (1968), The case for case. In E. Bach and R. Harms (Eds.), Universals in Linguistic Theory. Holt, Rinehart and Winston, New York.

Fodor, J. (1966), How to learn to talk: some simple ways. In F. Smith and G.A. Miller (Eds.), The Genesis of Language. M.I.T. Press, Cambridge.

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

Hamburger, H. and K. Wexler (1973a), Identifiability of a class of transformational grammars. In K.J.J. Hintikka, J.M.E. Moravcsik and P. Suppes (Eds.), Approaches to Natural Language. D. Reidel, Dordrecht, Holland.

Hamburger, H. and K. Wexler (1973b), A mathematical theory of learning transformational grammar. Social Sciences Working Paper #47, University of California, Irvine. To appear in the Journal of Mathematical Psychology.

Hankamer, J. (1973), Unacceptable ambiguity. Linguistic Inquiry, 4, 1, 17-68.

Higgins, F.R. (1973), On J. Emonds' Analysis of Extraposition. In J.P. Kimball (Ed.), Syntax and Semantics, Volume 2. Seminar Press, New York.

Horning, J.J. (1969), A study of grammatical inference. Technical Report No. CS 139, Department of Computer Science, Stanford University.

Kelley, K.L. (1967), Early syntactic acquisition. RAND Corporation Report No. P-3719, 1967.

Kernan, K.T. (1969), The Acquisition of Language by Samoan Children. Ph.D. dissertation, University of California, Berkeley.

Kernan, K.T. (1970), Semantic relationships and the child's acquisition of language. Anthropological Linguistics, 12, 5.

Klein, S. and M.A. Kuppin, (1970), An interactive heuristic program for learning transformational grammars. Computer Studies in the Humanities and Verbal Behavior, III, 3.

Klima, E. (1964), Negation in English. In J.A. Fodor and J.J. Katz (Eds.) The Structure of Language: Readings in the Philosophy of Language. Prentice-Hall, Englewood Cliffs, New Jersey.

Kuno, S. (1973), Constraints on internal clauses and sentential subjects. Linguistic Inquiry, IV, 3, 363-385.

Lakoff, G. (1970), Global rules. Language, 46.627-639.

Lakoff, G. (1970a), Linguistics and natural logic. Synthese, 22, 151-271.

Langendoen, D.T. (1973), Review of M.K. Burt, From Deep to Surface Structure. Language, 49, 3, 714-725.

McCawley, J. (1970), English as a VSO language. Language, 46, 2, 286-299.

McNeill, D. (1966), Developmental psycholinguistics. In F. Smith and G.A. Miller (Eds.), The Genesis of Language. M.I.T. Press, Cambridge, Massachusetts.

Mirikitani, L.T. (1972), Kapampangan Syntax, University of Hawaii Press, Honolulu.

Montague, R. (1973), The Proper Treatment of Quantification in English. In K.J.J. Hintikka, J.M.E. Moravcsik and P. Suppes (Eds.), Approaches to Natural Language, D. Reidel, Dordrecht, Holland.

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

Postal, P. (1972), On some rules that are not successive cyclic. Linguistic Inquiry. III, 2, 211-222.

Postal, P. (1974), On Raising: One Rule of English Grammar and its Theoretical Implications. M.I.T. Press, Cambridge, Massachusetts.

Postal, P. (ms., 1973), A paper which some may say is worse than it is. I.B.M. Watson Research Center.

Putnam, H. (1967), The 'Innateness Hypothesis' and explanatory models in linguistics. Synthese, XVII, 1.12-22.

Ross, J.R. (1967), Constraints on Variables in Syntax, Ph.D. dissertation, M.I.T., Cambridge, Massachusetts.

Ross, J.R. (1972), Act. in D. Davidson and G. Harmon (Eds.) Semantics of Natural Languages, Reidel, Dordrecht, Holland.

Schwartz, A. (1972), The VP-constituent of SVO languages. In J.P. Kimball (Ed.), Syntax and Semantics, Volume 1, Seminar Press, New York.

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

Snow, K. (1972), Mother's speech to children learning language. Child Development, 43:549-566.

Venneman, T. (1973), Explanation in Syntax. In J.P. Kimball (Ed.), Syntax and Semantics, Volume 2, Seminar Press, New York.

Watt, W.C. (1970), On two hypotheses concerning psycholinguistics. In J.R. Hayes (Ed.), Cognition and the Development of Language. Wiley, New York.

Wexler, K. and P. Culicover (1973), Two applications of the Freezing Principle in English, Social Sciences Working Paper #38, University of California, Irvine.

Wexler, K. and P. Culicover (1974), The semantic basis for language acquisition: The Invariance Principle as a replacement for the Universal Base Hypothesis. Social Sciences Working Paper #50, University of California, Irvine.

Wexler, K., P. Culicover and H. Hamburger (in preparation), Formal Principles of Language Acquisition.

Wexler, K. and H. Hamburger (1973), On the insufficiency of surface data for the learning of transformation languages. In K.J.J. Hintikka, J.M.E. Moravcsik and P. Suppes (Eds.), Approaches to Natural Language. D. Reidel, Dordrecht, Holland.

DISTRIBUTION LIST

Navy

4 Dr. Marshall J. Farr, Director
Personnel and Training Research Programs
Office of Naval Research (Code 458)
Arlington, VA 22217

1 ONR Branch Office
495 Summer Street
Boston, MA 02210
ATTN: J. Lester

1 ONR Branch Office
1030 East Green Street
Pasadena, CA 91101
ATTN: E.E. Gloye

1 ONR Branch Office
536 South Clark Street
Chicago, IL 60605
ATTN: M.A. Bertin

6 Director
Naval Research Laboratory
Code 2627
Washington, DC 20390

12 Defense Documentation Center
Cameron Station, Building 5
5010 Duke Street
Alexandria, VA 22314

1 Special Assistant for Manpower
OASN (M&RA)
Pentagon, Room 4E794
Washington, DC 20350

1 LCDR Charles J. Theisen, Jr., MSC, USN
4024
Naval Air Development Center
Warminster, PA 18974

1 Chief of Naval Reserve
Code 3055
New Orleans, LA 70146

1 Dr. Lee Miller
Naval Air Systems Command
AIR-413E
Washington, DC 20361

1 CAPT John F. Riley, USN
Commanding Officer
U.S. Naval Amphibious School
Coronado, CA 92155

1 Chairman
Behavioral Science Department
Naval Command & Management Division
U.S. Naval Academy
Luce Hall
Annapolis, MD 21402

1 Chief of Naval Education & Training
Naval Air Station
Pensacola, FL 32508
ATTN: CAPT Bruce Stone, USN

1 Mr. Arnold Rubinstein
Naval Material Command (NAVMAT 03424)
Room 820, Crystal Plaza #6
Washington, DC 20360

1 Commanding Officer
Naval Medical Neuropsychiatric
Research Unit
San Diego, CA 92152

1 Director, Navy Occupational Task
Analysis Program (NOTAP)
Navy Personnel Program Support
Activity
Building 1304, Bolling AFB
Washington, DC 20336

1 Dr. Richard J. Niehaus
Office of Civilian Manpower Management
Code 06A
Washington, DC 20390

1 Department of the Navy
Office of Civilian Manpower Management
Code 263
Washington, DC 20390

1 Chief of Naval Operations (OP-987E)
Department of the Navy
Washington, DC 20350

1 Superintendent
Naval Postgraduate School
Monterey, CA 93940
ATTN: Library (Code 2124)

1 Commander, Navy Recruiting Command
4015 Wilson Boulevard
Arlington, VA 22203
ATTN: Code 015

1 Mr. George N. Graine
Naval Ship Systems Command
SHIPS 047C12
Washington, DC 20362

1 Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054
ATTN: Dr. Norman J. Kerr

1 Commanding Officer
Service School Command
U.S. Naval Training Center
San Diego, CA 92133
ATTN: Code 3030

1 Dr. William L. Maloy
Principal Civilian Advisor
for Education & Training
Naval Training Command, Code 01A
Pensacola, FL 32508

1 Dr. Alfred F. Smode, Staff Consultant
Training Analysis & Evaluation Group
Naval Training Equipment Center
Code N-00T
Orlando, FL 32813

1 Dr. Hanns H. Wolff
Technical Director (Code N-2)
Naval Training Equipment Center
Orlando, FL 32813

1 Chief of Naval Training Support
Code N-21
Building 45
Naval Air Station
Pensacola, FL 32508

1 Dr. Robert French
Naval Undersea Center
San Diego, CA 92132

1 CDR Richard L. Martin, USN
Fighter Squadron 124
NAS Miramar, CA 92145

1 Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

5 Navy Personnel R&D Center
San Diego, CA 92152
ATTN: Code 10

1 D.M. Gragg, CAPT, MC, USN
Head, Educational Programs
Development Department
Naval Health Sciences Education
and Training Command
Bethesda, MD 20114

Army

1 Headquarters
U.S. Army Administration Center
Personnel Administration Combat
Development Activity
ATCP-HRO
Ft. Benjamin Harrison, IN 46249

1 Armed Forces Staff College
Norfolk, VA 23511
ATTN: Library

1 Director of Research
U.S. Army Armor Human Research Unit
Building 2422 Morade Street
Fort Knox, KY 40121
ATTN: Library

1 Commandant
United States Army Infantry School
ATTN: ATSH-DET
Fort Benning, GA 31905

1 Deputy Commander
U.S. Army Institute of Administration
Fort Benjamin Harrison, IN 46216
ATTN: EA

1 Dr. Frank J. Harris
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

1 Dr. Ralph Dusek
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

1 Mr. Edmund F. Fuchs
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA

1 Dr. Leon H. Nawrocki
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

1 Dr. J.E. Uhlaner, Technical Director
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

1 Dr. Joseph Ward
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

1 HQ USAREUR & 7th Army
ODCSOPS
USAREUR Director of GED
APO New York 09403

Air Force

1 Research Branch
AF/DPMYAR
Randolph AFB, TX 78148

1 Dr. G.A. Eckstrand (AFHRL/AS)
Wright-Patterson AFB
Ohio 45433

1 Dr. Ross L. Morgan (AFHRL/AST)
Wright-Patterson AFB
Ohio 45433

1 AFHRL/DOJN
Stop #63
Lackland AFB, TX 78236

1 Dr. Martin Rockway (AFHRL/TT)
Lowry AFB
Colorado 80230

1 Major P.J. DeLeo
Instructional Technology Branch
AF Human Resources Laboratory
Lowry AFB, CO 80230

1 AFOSR/NL
1400 Wilson Boulevard
Arlington, VA 22209

1 Commandant
USAF School of Aerospace Medicine
Aeromedical Library (SUL-4)
Brooks AFB, TX 78235

1 Dr. Sylvia R. Mayer (MCIT)
Headquarters Electronic Systems Division
LG Hanscom Field
Bedford, MA 01730

1 CAPT Jack Thorpe, USAF
Flying Training Division (HRL)
Williams AFB, AZ 85224

Marine Corps

1 Mr. E.A. Dover
Manpower Measurement Unit (Code MPI)
Arlington Annex, Room 2413
Arlington, VA 20380

1 Commandant of the Marine Corps
Headquarters, U.S. Marine Corps
Code MPI-20
Washington, DC 20380

1 Director, Office of Manpower Utilization
Headquarters, Marine Corps (Code MPU)
MCB (Building 2009)
Quantico, VA 22134

1 Dr. A.L. Slafkosky
Scientific Advisor (Code RD-1)
Headquarters, U.S. Marine Corps
Washington, DC 20380

Coast Guard

1 Mr. Joseph J. Cowan, Chief
Psychological Research Branch (G-P-1/62)
U.S. Coast Guard Headquarters
Washington, DC 20590

Other DOD

1 Lt. Col. Henry L. Taylor, USAF
Military Assistant for Human Resources
OAD (E&LS) ODDR&E
Pentagon, Room 3D129
Washington, DC 20301

1 Mr. William J. Stormer
DOD Computer Institute
Washington Navy Yard, Building 175
Washington, DC 20374

1 Col. Austin W. Kibler
Advanced Research Projects Agency
Human Resources Research Office
1400 Wilson Boulevard
Arlington, VA 22209

1 Helga L. Yeich
Advanced Research Projects Agency
Manpower Management Office
1400 Wilson Boulevard
Arlington, VA 22209

Other Government

1 Dr. Eric McWilliams, Program Manager
Technology and Systems, TIE
National Science Foundation
Washington, DC 20550

1 Dr. Andrew R. Molnar
Technological Innovations in Education
National Science Foundation
Washington, DC 20550

1 Dr. Marshall S. Smith
Asst Acting Director
Program on Essential Skills
National Institute of Education
Brown Bldg., Rm 815
19th and M St., N.W.
Washington, DC 20208

Miscellaneous

1 Dr. Scarvia B. Anderson
Educational Testing Service
17 Executive Park Drive, N.E.
Atlanta, GA 30329

1 Dr. John Annett
The Open University
Milton Keynes
Buckinghamshire
ENGLAND

1 Dr. Richard C. Atkinson
Stanford University
Department of Psychology
Stanford, CA 94305

1 Dr. David G. Bowers
University of Michigan
Institute for Social Research
Ann Arbor, MI 48106

1 Mr. Kenneth M. Bromberg
Manager - Washington Operations
Information Concepts, Inc.
1701 North Fort Myer Drive
Arlington, VA 22209

1 Centry Research Corporation
4113 Lee Highway
Arlington, VA 22207

1 Dr. Allan M. Collins
Bolt Beranek and Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

1 Dr. H. Peter Dachler
University of Maryland
Department of Psychology
College Park, MD 20742

1 Dr. Rene' V. Dawis
University of Minnesota
Department of Psychology
Minneapolis, MN 55455

1 ERIC
Processing and Reference Facility
4833 Rugby Avenue
Bethesda, MD 20014

1 Dr. Victor Fields
Montgomery College
Department of Psychology
Rockville, MD 20850

1 Dr. Edwin A. Fleishman
American Institutes for Research
Foxhall Square
3301 New Mexico Avenue, N.W.
Washington, DC 20016

ERIC

1 Dr. Robert Glaser, Director
University of Pittsburgh
Learning Research & Development Center
Pittsburgh, PA 15213

1 Dr. M.D. Havron
Human Sciences Research, Inc.
7710 Old Spring House Road
West Gate Industrial Park
McLean, VA 22101

1 HumRRO
Division No. 3
P.O. Box 5787
Presidio of Monterey, CA 93940

1 HumRRO
Division No. 4, Infantry
P.O. Box 2086
Fort Benning, GA 31905

1 HumRRO
Division No. 5, Air Defense
P.O. Box 6057
Fort Bliss, TX

1 Dr. Lawrence B. Johnson
Lawrence Johnson & Associates, Inc.
200 S. Street, N.W., Suite 502
Washington, DC 20009

1 Dr. Milton S. Katz
MITRE Corporation
Westgate Research Center
McLean, VA 22101

1 Dr. Steven W. Keele
University of Oregon
Department of Psychology
Eugene, OR 97403

1 Dr. David Klahr
Carnegie-Mellon University
Department of Psychology
Pittsburgh, PA 15213

1 Dr. Alma E. Lantz
University of Denver
Denver Research Institute
Industrial Economics Division
Denver, CO 80210

1 Dr. Robert R. Mackie
Human Factors Research, Inc.
6780 Cortona Drive
Santa Barbara Research Park
Goleta, CA 93017

1 Dr. Donald A. Norman
University of California, San Diego
Center for Human Information Processing
LaJolla, CA 92037

1 Mr. Brian McNally
Educational Testing Service
Princeton, NJ 08540

1 Mr. A.J. Pesch, President
Eclectech Associates, Inc.
P.O. Box 178
North Stonington, CT 06359

1 Mr. Luigi Petrullo
2431 North Edgewood Street
Arlington, VA 22207

1 Dr. Joseph W. Rigney
University of Southern California
Behavioral Technology Laboratories
3717 South Grand
Los Angeles, CA 90007

1 Dr. Leonard L. Rosenbaum, Chairman
Montgomery College
Department of Psychology
Rockville, MD 20850

1 Dr. George E. Rowland
Rowland and Company, Inc.
P.O. Box 61
Haddonfield, NJ 08033

1 Dr. Arthur I. Siegel
Applied Psychological Services
404 East Lancaster Avenue
Wayne, PA 19087

1 Dr. C. Harold Stone
1428 Virginia Avenue
Glendale, CA 91202

1 Mr. Dennis J. Sullivan
725 Benson Way
Thousand Oaks, CA 91360

ERIC

1 Dr. Benton J. Underwood
Northwestern University
Department of Psychology
Evanston, IL 60201

1 Dr. David J. Weiss
University of Minnesota
Department of Psychology
Minneapolis, MN 55455

1 Dr. Anita West
Denver Research Institute
University of Denver
Denver, CO 80210

1 Dr. Arnold F. Kanarick
Honeywell Inc.
2600 Ridge Parkway
Minneapolis, MN 55413

ERIC
Full Text Provided by ERIC