ESD-TR-67-633
ESTI FILE COPY

ESD-TR-67-633
ESD RECORD COPY

RETURN TO
SCIENTIFIC & TECHNICAL INFORMATION DIVISION
(ESTI), BUILDING 1211

ESD ACCESSION LIST
ESTI Call No. AL 6060
Copy No. 1 of 2 cys.

THE DEVELOPMENT OF PSYCHOLOGICAL MODELS FOR SPEECH RECOGNITION

J. A. Fodor
T. G. Bever
M. Garrett

January 1968

1 of 2
ESL S

DEPUTY FOR COMMAND SYSTEMS
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

This document has been approved for public release and sale; its distribution is unlimited.

(Prepared under Contract No. AF 19(628)-5705 by Massachusetts Institute of Technology, Department of Psychology, Cambridge, Massachusetts.)

AD0669360

LEGAL NOTICE

OTHER NOTICES

ESD-TR-67-633

THE DEVELOPMENT OF PSYCHOLOGICAL MODELS FOR SPEECH RECOGNITION

J. A. Fodor
T. G. Bever
M. Garrett

January 1968

DEPUTY FOR COMMAND SYSTEMS
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

(Prepared under Contract No. AF 19(628)-5705 by Massachusetts Institute of Technology, Department of Psychology, Cambridge, Massachusetts.)

FOREWORD

This report covers research done with support in whole or part from a United State Air Force grant (Contract Number AF 19(628)-5705) Project No. 2801 during the period September 1, 1965, to August 31, 1967. Report submitted October 14, 1967, by J.A. Fodor, T.G. Bever and M. Garret, Massachusetts Institute of Technology, Department of Psychology, Cambridge, Massachusetts, Mr. John B. Goodenough (ESLFS) was the Contract Monitor.

This technical report has been reviewed and approved.

Sylvia R. Mayer
SYLVIA R. MAYER
Project Officer

James L Blilie
JAMES L. BLILIE, Lt. Col., USAF
Chief, Engineering Support Branch

Abstract

This report has two principal sections. In Section I the theoretical issues relating to the development of syntax recognition routines based on psychological models of human speech recognition are discussed and the relevant psychological literature reviewed. The research reported in this section deals with attempts to relate various syntactic variables to measures of the perceptual complexity of sentences. The results of the research indicate that (1) analysis by synthesis routines are probably not appropriate to models of the system employed by speakers of natural languages for speech recognition, (2) the complexity of sentences is not related in any direct way to the number of operations required by a grammar to produce them (3) both the lexical structure of verbs and the relation of certain other lexical formatives to transformational operations of the grammar are significantly related to the ease of understanding sentences. Section II of the report deals with research concerning the perceptual segmentation strategies employed by speakers in their analysis of continuous speech. The report details the development of a particular investigative technique and its application to the determination of segmentation strategies. The research reported indicates that speech signals are initially analyzed into segments that approximate the clause. The experiments reviewed explore the relationship between the constituent structure of sentences and this clause-like segmentation.

Introduction

Solving the problem of communicating with machines in a natural language is a precondition to the use of computers for a variety of important purposes: Among them information retrieval and machine translation. The problem has not, however, thus far received a solution that is fully satisfactory from either the theoretical or the engineering standpoint. On the contrary, the recent development of conceptually rich analyses of the grammatical structures of natural languages has dramatized the difficulty of that part of the problem which involves machine recovery of the syntactic structures of sentences. On the one hand, it appears sufficiently clear that the decoding of a sentence (e.g., semantic interpretation in the sense of Katz and Fodor, 1963) presupposes the recovery of the syntactic relations which obtain among its parts. But it equally appears that the recovery of such relations involves assigning to the sentence a 'structural description' of considerable abstractness and complexity. A pressing problem in the employment of natural languages for man/machine communication is thus that of designing an algorithm for the mechanical assignment of grammatical descriptions to arbitrary sentences drawn from a natural language.

Two kinds of approaches have dominated research on algorithms for syntax recognition. The first has treated that problem as one of 'artificial intelligence'. In particular, an attempt has been made to design computer programs for the assignment of structural descriptions to sentences. These programs have not been viewed as primarily simulations of human sentence processing capacities, nor have the theorists who have devised them been primarily concerned with the psychological phenomena of human sentence recognition (for a review, see Keyser and Petrick, 1967).

The second approach to machine recognition of natural languages has been via a sustained investigation of the processes humans employ in sentence perception. A variety of the aspects of such psycholinguistic research have proved pertinent insofar as the research has had for its goal the characterization of the computational capacities fluent speakers employ to process ordinary language inputs. The work on the present project has been entirely directed to the analysis of human language processing capacities. As this report will make clear, investigations of certain relations between the syntactic structure of sentences and the complexity of sentences have suggested a number of hypotheses about the sorts of strategies speakers may employ for sentence recognition. These investigations have led to three general conclusions which this report will seek to substantiate:

1) In all probability, analysis by synthesis procedures are not the best types of models for the way human speakers use grammatical information for purposes of sentence recognition. A more abstract relation between the grammar and the sentence recognition device than such models represent is probably required for adequate simulation of human information processing during sentence recognition. In par-

ticular, information about the 'lexical structure' of individual formatives is very likely essential to the sentence recognition process.

2) The complexity of sentences is, correspondingly, not predicted by such simple parameters of their structural descriptions as length. Use of grammatical variables in complexity measures for stylistic purposes or for the facilitation of information exchange should not, therefore, depend primarily on the use of such variables.

3) Sentence recognition procedures probably involve pre-analysis of sentence units into elements approximating the clause. The precise character of this pre-analysis is of considerable theoretical and practical significance and would justify sustained research.

This report deals with the evidence for the three preceding conclusions.

Section I: Syntactic Analysis

If we consider the syntactic models proposed by contemporary generative grammarians, we notice that there are two aspects of any syntactic structural description. On the one hand, a sentence is assigned a surface structure, in terms of which it is phonologically interpreted (and which heavily determines its perceived prosodic features) and on the other hand, it is assigned a 'deep structure' in which the grammatical relations holding between the segments of the sentence are marked. (See fig. 1) The problem of syntax recognition is the problem of characterizing a device which takes a string of morphemes that constitutes a sentence of some language and assigns to it the appropriate deep and surface structure descriptions.

We can assume that the sentence recognizer must have at least two kinds of information at its disposal: Information about the acoustic properties of the input it is processing, and 'background information' about the general properties of the system that produced the input. This is to say that the device must be pictured as 'knowing' not only the acoustic character of the present input signal, but also as

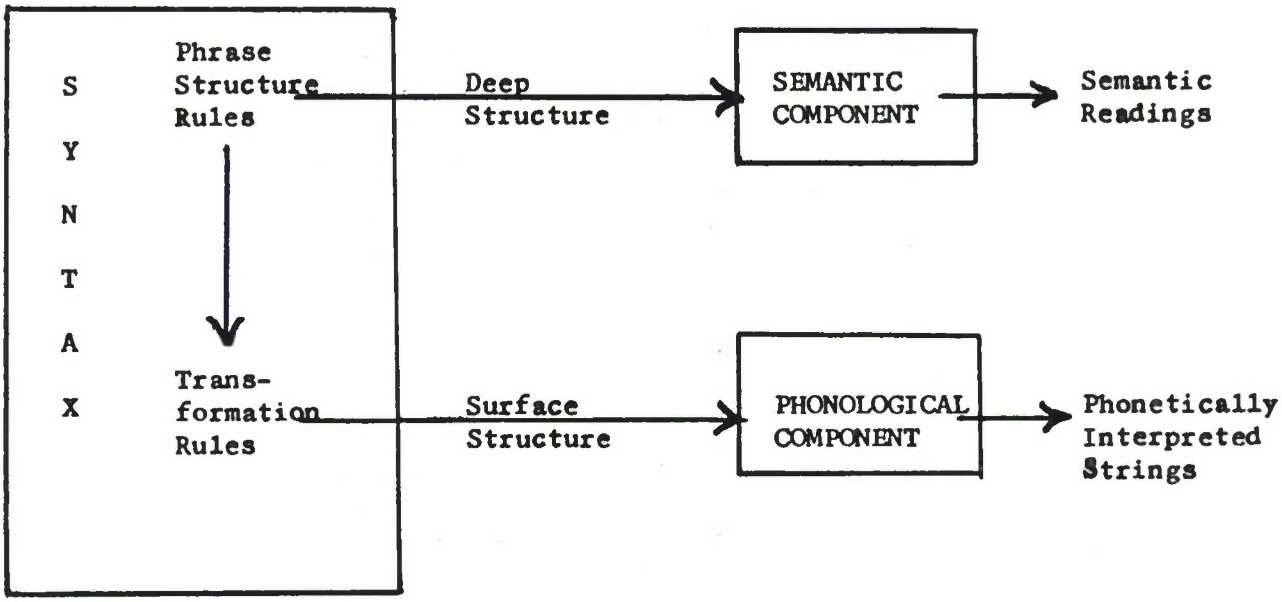


Figure 1. The components of a linguistic description.

having knowledge of the domain from which the signal is drawn. This 'background information' may be thought of as determining a perceptual 'set': a system of presuppositions the device makes about the properties of the signals it is required to recognize.

In order to recover the syntactic structure of a sequence of formatives a syntax recognition device must employ information about the grammatical structure of the language from which the sequence is drawn. This is most easily seen by considering that, by definition, the representation of a sentence a recognition device is required to recover is a syntactic structural description. Such a description is formulated in terms of familiar linguistic constructs like phrase markers and transformations. However, to say of a sentence that it contains a noun phrase, or that it is a transform of a certain type, etc., is at least to say of it that it bears a specified relation to other sentences in the language.

Thus, in providing a structural description one is implicitly characterizing the relation between the sentences described and other linguistic structures. It is patent, therefore, that a device capable of recovering structural descriptions will have to have available to it a representation of the structural relations permissible in the language, which is in effect to say that the syntax recognition device will require access to information of the type formulated by a grammar.

It is possible to conceptualize the syntax recognition device as defined by an infinite set of triples, each of the form S_i G_i SD_i where S_i is the ith sentential input, G_i is a specified grammar, and SD_i is the structural description G_i provides for the ith input. That is, we think of the recognition device as answering the question: what structural description does grammar G_i assign S_i for each of infinitely many S_is?

Among the most interesting questions we are now able to formulate about syntax recognition devices is, "How do such devices employ the information formulated in a grammar?" What makes this question interesting is, first, that we must solve it if we are to construct such a device, and second that it implicates the general problem of the nature of a grammar considered as a psychological theory. That is, a way of making the question, "What kind of a psychological model is a grammar?" more precise is to reformulate it as, "What is the exact relation between a grammar and a recognition routine?"

The least abstract way of realizing a grammar as part of a perceptual device is to employ some version of an analysis by synthesis procedure. In such procedures, the grammar is used to generate a 'search space' of candidate structural descriptions which, in effect, are tested one by one against the input string. (See Halle & Stevens, 1964; Matthews, 1962.) The grammar is employed as a source of internally generated sentential analyses. These analyses are compared with some representation of the input signal to be recognized. The comparison procedure halts when a match is effected between the internally generated signal and the input. The structural analysis of the input is determined by reference to the grammatical rules employed in generating the successful matching signal.

It should be noted that such procedures are, in principle, capable of being made fail proof. That is, appropriate formal maneuvers will guarantee that these procedures will in fact find the correct grammatical description of any input. Thus analysis by synthesis procedures are probably capable of providing a representation of the competence of an ideal hearer. The important question is whether they are capable of providing the <u>correct</u> representation.

Clearly, an acceptable recognition model must do more than assign structural descriptions correctly in a finite time. It ought also to predict detailed features of speakers' performance, such as the relative difficulty they have in processing sentences of various syntactic forms. That is, it is a reasonable constraint to place upon such a model that if, in computing structures of type A it requires more operations, or more complicated operations than those it requires for computing structures of type B, then speakers ought to have more difficulty understanding A structures than B structures, <u>ceteris paribus</u>.

If we indeed place upon our recognition models the requirement that they correctly predict 'performance properties' like the complexity of sentences, we can derive an important testable consequence of analysis by synthesis models. The prediction such models make is this: all other things being equal, the more grammatical operations required to generate a sentence, the more difficult the sentence ought to be to understand. It should be noted that this prediction is characteristic not only of analysis by synthesis models but, in general, any model that assumes either that in understanding a sentence the hearer literally runs through the set of operations a grammar employs in generating it, or at least that for each such grammatical operation there is a corresponding psychological operation in the decoding process.

A number of early psycholinguistic studies of generative grammar appear to have been motivated precisely by the hypothesis that, insofar as sentential complexity is a function of syntactic variables, the complexity of a sentence is measured by the number of grammatical rules employed in its derivation. We shall refer to this hypothesis as the Derivational Theory of Complexity (DTC).

DTC can be made explicit in the following way. Consider a generative grammar of the language L and a sentence S in the range of G. It is possible to define a metric which, for every pair G_i S_i specifies the number N of rules G_i requires to generate S_i. (It makes no difference, for present purposes whether the metric is defined, e.g., over the number of grammatical rules employed, over the number of elementary operations, over the number of transformations, etc. All these measures give convergent predictions for the studies we shall be discussing.) DTC in its strongest form is the claim that N is an index of the complexity of S_i. In particular, two sentences assigned the same number are predicted to be equally complex, and of two sentences assigned different numbers, the sentence predicted to be most complex is the one to which the larger number is assigned.

We may commence our discussion of the character of the sentence recognition device by examining the status of the experimental evidence for DTC since as we have seen, the adequacy of a wide variety of types of recognition models is at issue in any test of DTC. If we can show that DTC is true, then it is plausible to argue that a recognition model which uses the grammar in the sort of way characteristic of analysis by synthesis is very likely to be correct. If, on the other hand, we can show DTC to be false, we have some reason for being suspicious of such models.

A wide variety of psycholinguistic studies which assume the generative model of grammar may in fact be interpreted as having some bearing on DTC. We shall presently turn to a review of some of these studies. First, however, a word needs to be said about some of the grammatical assumptions that underlie these experiments.

It was at one time widely supposed by psycholinguists that one particular sentence form (the simple, affirmative, active, declarative, hereafter written as D) is of pre-eminent grammatical significance. This was because the D was considered to serve as the transformational source or basis for the production of all other sentence types. This assumption was due, in part, to a misunderstanding about the nature of the grammatical formalism and in part to a linguistic account which has since been discarded as mistaken.

In the earlier formations of the grammar (Chomsky, 1957) the output of the phrase structure component was modified by a transformational component just as indicated above in figure 1. However, certain functions now assigned to the phrase structure component of the grammar were at that time handled transformationally. In particular, such complexities of sentence structure as question (Q) negative (N) and passive (P) were then introduced by transformational rules which operated optionally on a given Phrase marker (P-marker), i.e., base structure tree. The general notion was that when no such options were exploited mandatory rules carried the P-marker into a D.

In current versions of the grammar, however, such forms as Q, N, and P are differentiated not only in their transformational histories, but also in their base trees. In particular, lexical markers representing these forms are introduced by base structure rules and the base structures into which they are introduced are then mandatorially transformed into the appropriate surface trees. Thus Q, N, P, and their family relations are _not_ now supposed to have a precisely identical base form.

Given the 1957 view that all sentences belonging to a common sentence family have the same base tree, and that the least transformationally elaborated form of that tree is the D, it is easy to make the mistake of supposing that the grammar is somehow committed to a special psychological centrality for that sentence form. In particular, psychologists widely confused the assertion that the base form of the D is also the base form of Q, P, N, etc., with the quite different assertion that the D itself is the form from which transformationally complex sentences are derived. This confusion is mentioned only because it has often led to misinterpretation of the work on the psychological relevance of transformational relationships among sentence forms. In particular, though there may well be _psychological_ evidence for the claim that Ds are in some sense basic sentence types, there is _no_ strictly linguistic motivation for this claim and no such claim was in fact made even during the early period of transformational analysis.

In short, the studies we are about to review can be brought to bear upon current versions of the grammar only with a little reinterpretation. If grammatical operations are psychologically real, then Ds ought, in general, to be less complex than their transformed counterparts--not, however, because the complex sentences are grammatically transforms of Ds, but only because, in general, fewer grammatical operations are required to specify a D than to specify a Q, P, N, etc. To this extent the identification of Ds with base structures is harmless since it produces predictions that are often convergent with the ones that would be generated by a more sophisticated understanding of the grammar.

A review of some sentence complexity studies

The first experiment aimed at revealing the psychological relevance of transformational operations, and hence the first experiment that needs to be considered in an evaluation of DTC, was performed by Miller, McKean and Slobin (Miller, 1962). Their attempt was to demonstrate that the relative difficulty of producing certain systematic changes in the surface character of sentences could be predicted as a function of the relative transformational complexity of the sentences as measured by the number of transformations they contain.

_S_s were given two columns of sentences and required to match the members of one with their counterparts (systematically altered) in the other. The sentences in the two columns were assumed to differ by one or two transformational operations and were, of course, randomly distributed between the columns. For example, a sentence in one column might have been *Jane liked the old woman* (D) while its counterpart in the other column was *The old woman was liked by Jane* (P). Before beginning the task, *S*s were instructed which operations were to be performed in a particular pair of columns (as Active: Passive, or Affirmative: Negative, etc.). Each pair of columns tested either one operation or a pair of operations (as in the case of Affirmative Active: Passive Negative). Base search time was determined by having *S*s locate untransformed versions of sentences in a scrambled list (as D: D, or P: P).

The assumption was that differences in the time taken to perform these matches would reflect differences in the time taken to perform the various transformational operations (or their inverse). That is, in order to find a match for a given sentence, it is assumed that *S* must perform the relevant transformation and then look for a new sentence which matches the result of that transformation.

Of the relationships among the types of sentences studied by Miller et al., two were considered to require two operations, and four, only one operation. Where a P sentence was required, for instance, and an N sentence given, it was assumed that one "undid" the work of the negative transformation and then applied the passive. But for the same initial condition (given an N sentence) where a PN sentence was required, *S*s were assumed to apply the passive while not being required to "undo" the result of the negative transformation. On this view, it would be predicted that the results of the experiment should find D: (N or P) and PN: (N or P) comparable while D: PN and N:P should both be more difficult (although comparable to each other). In fact, the order in the results was just that (see table 1). The view that the difficulty of processing these sentences can be indexed by the number of steps in their derivational history thus seemed to be supported. Further, there was the suggestion that these operations produced a linearly additive complication (sentences involving both negative and passive transformations required a time approximately equal to the sum of the average time required for negative and passive applied separately).

Because of some dissatisfaction with the pencil and paper method used above, Miller and McKean carried out a refined version of this same experiment (1964). In the latter version a sentence was presented tachistoscopically; when *S* had performed the required transformation of the sentence, he pressed a button which (1) presented a search list and (2) stopped a timer which had been started on presentation of the original sentence. In this technique the search time is separated from the presumed processing time. In this way an independent measure (*S*'s subjective estimation) of the transformation time is obtained, and variance introduced by the search is eliminated.

Table 1

"Average transformation times"

(from Miller & McKean, 1964, p. 300)

Sentence Change	Seconds more than base search time
D - N	1.13
D - P	1.43
P -- PN	1.66
N -- PN	1.87
D - PN	2.74
N -- P	3.50

The results obtained here were comparable to those with the pencil and paper method. However, though the equivalence sets were ordered as before, there was some rearrangement among their members. That is, those sentences requiring two operations remained more complicated than those requiring one, but the rank order within the classes was altered.

In fact, however, the two preceding studies cannot be supposed to show anything about the relationship of surface and deep structure, as these notions are characterized by the grammar, nor do they yield any very dirrect information about the relative difficulty of the various transformations *as they are formalized in the grammar*. For, in order to maintain the predictions confirmed by the results, one has to assume that *S*s were operating with some *ad hoc* strategies which do *not* apply to the structures specified as domains for true grammatical transformations. The relationship between what *S*s are doing in the experimental situation and the operations defined by a grammar is thus quite unclear.

Consider the sequence of operations which would have to occur if *S*s were using the transformations as they are (or were) formulated in the 1957 version of the grammar. First, *S* is required to recover the underlying ("kernel") structure of the sentence he is presented with. Having recovered this deep structure, he can then perform the transformations necessary to produce one or another derived sentence form. Notice that he *must* recover deep structure in order to apply the transformations, for it is only for deep structure domains that transformations are defined. Having recovered the deep structure, *S* has then to generate the required derived structure. This may include carrying out any transformational operations he has "undone" in recovering the deep structure (UP). This will be the case whenever a structural feature present in the stimulus sentence is also found in the search sentence. For example, on Miller and McKean's way of counting N : P is more complex than N : PN--presumably because it was supposed that one must undo the effect of the negative transformation in one case but not in the other. But, in fact, to recover the deep structure in *either* case, one would have to perform an inverse negative transform.

Hence:

Neg. sentence __________________ UP

UP___generation of________ (apply passive transform)________Passive Sentence
surface structure

AND:

UP___generation of___ (apply passive and negative transforms)________Negative Passive Sentence
surface structure

This would make it appear that N : NP is <u>more</u> complicated than N : P.

In passing, we may note that accepting current versions of the grammar will not yield Miller and McKeans' prediction. If we assume that the deep structures of negative sentences contain a "Negative Marker" that must be deleted in going from N to P, then N : P is represented as requiring the <u>same</u> number of operations as N : NP. Hence:

$\text{Neg. sentence}\,\underline{\hspace{3em}}\,\text{UP}_{\text{neg}}\,\underline{\text{(delete neg)}}\;\text{UP}\;\;\underline{\text{ (add pass)}}\;\;\text{UP}_{\text{pass}}.$

$\text{UP}_{\text{pass}}\,\underline{\text{(do pass)}\hspace{3em}}\;\;\text{Passive sentence}$

AND:

$\text{Neg. sentence}\,\underline{\hspace{6em}}\;\text{UP}_{\text{neg}}\,\underline{\text{ (add pass)}\hspace{3em}}\;\;\text{UP}_{\text{pass neg}}$

$\text{UP}_{\text{pass-neg}}\,\underline{\text{(do pass, do neg)}\hspace{4em}}\;\;\text{Negative-Passive Sentence}$

We are not, of course, suggesting that it is reasonable to claim that subjects in Miller and McKean's experimental paradigm <u>do</u> get from N to NP by first deleting and then restoring the negative marker. The point is rather that an interpretation of miller and McKean's results confronts one with a dilemma. If these results are to be understood as relevant to such grammatical operations as passivization, negation, etc., then the subjects must indeed be understood to be applying these operations to base structures, since those are the only structures for which the grammatical transformations are defined. If, on the other hand, we accept the data, it looks as though subjects do no such thing: e.g., they go from negative to negative passive in 'one move' rather than three. Hence, if one is to account for the Miller and McKean data, it seems necessary to hold that the strategies subjects employ in the experimental situation are <u>not</u>, in fact, directly derived from the rules of the grammar. What is needed, of course, is some theoretically motivated way of saying which of the grammatically defined operations are psychologically relevant -- e.g., which ones should be counted in making predictions about the difficulty of perceiving, producing, or recalling sentences. Unfortunately, however, we have no very good account of how grammatical information is brought to bear in the development of the strategies employed in experimental situations like that of Miller and McKean. Indeed, if their data were to be relied upon, they suggest that any such account might turn out to be quite complicated.

Another experiment (McMahon, 1963) with negative and passive sentence types provides information of a different kind. McMahon's test of these sentence types required <u>S</u>s to judge whether a presented sentence was true or false. His sentences were of forms such as, <u>5 precedes 13</u> or <u>3 is preceded by 7</u>, etc.

McMahon found that it required longer to judge the truth of negative sentences than of passives or active affirmatives. The order of difficulty from easiest to most difficult was:

Act Aff ------------------ small difference
Pass Aff
------------------ large difference
Act Neg ------------------ small difference
Pass Neg

These results are interesting when compared to one of those of Miller and McKean. Miller & McKean found their <u>shortest</u> search times to be associated with negation. The transformation easiest to perform is thus apparently the one whose truth value is most difficult to determine. It is not obvious how to interpret this difference between the Miller-McKean results and those of McMahon unless it is assumed that semantic considerations play a central role when questions of truth value are raised in the experimental situation. It is worth noting that on that assumption, one can account for a fact that is observed in a wide variety of experimental situations: the similarity of <u>S</u>s performance on synonymous forms.

An extensive study of the effects on recall of differences between D, Q, P, and N sentences was carried out by Mehler (1964). This study, and a number of others inspired by it, are more clearly concerned with interactions between syntactic factors and memory than with DTC <u>per se</u>. Nevertheless, the convergence between Mahler's results and those of McMahon and of Miller et al. is striking.

In Mehler's study, <u>S</u> was exposed to lists of eight sentences, the sentences being presented seriatum. Each list contained one each of the following sentence types: D, Q, P, N, PN, PQ, NQ, and PQN. After he had seen all eight sentences <u>S</u> attempted to recall them. The sentences were then presented again and <u>S</u> was again tested for recall--and so on for five trials.

In these results, the sentences were ordered roughly as they were in the Miller and McKean study, that is, the order of sentence difficulty in a recall task turns out to be similar to ease of transformation. D was much the best recalled, those cases with only one transformation were next, followed by those with multiple transformations. This indicated that for these types of sentences, length of derivations is related to the ease of recall. It is worth noting that in none of these studies is sentence length an adequate predictor of performance (for example, questions like <u>Did the boy eat the apple</u>? characteristically elicit worse performance than affirmatives like <u>The boy has eaten an apple</u>).

There are three more types of studies which can be interpreted as particularly relevant to our evaluations of DTC. The first of these is by Savin and Perchonock (1965). It represents an attempt to relate the storage requirements of the various sentence types to aspects of their structural descriptions. The assumption is that the greater the complexity of a sentence's description, as indexed by the number of rules required for its production, the greater will be the demands on storage. Savin and Perchonock sought to determine this difference by requiring *S*s to recall both a sentence and a set of unrelated words. *S*s were presented with a sentence followed by a string of eight unrelated words. *S*s had to repeat the sentence and then as many of the eight words as he could recall. The number of words successfully recalled when the sentence was correctly recalled was the measure of storage requirements for that sentence type. Savin used the same types of sentences as did Mehler plus emphatic (E-heightened stress on *aux*) and *wh* forms (What has the boy eaten?).

The ordering of sentence types from the results of this experiment was as shown in table 2 (ordering is presumably from least storage requirement to greatest):

Table 2

Mean number of words

recalled for each sentence type

Sentence type	Words recalled	
D	5.27	(Presumably one transformation)
Wh	4.78	
Q	4.67	
P	4.55	
N	4.44	
Q_{neg}	4.39	
E	4.30	
PQ_{neg}	4.02	(Presumably two transformations)
PQ	3.85	
EP	3.74	
NP	3.48	

Note that in every case those sentences with one transformational operation required less storage (interfered less with recall of word strings) than those with two operations and that D was least interfering of all.

More strikingly, Savin and Perchonock found a constant effect of given transformations. That is, a particular transformational operation apparently took the same storage space whatever other transformations it was associated with (thus Q added the same degree of difficulty whether added to P or E or D, etc.). In this experiment as in the ones previously discussed, length of the sentence in words will not account for the results.

The results of Savin and Perchonock and the Mehler studies are most persuasive as an argument for the relation of aspects of the derivational history provided by the grammar to the determination of sentence complexity. It should be noted, for example, the kinds of objections raised to the Miller-McKean results do not apply here since the ordering of sentence types depends only on the number of operations required for deriving the sentences and not upon questionable assumptions about the operations required to change one sentence type to another.

The results just reviewed are examples drawn from a variety of studies which, taken together, offer a rather persuasive argument for some view of the relations between syntactic variables and sentence complexity like the one embodied in DTC. Yet, despite the persuasiveness of these experimental findings, and of the analysis by synthesis model of syntax recovery with which DTC comports, we shall presently see that there are very strong considerations militating against accepting any version of DTC that has thus far been proposed. Since we will be urging the rejection of DTC and, ultimately, of analysis by synthesis models of perceptual processes at the syntactic level, it is worth while to pause to underscore the distinction between DTC and the grammar *per se*.

It must be emphasized that adherence to the transformational view of grammar does *not ipso facto* require accepting DTC, nor would the confirmation of DTC be in any direct sense a confirmation of the grammar. In particular, the grammar is not committed to a correspondence between the complexity of the *grammatical* operation of converting a base structure into a surface structure and the complexity of the *perceptual* operation of converting a surface structure into a base structure.

Though it is widely accepted that grammars are theories neither of speakers nor of hearers, much past work in psycholinguistics suggests a failure to appreciate the full implications of this fact, namely, that the grammar constrains only the *output* of the perceptual model (i.e., the latter must recognize the set of structural descriptions generated by the former). In particular, the grammar in no way imposes constraints either on the set of facts about the input the perceptual model may exploit in computing structural descriptions, or upon the operations it runs through in the course of such computations. These constraints are consequences not of the grammatical analysis, but of the particular way we chose to realize the grammar in the perceptual model.

What is at issue is ultimately the questions whether psychological reality is to be claimed not only for the structural descriptions the grammar enumerates, but also for the rules and operations employed by the grammar to effect that enumeration. It is critically important to understanding the nature of the relations between linguistic and psycholinguistic theories to grasp the difference between these claims. There exists abundant intuitive and experimental evidence for the psychological reality of structural descriptions. We shall now see that the status of the evidence for the psychological reality of the grammatical operations employed in generative structural descriptions is hightly equivocal. Let us remark again that, insofar as this evidence is negative, it militates against any view which, like the analysis by synthesis model, employs the grammar directly as a source of candidate analyses in the assignment of structural descriptions. That is, it militates against one kind of theory of the way that the information represented by a grammar might be explained for purposes of perceptual integration. It does not, however, impugn the accuracy with which the grammar represents that information.

Recent results owing to Bever, Fodor, Garrett and Mehler suggest that increase in performance complexity is not uniformly associated with increase in the length of the derivational history of a sentence in certain cases where synonymy and morpheme length are fully controlled. In pilot experiments, a task analogous to that employed by Savin and Perchonock was used, except that S was required to identify one of a series of four tones heard prior to a sentence. Under this condition, no performance differences could be reliably associated with transformational differences like the one between John phoned up the girl and John phoned the girl up or like the one between The bus driver was fired after the wreck and The bus driver was nervous after the wreck. Further investigations in this area are currently being undertaken.

In an experiment by Fodor, Jenkins and Saporta (1965) similar sorts of structural differences were evaluated for their effect on performance variables. For example, variants of the comparative such as, (1) John swims faster than Bob swims, (2) John swims faster than Bob, (3) John swims faster than Bob does. In terms of DTC, the order of complexity would increase from (1) to (3); sentence (1) is fewer transformational steps removed from its base structure than are either (2) or (3); similarly, (3) requires one step more than (2) in its derivation ('do support' after deletion of the second verb). When recognition latencies were measured for sentences varying in this way, however, sentences like (1) turned out to be most difficult, while those like (2) and (3) were indistinguishable. Similar tests for differences between sentences with displaced particles and those with particles in the untransformed position showed no significant differences in recognition latencies.

Another result of Miller and McKean is relevant in light of the preceding. They carried out tests of sentences that differed in the expansion of the verb auxiliary. Sentences differed as follows: *Joe warned the boy*; *Joe had warned the boy*; *Joe was warning the boy*; *Joe had been warning the boy*. These differ by (1) adding *had*, (2) adding some form of *be*, of (3) adding both *had* and *be*. The predictions here were analogous to those for the transformations--difficulty was expected to increase with complexity of the verb phrase. The result, however, was that there were no significant differences as a function of the changes introduced, with the exception that addition of *had* was very much easier than any other operation.

These results suggest at least that there is no *general* correspondence between the complexity of the sentence and the number of grammatical operations involved in its derivation. It is, of course, compatible with the suggestions that for some reasons, operations involved in specifying base structure entities (like aux) 'cost' less than transformational operations. That is, that the operations in the grammar which are related to Q, N, P, etc., are somehow more relevant to the determination of performance variables than the operations involved in the development of the base tree. *Prima facie*, however, it is extremely difficult to understand why this should be the case. Base structure rules are, after all, simply specially restricted kinds of transformations. The elementary operations in terms of which the former are specified are the same elementary operations used to specify the latter. Why, then if performance variables are sensitive to the number of operations in a derivational history, should they be responsive to transformational elaboration but not to elaboration in the base?

The present case provides a good example of the problems that arise when one attempts to interpret negative results concerning the psychological reality of grammatical operations. It is difficult to know whether to blame the experimental procedures, the grammatical theories, or the specifically psycholinguistic assumptions about how grammatical and psychological constructs ought to interact in the perceptual model.

That we are, at any event, not dealing with experimental artifact is suggested by the similar findings of Mehler who tested the effect of expansion of *aux* on recall. More complicated expansions (involving modals) were used than those of Miller and McKean (*The boy hit the ball*; *The boy has hit the ball*; ...*could hit*...;...*was hitting*...

could have hit...could be hitting...has been hitting...;...could have been hitting...).

Subjects did not show any significant tendency to learn sentences with simpler aux expansions more readily than those with complex expansions. Further, there was no tendency for response errors to be simplifications of the presented sentence; a tendency often found when transformational complexity is the main sentence variable. 155 errors were simplifications, but 157 involved complications of the aux.

Experiments on sentence complexity

The results we have just reviewed suggest something less than a perfect correspondence between the number of grammatical operations required to generate a sentence and the relative complexity of that sentence. They, thus, cast some doubt upon the derivational theory of complexity; i.e., upon the view that an isomorphism exists between generative grammatical operations and psychological operations involved in decoding. Two urgent questions thus present themselves to the theorist confronted with these results. In the first place, it is essential to provide direct and reasonably conclusive evidence, pro or con, on the crrectness of DTC. Second, if we are to abandon DTC, what is to be put in its place? In particular, if the grammatical operations involved in the generation of a sentence are not isomorphic to decoding operations, what alternative models of decoding might it be reasonable to envision?

In experiments discussed here, we have attempted to directly investigate the relation between the complexity of a sentence and the length of its derivational history. Our initial hypothesis was that the complexity of a sentence is a function not (or not only) of the transformational distance from its base structure but also of the degree to which the arrangement of elements in the surface structure provides clues to the relations of elements in the deep structure.

To a certain extent this hypothesis about complexity and DTC yield convergent experimental predictions. This is because increasing the distance from base to surface structure (e.g., adding transformations) tends, on the whole, to obliterate surface structure clues to deep structure relations. Transformations deform the base structure trees that they apply to. Nevertheless, insofar as increasing the number of transofrmations tends to increase complexity, it is suggested that this is not because of the increased transformational distance between base and surface structure *per se*, but rather because of the consequent obliteration of the surface structure clues upon which the reconstruction of deep structure depends. We shall see that profound theoretical questions about the relation of a generative grammar to a psychological model of speech perception hang upon this distinction.

In experiments 1 through 5, the experimental procedures turn upon the fact that certain lexical items, which appear in sentences as the result of embedding transformations, are optionally deletable. For example, the "self-embedded" sentences (1) and (2) are fully grammatical and fully synonymous. In (1), however, the word whom, which results from the embedding of the dog bit the man in the man died, has been deleted.

(1) The man the dog bit died.

(2) The man whom the dog bit died.

It was hypothesized that the presence of the relative pronoun in sentences like (2) should make them perceptually less complex than corresponding sentences like (1). This is because the relative pronoun provides a surface structure clue to semantically crucial deep structure relationships that obtain between the clauses. Specifically, the embedding transformation, which introduces the relative pronoun, can apply only where certain grammatical relations hold between the NP's in the sentence.

For the lexical material employed, the surface structure constellation NP_1 relative pronoun, NP_2 can appear grammatically only where NP_2 is the base structure subject of a transitive verb of which NP_1 is the base structure object. Thus, the relative pronoun in a sentence like (2) provides an immediate clue to the deep structure of the sentence. This clue is absent in sentences like (1), where the initial constellation of NP's is compatible, e.g., with the list structure displayed in (3).

(3) The man the dog and the girl went swimming.

The restriction of this analysis to sentences containing lexical material of the type employed is essential. There exist types of verbs that can follow the constellation NP_1 rel NP_2 in sentences where the NP's do not exhibit the sorts of subject/object relations exemplified by (1) and (2). For example, verbs that take complements, as in (4), middle verbs, as in (5), and verbs that take indirect objects, as in (6).

(4) The boy that the man wanted Joan to meet was ill.

(5) The amount that the book cost was excessive.

(6) The girl that the boy gave the book to was pretty.

On the other hand, an heuristic based on the presence of the relative pronoun in the sequence NP_1 rel NP_2 will work without exception for sentences containing only simple transitive verbs. The interaction of this heuristic with the verb structure of the stimulus sentence turns out to have extremely interesting consequences to thich we shall presently return. In short, if the heuristic (i.e., "given a sequence $\underline{NP_1 \text{ rel } NP_2}$, assume the NP's are related to each other as object and subject respectively, of the same verb.") is to be useful to *S*s, it must interact with their knowledge of the character of the verb(s) in the stimulus sentences.

We have seen that self-embedded sentences like (1) and (2) above exhibit features that are relevant to testing the significance of certain types of surface structure clues to deep structure configurations. This type of sentence was employed in our experiments not only for this reason, but also because, with iteration of the self-embedding operations, *S*s have great difficulty in understanding them. This provides an opportunity for the presumed facilitatory effects of surface structure clues to be revealed more strongly than in the case of sentences which *S*s find easy to understand.

In the first experiment described below, the presence of the presumed facilitatory effect of the presence of relative pronouns is demonstrated. Subsequent experiments deal with possible alternative explanations of this effect.

Experiment 1. Nine pairs of sentences with two embeddings each were constructed. The list below gives the sentences with the optionally deletable items in brackets.

1. The pen (which) the author(whom) the editor liked used was new.

2. The tiger (which) the lion (that) the gorilla chased killed was ferocious.

3. The boats (which) the rocks (that) the waves covered sank were large.

4. The cigarette (which) the match (that) the flame ignited lit smoked.

5. The car (which) the man (whom) the dog bit drove crashed.

6. The window (which) the ball (that) the boy threw hit broke.

7. The bicycle (which) the boy (whom) the policeman stopped turned was lost.

8. The shot (which) the soldier (that) the mosquito bit fired missed.

9. The man (whom) the girl (that) my friend married knew died.

Two groups of sentences were derived from this list. The group in which the relative pronouns were deleted will be referred to as Group I; the group in which the relative pronouns are retained will be referred to as Group II. Sentence Groups I and II were tape-recorded for presentation to Ss. In the materials used in experiment 1, the sentences were read in what E judged to be a "neutral" manner; that is, the intonation contour was kept flat and expressive devices like emphatic stress, long pause, etc., were avoided as far as possible.

Twenty Ss were randomly assigned to each of two groups of ten. One subject group heard sentence Group I and the other heard sentence Group II. Ss were run individually; the material was presented binaurally through earphones in a quiet environment.

Ss were all MIT undergraduates paid for their voluntary participation. They were instructed to restate each sentence in their own words immediately following its presentation. Ss were not allowed a rote repetition of the sentence. This was because it was found in preliminary work that Ss could produce rote repetitions of sentences of the type used here even when they professed not to understand the sentences and could not paraphrase them. Clearly, the present hypothesis is relevant only in cases where the material is treated as a sentence, not merely as a list for rote recall. Ss had five successive attempts at each sentence. The responses were tape-recorded.

Two basic scores were computed. The first was the interval between the end of the presentation of the stimulus sentence and the onset of S's vocalization. We will refer to this interval as the "response delay". Second, S's paraphrase was scored to determine his grasp of the subject-object relations among the NP's. The results will be reported here, however, in terms of the grammatical relations scores weighted by the response delay scores. That is, the value reported for each sentence will be the mean number of correct subject-object relations reported per trial, divided by the median response delay. This value may be thought of as the number of grammatical relations correctly recovered per second of response delay (it must be borne in mind, however, that the values were determined from group means). Here and subsequently the pattern of results for both the response measures is substantially the same. Where there is any difference between the combined score and either of the two independent measures, it will be noted.

Results: Table 3 presents the combined score measures for each of eight stimulus sentences. The difference Between Groups I and II is significant for $\alpha = .05$ (Wilcoxon test). All eight sentences showed larger combined scores for the group with relatives present. In fact, the presence of the relatives made the sentences easier to understand whether one takes the response delay or the accuracy of information retrieval as the criterion.

Table 3

Mean number of subject-object relationships recovered per second of response delay [1]

		Group I	Group II			Group I	Group II
1.	The pen	.14	.49	5.	The car	.66	.88
2.	The tiger	.15	.47	6.	The ball	.80	1.00
3.	The boat	.29	.48	8.	The shot	.51	.99
4.	The cigarette	.18	.57	9.	The man	.24	.38

Seven of the eight sentence pairs showed response delay differences in the predicted direction; the one reversal is by far the smallest of the differences. All eight sentences show the greater number of grammatical relations correctly reported for the versions in which relatives were present.

Discussion. There are four important alternative explanations of the difference between Groups I and II which must be considered. First, one might suppose that inadvertent changes in the prosodic features were introduced during the reading of the Group II sentences. That is, one might suppose that is is not the relative pronous per se that produced the facilitation for Group II, but rather that their introduction affected the reading of the stimulus sentences (i.e., a more expressive reading was given for Group II than for Group I). It is clear that the prosodic features are capable of conveying some of the sorts of information we have hypothesized is provided by the presence of relative pronouns. Hence, though in the judgment of the experimenters the readings for both groups in Experiment 1 were comparably expressionless, the possibility of inadvertent prosodic cues must be considered.

---

[1] Sentence 7 is excluded from the analysis here and subsequently. The sentence was badly conceived (as the reader can determine for himself by reference to page 30). Ss were virtually unanimous in their indignant assertions that the sentence was "meaningless" -- even those Ss who eventually succeeded in reporting the correct relationships among the sentence elements. In fact, the sentence is best described as a non sequitur. Exclusion of sentence 7 does not significantly alter the results.

Second, it might be argued that the presence of the relatives somehow provides a cue to the segmentation of the sentence. That is, it could be claimed that the difficulty with these sentences arises not out of any differential availability of the grammatical relationships, but in the preprocessing required to establish segmentation of the utterance. The difficulty with the deleted relatives sentences would thus be attributed to the subject's undertainty about which objects in the stimulus sentences to treat as units in its analysis, rather than to his inability to grasp the grammatical relations which obtain among them.

The third alternative is that, since the relatives are ordinarily redundant (the information they provide is carried positionally and by the prosodic features), their presence thus reduces the rate of presentation of information and gives S more time to retrieve the grammatical facts about the sentence.

The fourth alternative is simply DTC, referred to earlier. That is, it might be assumed that the addition of the relative pronoun deletion transformation to the derivational history of the sentences in Group I makes them more complex than the sentences of Group II.

Four variations of Experiment 1 were run in order to evaluate these alternatives. In order to determine the relevance of prosodic cues to the facilitation effects found in Experiment 1, such cues were intentionally introduced in Experiment 2. In Experiment 3, rate of presentation is made slower for sentences in Group I. In Experiment 4, the relevance of the serial input and fixed rate of information presentation required by the auditory mode is evaluated through a visual analagon of the experiment. In Experiment 5, the length of the derivational history for sentences in Group I is increased by several transformations.

Experiment 2. All the stimulus materials and conditions of presentation of sentences were as in Experiment 1. The sentences were rerecorded, however. Instead of a neutral reading as in Experiment 1, an effort was made to read the sentences expressively. If the prosody is capable of doing the work of the relatives, we should expect to eliminate any differences between Groups I and II not attributable to other factors (as slower rate of presentation).

Results: Table 4 reports the combined scores for the eight sentences. The difference between the two groups is significant for $\alpha = .05$. (If response delay and grammatical relations measures are considered separately, we find both in the predicted direction. The difference between the two groups is significant for response delay but not for the grammatical relations measure.)

Table 4

Mean number of subject-object relationships recovered per second of response delay

	Group I	Group II		Group I	Group II
1. The pen	.31	.78	5. The car	.71	.86
2. The tiger	.36	.39	6. The ball	.70	.80
3. The boat	.51	.48	8. The shot	.68	.76
4. The cigarette	.45	.71	9. The man	.29	.36

If one compares performance in these groups with that of the groups in Experiment 1, it is clear that the prosodic cues improved scores significantly in the relatives deleted group. In spite of this reduction, the differences between the relatives deleted and the relatives present groups remain. Further, Group II (relatives present) <u>without</u> prosodic cues (i.e., Group II in Experiment 1) is significantly superior ($p < .01$) to Group I (relatives deleted) <u>with</u> prosodic cues (i.e., Group I in Experiment 2). Evidently, the prosodic features are helpful, but they cannot account for the difference between Group I and Group II observed in Experiment 1.

<u>Experiment 3</u>. A copy of the tape containing stimulus sentences for Group I in Experiment 1 was made. At the points in the sentence where relatives occur in Group II, a piece of blank tape of appropriate length was spliced into the sentence (mean values for the relatives' durations; 330 ms for the first relative position and 260 ms for the second relative position). This made the durations of the sentences in Group I approximately equal to thos in Group II. All other conditions were as in Experiment 1.

Results: Table 5 presents the values on the combined measure for the eight stimulus sentences. When compared with the values for Group I (see Table 1) we see that there is no significant improvement in the scores produced by the reduced rate of presentation of the sentences.

Table 5

Mean number of subject-object relationships recovered per second of response delay

1. The pen	.21	5. The car	.56
2. The tiger	.20	6. The ball	.55
3. The boat	.30	8. The shot	.59
4. The cigarette	.25	9. The man	.26

Experiment 4. As an additional control on the possible effects of rate of presentation, a visual version of the experiment was performed. The sentences were typed separately on 3 x 5-inch file cards in capital letters. Ss were presented with each card for a period of 3 seconds. The card was then removed and Ss were required to respond as they were in the other experiments. All other conditions were as in Experiment 1.

Results: Table 6 presents the values on the combined score for sentence Groups I and II in visual presentation. Just as in the auditory presentation the version in which the relative pronouns are present is significantly better than the group in which they are absent. Since in this version of the experiment the rate of presentation of information is presumably equivalent for the two groups, the significant differences in Ss' performance evidently cannot be attributed to a segmentation problem peculiar to the auditory mode.

Table 6

Mean number of subject-object relationships

recovered per second of response delay

		Visual Group I	Visual Group II			Visual Group I	Visual Group II
1.	The pen	.35	.48	5.	The car	.81	1.17
2.	The tiger	.25	.29	6.	The ball	1.19	1.13
3.	The boat	.23	.43	8.	The shot	1.00	1.25
4.	The cigarette	.33	.64	9.	The man	.30	.43

Experiment 5. In this experiment, the length of derivational history of the sentences in Group I was greatly increased by the addition of an adjective before each of the first two noun phrases. Once again, sentences were read in a neutral intonation to permit comparison with group I in Experiment 1.

To understand the nature of this control, we must consider the syntactic analysis of adjective-noun sequences. Consider, in particular, the sequence "tired soldier" occurring in the sentence: "The first shot the tired soldier the mosquito bit fired missed." The underlying structure is the configuration (The soldier rel fired the shot), where rel dominates (the soldier is tired) (see Figure 2). The order of operations in the derivation is roughly the following:

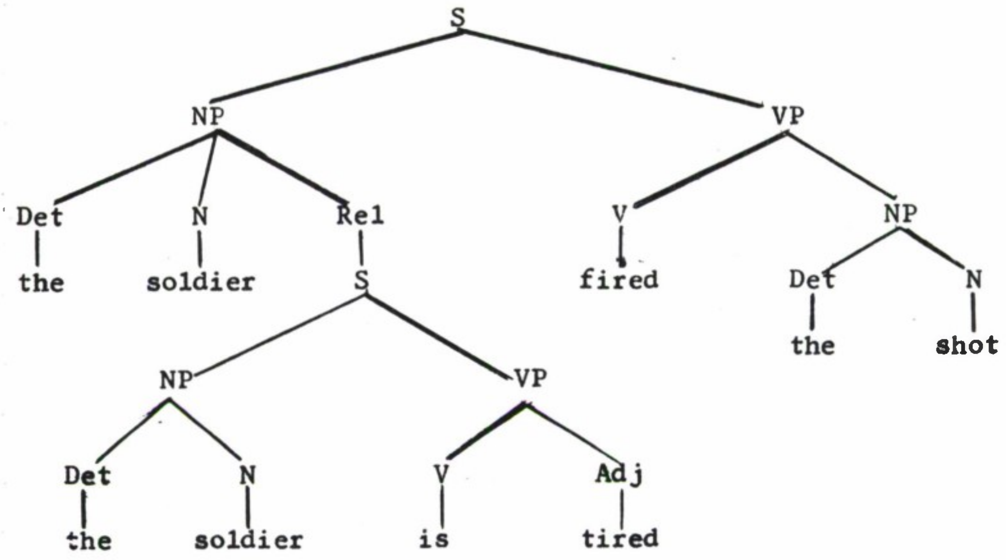


Figure 2. Schematic deep structure tree for the sentence: *the tired soldier fired the shot*.

Relativization applies to produce (the soldier who is tired). *Who is* deletes optionally to yield the intermediate form (the soldier tired) which by a mandatory permutation of the adjective with the noun yields *the tired soldier*.

In short, each of the two added adjectives introduces three transformations, so that each adjective control sentence has six more transformations in its derivational history than the comparable Group I version and eight more than the comparable Group II version. If, then, increase in the length of the derivational history by *two* deletion operations is presumed to account for the difference in complexity between the relative and delected relative version of the stimulus sentences, we would expect that the increase of length by an additional *six* operations ought to make *S*'s performance on the adjective versions far the worst in our study.

Results: The comparison of the adjective and deleted relatives groups on the weighted grammatical relations measure is given in Table 7. Clearly, the addition of the adjectives fails to produce a performance decrement.

Table 7

Mean number of subject-object relationships

recovered per second of response delay

1.	The pen	.42	5.	The car	.62
2.	The tiger	.21	6.	The ball	.82
3.	The boat	.48	8.	The shot	.84
4.	The cigarette	.21	9.	The man	.20

On the contrary, the adjective group is better than the deleted relatives group ($p < .13$). This, of course, is just the reverse of the DTC prediction.

In order to determine whether this result was consistent, we reran the adjective materials with visual presentation. The results were that performance was slightly but not significantly ($p < .20$) worse than for the visual presentation of Group I (cf. Experiment 4). In neither the auditory nor the visual presentations were the differences between Group I values and values for the adjective versions of the stimulus sentences as great as those produced by the presence of the relative pronouns. The mean changes across sentences effected by the presence of relative pronouns was .29 for auditory presentation and .17 for visual presentation. The changes effected by the adjectives in Experiment 5 were an improvement of .10 for auditory presentation and a reduction of .02 for visual presentation.

It is possible that the improved performance for the auditory presentation of the adjective sentences is attributable to additional prosodic cues in that group. As Experiment 2 showed, performance on this task can be significantly improved by such cues. In any event, it seems clear from the two experiments with the adjective versions of the stimulus sentences that the drastic performance decrement predicted by DTC is not forthcoming.

Discussion. A convenient overview of the several experiments can be obtained by relating the experimental groups to each other on each of the stimulus sentences. Using the combined measure (grammatical relations scores weighted by response delays), the scores for each auditory condition can be ranked for each sentence. For example, on Sentence 1 the experimental groups rank as follows (lowest scores to highest):

Sentence Number 1

Group I, No Intonation	Group I, Blank Tape	Group I, With Intonation	Group I, With Adjectives	Group II, No Intonation	Group II, With Intonation

If one does similarly for the remaining sentences and then sums the ranks for each condition, an ordering of the experimental groups is determined as follows:

Exp. Cond:	No into no rels	Blank tape	Aud. Adj.	Into no rels	Into rels	No into rels
Sum of ranks:	13.5	16	26	30	38.5	44

No pair of adjacent groups in the above array are significantly different except where the relatives are introduced. More strikingly there is no significant difference between the groups with relatives, even though the corresponding difference between the deleted relatives groups is significant. That is, though the addition of prosodic cues produces significant increases in scores for the deleted relatives groups, this is not the case for groups where relatives were present. In short, in the presence of the relative pronouns apparently the other facilitating variables have little effect.

In the analysis of the results above, the grammatical relations score used referred only to *S*'s performance on subject-object relations among the NPs. Further analysis reveals facilitation for relations between nouns and verbs as well, e.g., *S*'s recovery of the relationship between NP_1 and VP_3 is facilitated. However, the facilitation for such relations was found to be significant in Experiment 1 but not in Experiment 2. That is, it appears that the facilitation effect of the relative pronouns extended to relations other than subject-object when there were few prosodic cues. When the prosodic cues were provided, however, *S*'s performance on the two groups of sentences is strongly discriminated only along the lines of the heuristic device we have suggested the relative pronoun permits.

Finally, in evaluating the effect of the relatives, sentences 5, 6 and 8 merit special considerations. In these sentences the nouns and verbs were chosen such that certain of their combinations are either grammatically impossible or intuitively absurd. Thus, in sentence 5, analyses like *the car bit* or *the dog drove* are presumably ruled out *a priori*. Schlesinger (1966) has shown that such restrictions enhance subject performance on self-embedded constructions presumably by narrowing the combinatorial possibilities *S* has to consider. In the present data, too, the occurrence of such restrictions clearly facilitates the subject's performance. It should be noticed, however, that even on these sentences the presence of the relatives significantly improves *S*'s performance.

If one assumes that the effect of retaining the relative pronouns is clearly established by these results, it remains to consider what implication these results have for a theory of how the syntactic structure of sentences is understood.

The most important consequence of the present results is, of course, not specific to the perception of self-embedded constructions. Rather, it concerns the question raised above about the relation between the grammatical operations (of transformation, etc.) employed by a syntax in the generation of representations of sentences and the mental operations that must be presumed on the part of the hearer who understands a sentence.

We saw that on theories like DTC this relation is assumed to be very close indeed. One predicts a correlation between the length of the derivational history of a sentence and its perceptual complexity for one of two reasons. Either one assumes that in understanding a sentence the hearer literally runs through the set of operations a grammar employs in generating the sentence or, at least, one assumes that for each grammatical operation there is a corresponding decoding operation.

Clearly, on the analysis by synthesis model or any other model which assumes that the operations of the grammar are isomorphic to some subset of the operations of the recognition device, DTC ought to be true. The present results, therefore, suggest that such models should be viewed with some suspicion. There are two relevant points. On the one hand, the results with the adjectives suggest that there are at least some cases where lengthening of the derivational history of a sentence fails to produce a significant increase in its perceptual complexity. On the other hand, the demonstrable facilitation introduced by the relative pronouns suggests alternatives to grammar-isomorphic recognition procedures.

In short, it appears that the subject has available heuristics for making fairly direct inductions of base structure configurations (i.e., of fundamental grammatical relations) given relevant surface structure information. These heuristics employ information represented _in_ the grammar, but they are not themselves grammatical rules in the usual sense of rules used to generate sentences. The present results suggest some insight into what the nature of these heuristics might be.

We have seen that it is a precondition of facilitation by the relative pronoun that _S_ has and employs an analysis of the types of base structure configurations into which verbs can enter. Initial support for this view derives from analysis of the erroneous responses in the two cases where our sentences contained verbs which can take deep structures other than NP_1 V NP_2; these are "like" (compare sentence 1 with "The pen the editor liked the author to use...") and "know" (compare sentence 9 with "The man my friend knew married the girl..."). In both these cases, but in only these cases, Ss produced errors compatible with deep structure analyses other than NP_1 V NP_2--that is, all other errors preserve the essential subject-verb-object base structure configuration.

It seems likely, on the basis of these results, that the program Ss use to recover the grammatical structure of sentences has at least the following two components; on the one hand, it must consult a lexicon which classifies the verb in the sentence according to the base structure configurations it may enter. Second, it must run through each such deep structure configuration, asking whether the surface material in the sentence can be analyzed as a transformed version of that deep structure. Thus, for example, in analyzing a sentence like "John expected Mary to leave," we assume that the subject must ask which of the deep structure configurations possible for "expect" can take for/to complementation must be facilitated by the presence of the "to" in the surface structure of the sentence.

The action of the surface structure markers in facilitating the recognition of such deep structure features as complement type can sometimes be quite dramatic. For example, that complements are the only ones within which tense can appear in English ("He thinks that John will be late" but "He believes John to be an idiot" and not* "He believes John to will be an idiot"). For this reason, the presence of marked tense is an extremely efficient clue to that complementation in reduced sentences (i.e., in sentences where the word "that" has been deleted). Notice that "He felt the child trembled" is heard as a reduced form of "He felt that the child trembled" and is unrelated to that (tenseless) "He felt the child tremble." Correspondingly, "He felt the children tremble" is ambiguous between the two versions, depending on whether "tremble" is heard as tensed. In these cases, it seems perfectly evident that an effective clue to the analysis of the complement is being given by the presence or absence of tense in the surface structure.

Notice that if these examples are typical of the sentence recognition strategies Ss employ, it seems their search routine need not require frequent analysis by synthesis loops through the grammatical rules. Rather, the types of structures to be examined in any given case would be drastically constrained if only by the subject's information about the deep structure capacities of the verb.

Clearly, this is no more than a sketch of the way syntax recognition might work. On the one hand, there will be cases where applying the surface structure tests for appropriateness of a putative base structure analysis requires the restoration of transformationally deleted material, and it is not at all obvious precisely how this is to be accomplished. Nevertheless, if this sort of account is at all plausible, it is self-evident that there would be no theoretical motivation whatever for theories of complexity analogous to DTC. Since the hypothesized relationship between grammatical rules and perceptual heuristics is extremely abstract, this model provides no basis for a notion of perceptual complexity defined in terms of the former.

Experiments with verb structure

The experiments we have discussed above suggest that the character of the verb in a sentence may play a central role in the determination of the strategies used for decoding that sentence. We have argued that a sentence recognition routine capable of inducing base structures, given formative strings, might consist of at least the following: (1) a component which projects candidate deep structure analyses of an input string by reference to the deep structure configurations which the lexical items in the string are capable of entering, and (2) a component which (when there is more than one such candidate analysis) chooses among candidate analyses by reference to explicit markers in the surface structure of the sentence. Roughly, the information exploited by the first of these processes is thought to drive from the lexical component of a generative grammar, while the information exploited by the second derives from its transformational component.

For example (as in the sentence types used in Experiments 1 through 5), the lexical item <u>whom</u> in (1) is the consequence of a transformation which derives (1) from a deep structure configuration in which <u>the man</u>

(1) the man whom the dog bit died

is object of a verb (<u>bit</u>) of which <u>the dog</u> is subject. The presence of <u>whom</u> in the surface structure of (1) may thus be thought of as a "spelling" of that deep structure configuration. This "spelling" is, however, ambiguous. That is, the presence of the surface structure configuration (2) does not uniquely determine the base structure configuration (3). Rather,

(2) NP_1 <u>whom</u> NP_2

the inference from surface structures like (2) to deep structures like (3)

(3) NP_2=subject Verb NP_1=object

requires at least information about the grammatical character of the verbs to which the NPs are related. (Notice that sentences like (4) exhibit the surface configuration (2), but not the base structure (3).

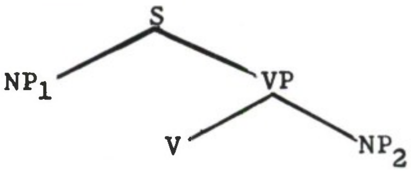
(4) the man whom the girl knows John likes got ill

Experiments 1 through 5 showed that the introduction of relative pronouns in the NP sequence of doubly embedded center-branching sentences facilitates their comprehension. It was argued that subjects' ability to exploit thepresence of the pronoun must depend on their application of a lexical analysis of the verb. In particular, *S*s must take account of the transitivity of the verbs in the sentence, since the inference from structures like (2) to structures like (3) is warranted only when the relation between NPs is mediated by transitive verbs.

Implicit in this analysis is the suggestion that the deep structure properties of the verb plays a major role in *S*'s determination of the base structure of an input sequence. It is this suggestion that leads to the present investigation.

Verbs may be classified in terms of the types of deep structure configurations they dominate. For Example, to say of the verb *V* that it is transitive is just to say that in the deep structure it accepts the configuration (5). To say that it is a pure *transitive* is to say that it accepts *only* deep structure configurations like (5).

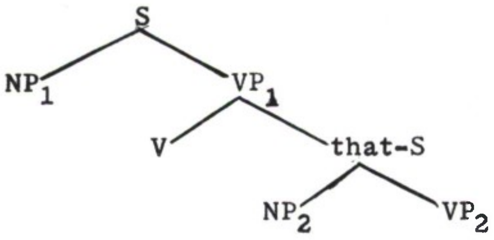
(5)



(John discussed the book)

Analogously, verbs may permit various sorts of complements (see Rosenbaum, 1967). That is to say, they may enter into various configurations of matrix and constituent sentences in the deep structure of a sentence. Thus, a verb like *remark* permits such deep stractures as (6), but not configurations like (5).

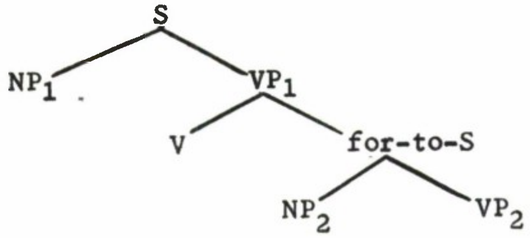
(6)



(John remarked that Mary came)

On the other hand, believe permits *either* configurations like (5) (e.g., John believed Mary), or like (6) (John believed that Mary came), or like (7).

(7)



(John believed Mary to be an idiot)

The proposal for a sentence recognition device of the sort discussed above, if correct, would suggest that these sorts of lexical properties of verbs are exploited in the assignment of base structure. Thus, given a sentence containing the main verb *discuss*, the subject can instantly determine the gross characteristics of its deep tree, i.e., it must have the general form of (5) since *discuss* is a pure transitive. On the other hand, given a sentence with the main verb *believe*, *S* may hypothesize *either* base configurations like (5), or like (6) or like (7) since the lexical structure of *believe* is compatible with all three. Presumably, given an unambiguous sentence in which *believe* is the main verb, *S* must decide between these various possible deep structures by reference to such explicit surface structure features as the occurrence of *that* in (6), of *to* in (7), of tense in the subordinate clause of (8) (cf. Fodor and Garrett, 1967), etc.

(8) John implied Mary came

There is thus an asymmetry between the lexical analysis of *discuss* and similar verbs and the lexical analysis of *believe* and similar verbs. *Certeris paribus,* verbs of the latter kind are compatible with a wider range of hypotheses about the deep structure of the sentences in which they appear than are verbs of the latter kind. This, in turn, suggests that sentences containing verbs like *believe* ought, *ceteris paribus*, to be more difficult for subjects than sentences containing verbs like *discuss*. This is a consequence of the hypothesis that the lexical character of the verb is an important determinant of the number of hypotheses about the base structure of a sentence which the subject must entertain. The following experiments test this prediction.

We have hypothesized that the heuristics employed in the recognition of sentences exploit information concerning the lexical structure of the sentence's verbs: In general, the greater the variety of deep structure configurations the lexicon associates with the main verb of a sentence, the more complicated the sentence should be. The ease with which a sentence is understood should, therefore, be in part a function of this variable.

Two experimental techniques were employed in order to test this hypothesis. We employed both the technique described for Experiments 1-5 and an anagram tast (cf. Marshall, 1964). In the first case, *Ss*' performance in paraphrasing self-embedded sentences is evaluated for both visual and auditory presentations. In the second, *Ss* are required to reconstruct sentences from scrambled word strings. The manipulated variable in each case is the presence of verbs which take a variety of complement types vs. pure transitive verbs. The experiment in which self- embedded sentences were paraphrased is reported first.

Experiment 6. Twelve pairs of sentences with two levels of self-embedding were constructed. The members of each pair differed only in that one contained a verb which which permits complement structures where the other contained a transitive verb.[2] The sentences are listed below with the complement verb underlined in each case.

(a) *knew*
1. The box the man the child (b) met carried was empty.

(a) mailed
2. The letter the secretary the manager employed (b) *expected* was late.

(a) *saw*
3. The actors the writer the agent sent (b) used are talented.

(a) followed
4. The deer the man the boy watched (b) *heard* were timid.

(a) *believes*
5. The plan the lawyer the client (b) hired proposed was impractical.

(a) cut
6. The material the tailor the designer used (b) *felt* was green.

(a) *prefers*
7. The tiger the natives the hunter (b) paid hated was fierce.

(a) evaded
8. The planes the sailors the enemy attacked (b) *feared* were bombers.

---

[2] We have designated the classes of verbs contrasted in this experiment and in experiments 7, 8 and 9 as "complement" verbs and "transitive" verbs. This classification refers to *object complementation*. There are other types of complementation which we have not discussed here (see Rosenbaum, 1967). Though the linguistic analysis is not clear in the case of every verb we have used, the verbs we refer to as complements are compatible with a wider range of structures than those referred to as transitives even when other complement types are taken into account.

(a) *suggested*
9. The tactics the general the soldiers admired (b) adopted
were stupid.

(a) discuss
10. The events the papers the man bought (b) *reported* are
unsettling.

(a) *intended*
11. The insult the waiter the lady summoned (b) provoked was
obvious.

(a) ignored
12. The results the scientist the committee appointed (b) *predicted*
are surprising.

The experimental groups, a and b, were derived from the list such that in each group half the sentences contain complement verbs and half do not. Under these conditions, the performance levels for each subject on the two types of sentences can be compared, as can the performance across *S*s on the two versions of each sentence.

Procedure. Each sentence was each typed onto a 3 x 5 file card in capital letters. *S*s were presented with a card for a three-second period. The card was then removed and *S* was required to restate the sentence in his own words as soon as he was able to do so. *S*s were told that we would measure both their accuracy and the time it took for them to do the task. *S*s were *not* permitted exact repetition, since it was found in earlier work that a rote repetition of sentences of this type and length was quite possible for *S*s even when they did not understand the sentence. *S*s had five successive attempts at each of the twelve sentences that were presented. Responses were tape-recorded.

Subjects: All subjects were M.I.T. undergraduates who were paid for their voluntary participation in the experiment. Twenty *S*s were randomly assigned to one of two groups. The groups were presented, respectively, with sentence Group A or sentence Group B. Subjects were run individually in sessions which lasted approximately twenty minutes.

Scoring. Two scores were computed: an accuracy of paraphrase score and a response delay score (i.e., the interval between the end of presentation of the sentence and the onset of *S*'s response). The accuracy of paraphrase score was determined from the number of Subject-verb- {object or / predicate / adjective} triples that were correctly reported.
triples that were correctly reported. For example, if *S* were shown the sentence:

The box the man the child knew carried was empty

and produced the paraphrase:

The child carried the box the man knew was empty

his score would be 1 for his correct recovery of the triple The box was empty. The S's paraphrase in this case would fail to correctly represent the triples the man carried the box and the child knew the man. Since each of the sentences presented to Ss had three propositions which could be recovered, the maximum score that could be achieved by an S for five presentations of a sentence was 15.

Results. Table 8 presents the accuracy of paraphrase scores for each of the sentences in its two versions. (Each sentence is referred to by a key word.) Scores overall were better in Group B than in Group A. Hence, for comparison of the two sentence versions, scores in Group B were transferred to the same mean as Group A. As Table 8 shows, in eleven of the twelve comparisons subjects produced more accurate paraphrases for the sentences which did not contain a complement verb. This pattern is significant for $\alpha = .05$ ($p < .005$, Wilcoxon test for matched pairs).

Table 8

Mean number of subject-verb-object triples correctly recovered

for each sentence in visual presentation

Sentence number	Complement version	Non-complement version	Sentence number	Complement version	Non-complement version
1. box	9.6	11.0	7. tiger	8.7	9.6
2. letter	8.1	11.8	8. planes	5.3	5.8
3. actor	5.9	6.5	9. tactics	9.1	12.0
4. deer	8.0	9.6	10. events	13.7	12.8
5. plan	8.8	11.1	11. insult	9.5	9.8
6. material	4.6	7.5	12. results	10.2	10.8

If the analysis is done by subject, we find that nineteen of the twenty subjects showed a difference between their scores on complement verb sentences and their scores on non-complement verb sentences. Fourteen of these subjects showed higher scores on the non-complement verb sentences, and five showed higher scores on the complement verb sentences (the six cases which do not support the hypothesis were distributed three in Group A and three in Group B). This pattern for the subject analysis is significant for $\alpha = .05$ ($p < .005$, Wilcoxon test).

When the response delay scores are considered, however, we find a reversal of the expected pattern; eight sentences showed greater response delays for the complement version than for the transitive version. Though the difference is not significant ($p = .61$), it is nonetheless an unanticipated trend, especially in light of the correspondence between the response delay measure and the paraphrase measure in an earlier study with a similar task.(Experiments 1 through 5).

Discussion. The experiment with visual presentation of the stimulus sentences seems to support the hypothesis that verb complexity significantly affects *S*'s performance on sentence comprehension. However, the results are unclear in several respects. First, there is the matter of response delay reversal mentioned above. Second, there are difficulties with the stimulus sentences. For example, sentences 4 and 5 have complement verbs in *both* versions; in the version with the greater predicted complexity, however, there are *two* such verbs, in the putatively simpler version, only one. Other, less serious difficulties were also discovered, such as the tendency of some subjects to confound *felt* the verb with *felt* the noun in sentence 6. Finally, there is the matter of ambiguity of the stimulus sentences. It is the unfortunate concomitant of the introduction of the verbs which take complements that one also creates the possibility of full and/or partial syntactic ambiguities in certain of the sentences. An examination of the stimulus sentences listed above will illustrate these difficulties. Thus, in the complement verb version of sentence 5, there is the possibility of construing the sequence *believed proposed* as a compound verb, on the model of "believes to be proposed." Though the sentence taken as a whole precludes this interpretation, such "local ambiguities" may, nevertheless, provide a source of confusion for *Ss*, thereby biasing the experiment against the complement verb versions of the stimulus material. The same difficulty can arise when the complement verb occurs one position later in the sentence. Thus, sentence 6 invites the interpretation "felt to be green" for the sequence *felt was green*.

A still more serious difficulty is that certain of the sentences are *fully* ambiguous in their complement version. An example is sentence 9, for which the intended reading is "the general suggested stupid tactics" but which permits the reading "the general suggested that the tactics were stupid." (The reader may have to read the sentence aloud to convince himself of the possibility of this interpretation. The following punctuation suggests the ambiguity: the tactics, the general the soldiers admired suggested, were stupid). Of the twelve stimulus sentences, sentences 3, 4, 6, 8, 9, and 10 permit this sort of interpretation in their complement versions. In the remaining sentences it is precluded either by tense restrictions or by the lexical items themselves.

We have already noted that of the twelve sentence comparisons eleven showed greater difficulty associated with the complement verb, although only six of the sentences are potentially ambiguous. Moreover, if the ambiguity were primarily responsible for the performance difference in the experiment, one would expect a weaker effect among the unambiguous sentences. In fact, the one reversal (i.e., poorer performance on the *non*-complement version) was an ambiguous sentence, and overall, the magnitude of the complement verb effect was slightly *greater* among the *unambiguous* sentences.

Experiment 7. In order to cope with some of the difficulties in Experiment 6, an auditory version of the study was performed. In this experiment, the stimulus sentences used in Experiment 6 were revised as shown in the list below.

(a) *knew*
1. The box the man the child (b) met carried was empty.

(a) mailed
2. The letter the secretary the manager employed (b) *expected* was late.

(a) *saw*
3. The actors the writer the agent sent (b) used were talented.

(a) fed
4. The deer the man the boy followed (b) *heard* were timid.

(a) *proposed*
5. The plan the lawyer the client interviewed (b) devised was impractical.

(a) cut
6. The material the tailor the designer used (b) *required* was green.

(a) *preferred*
7. The tiger the natives the hunter (b) paid hated was fierce.

(a) evaded
8. The planes the sailors the enemy attacked (b) *feared* were bombers.

(a) *suggested*
9. The tactics the general the soldiers admired (b) adopted were stupid.

(a) discussed
10. The events the papers the man bought (b) *reported* were unsettling.

(a) intended
11. The insult the waiter the lady summoned (b) provoked
was obvious.

12. The results the scientist the committee appointed
(a) ignored
(b) predicted were surprising.

In this revised list no sentence has more than one complement verb, and various tense adjustments have been made in an attempt to reduce partial ambiguities.

The sentences were tape-recorded. They were read with full intonation. Also, emphatic stress and pause were used to indicate the intended sentence structure. It was intended thereby to preclude the irrelevant interpretation of fully or partially ambiguous sentences insofar as possible. (The "normal" intonation pattern is quite different for the two versions of the ambiguous sentences, whether the ambiguity is full or partial.)

The sentences were presented to 30 Ss randomly assigned to Groups A and B. The Ss were run individually with the stimulus material reproduced over loudspeakers. Responses were tape-recorded.

Results: Table 9 reports the paraphrase scores for the twelve sentence pairs. Each sentence is referred to by a key word, as in Table 8.

Table 9

Mean number of subject-verb-object triples correctly recovered

per sentence for auditory presentations

Sentence number	Complement version	Non-complement version	Sentence number	Complement version	Non-complement version
1. box	10.9	11.3	7. tiger	9.9	11.3
2. letter	10.4	11.1	8. planes	7.6	8.4
3. actor	6.5	6.7	9. tactics	13.1	12.8
4. deer	11.3	10.9	10. events	11.8	13.8
5. plan	9.3	10.9	11. insult	11.3	12.1
6. material	9.4	10.1	12. results	11.9	12.1

As in Experiment 6, performance overall is better in Group B than in Group A. In order to permit a sentence-by-sentence comparison, the scores of Ss in Group B were transformed to the same mean as those in Group A. As Table 9 shows, for the sentence-by-sentence comparison, the versions containing a complement verb have a lower score on the paraphrase measure than do their counterparts without complement verbs in 10 out of 12 instances ($p. < .005$ Wilcoxon, one tail). There are five confirming and one disconfirming comparisons for each group.

When we consider the response delays for the auditory experiment, we find that (unlike the visual experiment) there is a non-significant tendency for the complement versions of the sentences to show relatively long response delays compared to their non-complement counterparts; 7 of 12 instances exhibit this pattern ($p = .38$).

Analysis of the results for individual subjects reveals the same patterns. Of the 29 Ss who showed a difference between their scores for the complement and non-complement versions 20 showed larger scores for the non-complement version on the paraphrase measure ($p < .005$, Wilcoxon, one tail). Response delay comparisons, however, showed no significant differences.

Discussion. Presentations of the stimulus material both visually and auditorily showed a performance decrement (on paraphrase scores) for those sentence versions containing the more complex verbs. This is true whether one looks at sentence-by-sentence contrasts across subjects' or at individual subject's performances on the two types of sentences. The pattern of results is less clear when one considers the response delay measure. Though response delay and paraphrase measures were highly correlated in our earlier work with similar experimental tasks, this was not true in the current experiment. In the visual presentation the trend was for slightly greater response delays for non-complement verbs; in the auditory presentation the trend was for slightly greater response delays for complement verbs. In neither instance were the differences significant. The best that one can say of the response measure in the current experiment is that its effectiveness may have been reduced by the greater tendency of subjects here to hesitate and flounder during their paraphrasing. Though we have made no quantitative assessment of the change, it is our impression that Ss were much less fluent than in the former experiments. Curiously, however, overall the response delays are lower in this experiment than in Fodor and Garrett (1967). Both these facts become reasonable if one assumes that Ss perceived the sentences to be easier than they found them to be. That is, Ss began to speak prematurely (thus decreasing response delays) and then discovered the difficulties afterwards (thus the increased nonfluency of Ss for this experiment).

Setting aside the nonsignificant response delay results, the paraphrase results appear to demonstrate clearly that the character of the verb structure is an important consideration in determining the perceptual complexity of sentences.

We earlier discussed the problem of ambiguity introduced by the use of complement verbs. The auditory version of the experiment was largely intended to cope with this problem. We argued above that even in the visual version ambiguity cannot fully account for our results. In the auditory version the effects of the prosodic cues also militate against the ambiguities. Further, as in experiment 6 we can reject this possible explanation of the difficulty of complement verb sentences on the grounds that (1) the unambiguous cases of complement verb sentences all show the predicted asymmetry, (2) the only two reversals of the expected effect occur among the ambiguous versions, and (3) the magnitude of the performance decrement is approximately the same for both the amiguous cases and the unambiguous cases (in fact, the mean effect for the unambiguous versions is slightly greater).

Further analysis of the paraphrase responses provides additional support for the view that it is the complex structure of the complement verbs that produces difficulty and not the potential ambiguities. Ss frequently reported only part of the stimulus sentences in their paraphrases--that is, portions of the input were omitted in their response. If one compares the incidence of verb deletions for complement and non-complement verbs in the same environments, it is found that Ss were much more likely to omit a complement verb than a non-complement verb. In the auditory presentation, for instance, complement verbs are dropped 266 times while non-complements are dropped 167 times. By sentence, the difference is significant; eight sentences show more instances of dropping complement verbs, three show more dropping of non-complements and one, no difference ($p < .05$, Wilcoxon, two tails).

There is some additional evidence which indicates the central importance of the verb in the analysis of self-embedded sentences. In a study by Bever (Harvard Center for Cognitive Studies Report, 1967) it was found that for doubly self-embedded sentences the paraphrase scores are higher for sentences with polysyllabic verbs than for their counterparts containing monosyllabic verbs. A similar variation in the length of nouns in such sentences, however, did not affect the accuracy of paraphrase. If it is assumed that the increase in length of a word provides an increase in the time available for computation of the sentence structure, Bever's findings indicate more computational activity while the verbs are being analyzed since an increase in computational time helps only when at the locus of the verbs.

Experiment 8. While the results of the experiments with self-embedded sentences appear to offer persuasive evidence for the role of verb structure in determining sentential complexity (and hence for the role of the main verb in determining the candidate analyses S considers as possible deep structures for a sentence), it seemed desirable to test this view of structure assignment with a different experimental paradigm. In particular, we wished to use stimulus materials which avoided the problems with ambiguity raised in the case of the self-embedded complement structures. Further, we desired that the stimulus materials be more "natural" examples of sentences than those used in experiments 6 and 7.

An experiment was, therefore, performed in which Ss" task was the construction of a sentence from a scrambled set of words. The use of an anagram task to evaluate the complexity of sentences is not novel. John Marshall (1964) found this sort of task yielded the same complexity ordering for optional singulary transformations as found by Miller and McKean (1964) and others.

Stimulus Materials. The stimulus sentences are given in the list below. Each sentence occurs in two versions differing only in that one version has a main verb which takes both complement structures and direct objects while the other version has a pure transitive.

1. The man whom the child (a) knew (b) met carried a box.

2. The girl in the movie really did (a) kick (b) like the salesman.

3. The letter which the secretary (a) expected (b) mailed was late.

4. Although he was very sick the dictator (a) disobeyed (b) resented the doctor's advice.

5. The actor whom the agent (a) saw (b) sent was talented.

6. The tired movie star (a) indulged (b) granted his public's requests for autographs.

7. The boy whom the man (a) found (b) followed was very ill.

8. He had good reasons not to (a) greet (b) acknowledge his old friends.

9. The ambassador (a) ordered (b) borrowed ten cases of brandy from his nephew.

(a) hired
10. The lawyer whom the client (b) believed was honest.

11. Congress quickly passed the controversial bill which the
(a) requested
president (b) drafted.

(a) needed
12. The tailors whom the designer (b) required belonged to
a union.

(a) decided
13. The judge (b) biased the case in favor of the corrupt
politician.

(a) paid
14. The natives whom the hunter (b) prefers are hard workers.

(a) reported
15. The janitor (b) hit the tenant who complained about the
high rent.

(a) relayed
16. The commander (b) announced the news of the armistice to
his troops.

(a) feared
17. The planes which the enemy (b) evaded were bombers.

(a) adopted
18. The tactics which the general (b) suggested were stupid.

(a) doubted
19. The chief of police (b) ignored the story of the watchman's brother.

(a) blamed
20. The manager (b) recommended John when the company ran short
of help.

(a) considered
21. The committee (b) deleted all the arguments for abolishing
private property.

(a) discussed
22. The events which the papers (b) reported are unsettling.

(a) revealed
23. The book was badly written but it (b) contained some
important facts.

(a) provoked
24. The insult which the waiter (b) *intended* was obvious.

(a) *demanded*
25. The condemened prisoner (b) obtained pardon from the
governor's office.

(a) hidden
26. The accomplice had (b) *warned* the murderer the moment the
police arrived.

(a) *permitted*
27. The government (b) ended the shipment of medical supplies
to the guerrillas.

(a) ignored
28. The results which the scientist (b) *predicted* are surprising.

(a) *required*
29. The old theory obviously (b) contained several false
assumptions about cosmology.

(a) killed
30. The villagers (b) *remembered* the evil witchdoctor from
the next district.

It will be noted that as in Experiments 1 and 2, none of the sentences actually contain a complement construction: i.e., whether or not the main verb is grammatically *capable* of accepting a complement, its role in the stimulus sentence is that of dominating a direct object.

In addition to the stimulus sentences listed, *S*s also received fifteen "padding" sentences. The padding sentences, which were distributed among the stimulus sentences, all contain a complement verb which, in fact, dominates a complement structure. The list of stimulus sentences and padding sentences received by a subject, therefore, was balanced in the following way: 15 sentences with pure transitive verbs and direct objects; 15 sentences with complement verbs and direct objects; 15 sentences with complement verbs and complement constructions. As in Experiments 1 and 2, therefore, both an analysis by sentence (the two versions of each test sentence compared) and an analysis by subject (the comparison of *S*'s performance on complement verb sentences with his performance on non-complement verb sentences) is possible.

Procedure. The stimuli presented to *S* consisted of individual words typed in capitals on small pieces of file card. The material was presented by placing the gragments haphazardly before the subject. *S* was requested to arrange the fragments serially so that they formed a grammatical and meaningful sentence. S was also instructed to perform the task as quickly as possible consonant with accuracy. In order to motivate his best possible performance, S was given a schedule of small cash incentives for rapid performance. No subject was allowed more than 60 seconds to complete a sentence. Thirty-two subjects were run individually and all were paid a minimum fee ($1.40) regardless of the level of their performance.

Subjects were timed with a stopwatch starting from the presentation of the stimulus materials. The watch was stopped when the subject announced his completion of the presented sentence.

Scoring. The subjects' responses fall into three categories: (1) instances in which a sentence was correctly constructed within the allotted 60 seconds; (2) instances in which S announced his completion within 60 seconds but had in fact failed to produce an acceptable sentence; (3) instances in which S failed to produce an acceptable sentence within 60 seconds. In the cases where S failed to produce an acceptable sentence, a record was kept of the sequence he produced. The data was thus analyzable in terms of number of correct completions, number of erroneously reported completions and response times for acceptable completions.

Results. If we consider only the response times for correct solutions to the sentence construction task, we find no difference between complement and non-complement versions of the sentences. There are thirty test sentences; median scores for each version of a test sentence were computed using only the values from S who constructed the sentence correctly. Of the thirty sentences, sixteen had larger medians for the complement version while fourteen had larger medians for the non complement versions.

However, if we consider those instances in which Ss either failed to complete the task or produced an incorrect sentence, we find a marked difference between the complement and non-complement versions of the sentences. Table 10 gives the sentence-by-sentence comparisons for false reports (S announces his solution to be a sentence when it is not) and failures to complete.

Of the thirty test sentences, thirteen showed more frequent failures to complete in their complement versions, four showed more in their non-complement versions and thirteen shoed no difference ($p < .05$, Wilcoxon test). The analysis by sentence for false reports shows the same pattern; fourteen sentences show more such errors in their complement versions than in their non-complement versions, eight show more for the non-complement version and eight show no difference ($p < .05$, Wilcoxon test). If we consider the sentence-by-sentence results for both measures combined, we find that on seventeen of the sentences there were more errors for complement than for non-complement versions and five sentences for which the reverse was true ($p < .005$, Wilcoxon test).

Table 10

Numbers of false reports and failures to complete for each stimulus sentence in both complement and non-complement versions

	Complement versions			Non-complement versions		
Sentence number	Failure to complete	False report	Combined score	Failure to complete	False report	Combined score
1.	3	5	8	2	5	7
2.	2	2	4	0	2	2
3.	0	1	1	0	1	1
4.	3	0	3	0	2	2
5.	0	7	7	3	5	8
6.	3	5	8	4	2	6
7.	1	2	3	0	3	3
8.	3	3	6	2	0	2
9.	0	2	2	0	1	1
10.	2	4	5	2	2	4
11.	2	1	3	0	2	2
12.	7	5	12	7	3	10
13.	3	2	5	5	2	7
14.	4	5	9	0	1	1
15.	2	0	2	1	1	2
16.	0	0	0	0	1	1
17.	1	5	6	1	1	2
18.	2	1	3	0	0	0
19.	2	3	5	2	3	5
20.	5	2	7	0	1	1
21.	1	1	2	1	2	3
22.	1	0	1	0	1	1
23.	1	2	3	1	0	1
24.	3	1	4	1	3	4
25.	0	4	4	0	0	0
26.	2	0	2	2	0	2
27.	0	2	2	2	1	3
28.	0	1	1	0	0	0
29.	2	1	3	1	1	2
30.	0	1	1	0	1	1

The greater difficulty of the complement versions of the sentences is also borne out when the performance of individual subjects is considered. Of thirty subjects, eighteeen made more errors on the complement sentences they received than on the non-complement sentences; six subjects showed the reverse effect and six performed equally well on both types of sentence.

Discussion. Subjects' performance in the solution of the anagram task seems strongly dependent on their correct assumption concerning the relationship of the verb to the rest of the sentence. Particularly convincing on this point are those instances in which *S*s misperceived the results of their "word shuffling" as sentences when in fact they were not. This misperception happened much more frequently for sentences involving complement verbs than for those not involving complement verbs. Moreover, an analysis of the *kinds* of errors that were made when *S*s incorrectly reported a completion reveals a marked difference between the complement verb cases and the non-complement verb cases. For the cases in which *S*s produced an incorrect solution (which he labelled correct) for a *non-complement* version of a sentence, the error was almost invariably one of getting an adjective or an article out of order (e.g., *the man old ran away* for *the old man ran away*, or *the commander relayed the the news of armistice to his troops*, etc.). When *S*s incorrectly reported completion for complement verb versions of sentences, however, it was frequently the case that there was a serious structural error (e.g., *the waiter intended which was obvious the insult* for *the insult which the waiter intended was obvious* or *a doctors treatment for minor cuts consider iodine to be poor now* for *doctors now consider iodine to be a poor treatment for minor cuts*, etc.).

These results appear to be compatible with the following analysis: when the subject isolates the main verb he makes a guess about the structure of the sentence in which it appears. The options available to him are governed by the lexical character of the verb (i.e., transitive, complement, or mixed). When he guesses correctly, no interference measurable by this paradigm is produced by the existence of irrelevant options: he performs as well with mixed verbs as with pure transitives. When, however, his guess is incorrect, various consequences of interference are evident; among these are failure to complete the task and misperception of the structure of the stimulus.

While it is impossible to prove that the anagram task illuminates specifically *perceptual* processes in sentence analysis, it does appear to illustrate the centrality of the lexical character of the main verb in the integration of linguistic objects. The anagram results thus appear to support the conclusions derived from Experiments 6 and 7.

Experiment 9. A more direct test of the perceptual consequences of the presence of complement verbs was devised using the stimulus materials like those of Experiment 8 (the anagram task). In this experiment subjects were required to recognize sentences under noise. Eighteen pairs of test sentences were prepared; the members of the pairs were identical to each other except for the substitution of a complement verb in one for the pure transitive appearing in the other. The list of test pairs is given below with the complement verb underlined.

1. The beautiful young farm girl in the movie really did
(a) *like*
(b) kick the unscrupulous traveling salesman.

(b) *resented*
2. Though on his deathbed, the dictator (a) disobeyed the
advice of his wife and his doctors.

3. Though tired from his performance, the movie star still
(a) *granted*
(b) indulged his public's requests for autographs.

4. When the president finally applied the necessary pressure
(b) *wrote*
the senate quickly (a) deleted the controversial amendment.

(a) *acknowledged*
5. There must have been many reasons why he never (b) greeted
his old friend from home.

(b) *ordered*
6. The charming but unpredictable ambassador (a) borrowed
a case of cherry brandy from the nearby consulate.

(a) *determined*
7. The corrupt judge (b) prejudiced the case in favor of the
biggest bribe.

(b) *helped*
8. The rich landlord (a) hit the tenant who complained about
the bad harvest and high rent.

(a) *reported*
9. As soon as the Armistice was signed, the general (b) relayed
the news to his troops.

10. The extremely clever chief of security immediately
(b) *doubted*
(a) detained the Minister of the Interior's younger brother.

(a) recommended
11. The personnel manager very quickly (b) blamed Bill when he

received an inquiry from his boss.

12. Though composed of wealthy business men, the committee
(b) admitted
(a) reviewed the arguments for abolishing private property.

13. Though it was badly written, the book on rhetoric
(a) revealed
(b) presented important facts about the nature of language.

14. Seconds before the hanging was to begin, the heroine
(b) demanded
(a) obtained a pardon from the king.
(a) warned
15. The accomplice had (b) called the murderer the moment the

police were brought into the case.
(b) permitted
16. The neutral government (a) ended the shipment of necessary medical supplies to the guerrillas last year.

17. Although it did account for all the experimental facts,
(a) required
the theory (b) contained several false assumptions.
(b) hated
18. The villagers (a) killed the maneating tiger who had been

roaming the district for several months.

Two stimulus lists were constructed (designated list a and list b); in each list half of the sentences contained a complement verb and half a pure transitive. Complement and transitive sentence types alternated in the lists. To each of the lists a set of nine "padding" sentences was added (every third sentence in the list was a padding sentence). These sentences all contained complement verbs with complement constructions. Hence, the stimulus lists were balanced in the following way: nine sentences with a complement verb and a direct object, nine sentences with pure transitive and a direct object, and nine sentences with a complement verb and a complement construction. Just as was the case with the anagram task, this balance was intended to preclude the formation of a set for the direct object construction.

The lists were recorded with normal intonation and presented to groups of subjects against a background of "speech noise." The signal to noise ratio was established such that about half of the lexical items were recovered from a set of warm-up sentences (S/N ≅ 0 db). Immediately after each sentence was presented, the subjects were required to write down whatever portions of the sentence they were able to produce.

Thirty-six subjects were assigned to two groups; one group of eighteen subjects heard list A, the other eighteen Ss heard list B. Hence, whereever Ss listening to list A received a sentence with a transitive verb, those listening to list B received the same sentence but with a complement verb, and conversely.

The score computed was the number of words correctly reported from the part of the sentence which followed the verb.

Results. If one simply looks at the number of lexical items correctly reported where complement verbs were used compared with the number reported for the same sentences where transitive verbs were used, we find that complement verb cases were harder overall; fewer lexical items were correctly reported for the sentences in their complement versions ($p = < .05$, binomial test, one tail).

If we look at the comparisons sentence by sentence, however, some interesting facts emerge. There was considerable variation among the sentences, both in their intelligibility and in the degree to which the expected asymmetry produced by the character of the verb was manifest. Considered pairwise, eleven of the eighteen sentence pairs showed larger numbers of lexical items recovered for transitive versions than for complement versions; seven sentences showed the reverse relationship. Although the differences where they were not in the predicted direction were generally small, there were some which were of the same magnitude as the positive cases. All of these instances, however, were also distinguished as cases in which there was a marked unbalance in the frequency of occurrence of the contrasted verbs. Although there was an attempt on our part to equate the members of each pair for their frequencies of occurrence in English, there were some instances in which we failed to do this. We can divide the verb pairs roughly into two groups; those which are within about two thousand words of each other on Thorndike-Lorge word count, and those that were much further apart. The ratio of frequencies of occurrence of the two verbs was for the first group about 2/1 at worst, while in the second group the ratios ranged from 6/1 to 100/1. There were five such cases of a gross imbalance in word frequency. In every case it was the complement verb which was the most frequent and all five cases were also cases in which the effect of the verb on perception was not in the expected direction. This strongly suggests that the imbalance in word frequency reduced the effect of the complement-transitive verb contrast. Overall, of course, such frequency differences cannot be appealed to in order to explain the reduced performance levels associated with the complement verbs. In addition to the five cases of gross imbalance noted above (where the effect of frequency runs counter to the prediction), in the remaining thirteen sentences, seven have higher frequencies for the complement verb (hence counter to the prediction), two have equivalent frequencies, and four have higher frequencies for the transitive verb. The precise nature of the interaction between word frequencies and the effect of the lexical structure of the verbs is not obvious. Further studies are at present being carried out to explore the nature of this interaction.

Whatever the effect of word frequency, it is clear that the nature of the verb had a significant effect on the ability of Ss to understand the sentences presented under noise. The results here, then, are compatible with those of Experiments 6, 7 and 8. The lexical character of the verb appears to be a significant determinant of the perceptual complexity of sentences.

General Discussion

The results reported here appear to provide support for the view advanced in Fodor & Garrett (1967). It appears that the exploration of the lexical analysis of the main verb of a sentence is a central heuristic in the strategy Ss use to recover its deep structure. This view in turn has a rather direct implication for further research both in linguistics and in psychology. We have presupposed as input to the sentence recognition process we outlined above a representation of the sentence which marks at least a crude segmentation, including the identification of the main verb. That is, for a syntax recognition device to employ the lexical structure of the main verb as a clue to the possible geometry of the deep structure tree underlying an input, it must at least have abailable some hypotheses about what the main verb of the input string is and about what substretches of the string constitute segments of the sentence which the main verb may dominate in deep structure. Hence, the postulation of some preanalysis of the sentence which marks putative relations between verb and noun phrases is a plausible hypothesis on the current view.

There are three sources of evidence pointing to the existence of this sort of preanalysis. First, there is the difficulty of center-embedded constructions which probably is contributed at least in part by the difficulty of determining which verb phrases are related to which noun phrases. That is, in such sentences there is no direct correspondence between surface structure clause adjacencies and deep structure relations (cf. Fodor & Garrett, 1967; Fodor, Bever and Garrett, 1967).

Second, there appears to be linguistic evidence for the existence, at least in English, of certain grammatical asymmetries between thepart of the sentence on the left hand side of the main verb and the part to its right. As Mr. R. Kirk has pointed out to us, in English the deletability of certain lexical items appears to be constrained by their position relative to the main verb of the sentence. For example, in the sentence it is obvious (that) John was bored, the presence of the lexical item that is optional. However, in the sentence that John was bored was obvious, the lexical item that may not be deleted; *John was bored was obvious. It is not implausible to assume that the point of such restrictions is to help the speaker distinguish between the case when the first verb in a sentence is its main verb and the case when it is merely the main verb of an embedded clause. This is by no means the only example of its kind (cf. Fodor, Bever and Garrett, forthcoming).

Third, there exists a certain amount of experimental evidence for the view that it is not primarily the immediate constituent but rather the clause that provides the perceptual unit in speech (see section II of this report). If this evidence is correct, it argues for the existence of a level of processing which provides just the sort of preanalysis of the input string presupposed by the view of deep structure recovery we have been presenting.

The following is suggested as an outline for a solution to the problem of how the deep structure of sentences is recovered. It is assumed that a verb may be thought of as dominating a characteristic array of labeled slots in the deep structure of any sentence in which it appears. Which types of arrays a given verb may in principle dominate is specified by what we have called its lexical analysis. It is further assumed that the lexical analysis of each verb in the language is part of the information a subject has about his language's structure. Applying this information to an input involves essentially attempting to analyze the input as a substitution instance of one or another of the slot arrays its verb is capable of dominating. The preanalysis routine ought thus to provide as much information as possible as to which structures in the sentences are NPs and VPs and which of the NPs each VP implicates. Given such information, a systematic exploitation of the slot analyses and of transformationally-introduced surface structure grammatical markers (such as inflection, relative pronouns, order, etc.) ought to be sufficient to uniquely specify an underlying geometry for each input sentence.

This conception of the sentence processing routine makes a strong claim concerning the role of the verb in sentence decoding. The experiments described in this report provide some evidence for that claim. It should be remarked, however, that the considerations that hold for verbs may well have analogues for other parts of speech. Thus, for example, it is quite conceivable that the lexical structure of nouns and/or adjectives is exploited by the sentence recognition device, in which case the lexical complexity of such formatives ought to contribute to the perceptual complexity of sentences. Experimental determination of the perceptual consequences of lexical complexity in the case of formatives other than verbs would therefore be most important for the further evaluation of models of the sort we have proposed.

Section II: The Perceptual Segmentation of Sentences

In the discussion of the recovery of deep structure we have given above we have referred to the necessity of some sort of "pre-analysis" routine which will yield information about the probable segmentation of a sentence into its major constituents. That is, the input to the analysis rountine suggested by the research reviewed in the first section of this report is a sequence of NPs and VPs which are grouped in terms of their probable grammatical relationship to each other. The second major line of research effort has thus been concerned with the determination of the perceptual segmentation of the acoustic signals representing utterances in a natural language (in this case, English).

A number of experimental investigations have been directed at this problem. We turn now to a detailed examination of their results and of the theoretical considerations which motivated them.

The problem of "units," as Miller, Galanter and Pribram remark, is of fundamental concern to those interested in human behavior.

> "Most psychologists take it for granted that a scientific account of the behavior of organisms must begin with the definition of fixed, recognizable, elementary units of behavior...Given a simple unit, complicated phenomena are then describable as lawful compounds. ...For the most part, serious students of behavior have had to ignore the problem of units entirely. Or they have had to modify their units so drastically for each new set of data that to speak of them as elementary would be the most unblushing sophistry." (Miller, Galanter and Pribram, 1960, pp. 21-22).

The question of what units are relevant to discussions of language behavior initially seems vexed with the same difficulty. Depending on whether one looks at encoding or decoding, at connected speech or the production of isolated words, at idealized discourse (as in formal written material) or spontaneous speech, the descriptive "units" may vary. Linguistic descriptions of utterances provide several simultaneous levels of analysis: phonemes, syllables, morphemes, words, stress groups, immediate constituents, etc. The problem is not a paucity of units in which to describe language events, but in the determination of the most revealing alternative.

The question of the psychological relevance of linguistic constructs is, of course, an old question. For example, Sapir (1947) considered it important to point out that phonemes are, in fact, psychologically relevant units--not just arbitrary linguistic constructs. That is, it is evident that the facts which linguists remark about language are <u>psychological</u> facts (cf. Bever, 1966, and Fodor and Garrett, 1966). A grammar explicates a certain variety of psychological facts--intuitions which speakers of a language have about their language--notions of what is or is not a sentence, notions of when a sentence has two or more interpretations and what they may be, notions of intersentential and intrasentential relationships--in short, a set of complexly interrelated observations about language structure. These facts are just those which linguists have required the structural descriptions produced by a grammar should mark. In that sense, it is unarguable that, to the extent the observation statements are sound, and to the extent the linguist has been successful in writing his grammar properly, the structural descriptions produced by the grammar are psychologically real.

We are concerned, however, with the question of which aspects of the structural description are relevant to explanations of particular performance tasks. It may be, for instance, that mechanisms relevant to recognition or perception of sentences are not the same as those relevant to recall of sentences over varying periods of time. It will be of concern to us subsequently to carefully distinquish the possible effects of recognizing, producing and recalling sentences in the interpretation of experimental results.

Experimental techniques: the "click" phenomenon

The acoustic signal generated by a speaker for the representation of some sentence is a continuous signal. We are primarily engaged here to review a series of experiments for the information they may provide regarding the effective segmentation of such acoustic representations of sentences during their processing for interpretation.

The experimental technique employed exploits the difficulty (first reported by Ladefoged and Broadbent, 1960) which individuals encounter in locating short bursts of noise ("clicks") superimposed on recordings of continuous speech material. The research which will be described here may be succinctly characterized as the investigation of the possibility that the perceptual and memory errors in such a task are a function of processes involved in the decoding of sentences. Supposing it to be the case that errors in the location of clicks superimposed on sentential material are influenced by the processing of the sentence, then the nature of the errors for particular sentences may be utilized to provide an indication of their units of analysis and perhaps of the order of processing events as well. Such an assumption has clear analogues in visual perception work, especially that in the Gestalt tradition. Visual closure experiments exploit the notion of the perceptual unit as a determiner of the extent to which a figure will resist interruption.

The technique described will presumably reveal something of the segmentation of utterances. First, however, note that the boundaries defined by the clustering of click location errors need not correspond to any linguistic analysis--for example, people could sample the signal in time and integrate arbitrary chunks determined by the storage capacities of an auditory "memory bank." Further, it should be borne in mind that there are facts about the analyses of sentences which cannot be revealed just in its segmentation. That is, we must be careful to distinguish between the segmentation of the utterance in terms of its surface structure and the analysis of the utterance in terms of the underlying configurations (untransformed) of constituents. Either or both of these aspects of structural description might be relevant in an account of the subjective location of clicks in sentences. It is not clear, however, how one might regularly associate some particular boundary of an utterance's surface structure with the underlying configurations of constituents. What we will have to say about the displacement of clicks applies primarily to the <u>segmentation</u> of the utterance rather than to the analysis of the segments. That is, the surface structure of the utterance reflects the organization of its parts insofar as that is representable by bracketing. Aspects of the structure which are not representable by such surface characterizations need not be associated with any particular boundary. Consider some subconfiguration XYZ with a surface bracketing (X) (Y) (Z). If the relationship in the deep structure is (XZ) (Y), what boundaries in the surface segmentation should be associated with this configuration of elements? Either surface boundary is plausibly affected. Sentences of the sort <u>The theorem was proved by induction</u> (<u>from someone proved the theorem by induction</u>) or <u>He looked the number up</u> or <u>He has always read the passage correctly</u> illustrate the difficulty. In short, the "chunks" into which an utterance is broken for purposes of analysis of the deep structure must be distinguished from the results of that analysis.

It would, of course, be possible to argue that surface boundaries which result from some particular transformation of deep structure (as the boundary created by the deletion between <u>the man</u> and <u>the dog bit</u> in the <u>man the dog bit died</u> are more or less likely to show intrusions of clicks than are untransformed boundaries, or perhaps that the boundaries of embedded configurations dominated by #S# should be more salient. There are, in fact, many such notions derivable from a consideration of particular operations in the phrase structure or transformational subcomponents of grammars. Such questions are of great interest since ultimately it is this sort which has significance for the determination of the nature of the processing mechanisms which are employed during the recognition and production of sentences. We will, however, begin with simpler notions of the source of click location errors and examine them carefully before essaying more exotic explanations.

Ladefoged and Broadbent recorded several groups of test strings (sentences and randomly-ordered sets of ten digits) with superimposed sounds and presented them to various groups of listeners. Two sorts of extraneous sounds (for which no differential results are reported) were used: a sharp transient and the speech sound /a/. They found that, in general, errors were smaller with superposition of the sounds on randomly-ordered series of digits than with sentences; further, the tendency was for listeners to report hearing the sounds prior to their actual point of occurrence. It was found, too, that for different types of sentences, different distributions of errors were observed. These authors interpreted their results as examples of "prior entry"--i.e., a stimulus the subject is predisposed to requires less time for recognition than a stimulus he is not predisposed to. They sought to relate the degree of predisposition to the amount of information in the stimulus material. Though they felt this would explain the tendency to preposition and the overall difference between accuracy for digit strings and accuracy for sentences, they conceded that differences among the sentences are not amenable to such an explanation. Ladefoged and Broadbent regarded their results as evidence that the units of decoding are probably larger than the duration of a single sound unit. They concluded that one could be no more explicit than this on the basis of their data, since the size of the errors was variable. Ladefoged and Broadbent focused on a single level of structure in their analysis. Evidently, they were thinking of the speech signal in terms of an analogy with written materials--i.e., letter-by-letter scanning - phoneme-by-phoneme scanning. They surmised quite rightly, of course, the speech signal is segmented in larger-than-phoneme chunks. But how much larger? And how shall we characterize the relevant criteria for decisions about perceptual analysis of the speech signal?

An examination of the Ladefoged and Broadbent data for which frequency distributions were provided (five sentences out of thirteen), reveals that the extraneous sound (hereafter referred to as a *click*) may tend to shift toward major syntactic boundaries. This possibility takes account of exactly the fact that different sentences produced different sorts of errors. Accordingly, our attention is first confined to the determination of a relationship between click location and the surface structures of sentences. We will define the following notions with reference to the hypothesized effect of structure on click location: *Constituent break* and *strength of constituent break*. The boundary between any adjacent pair of words not a member of the same constituent is a constituent break. In the sentence

The ^{1} men ^{2} who ^{3} whistled ^{4} were ^{5} happy

2 and 4 are constituent breaks, for example. Strength of a constituent break is defined as being directly related to the number of constituents whose boundaries coincide at the break. In the example sentence, boundary number 4 is the strongest, since it is a boundary for five constituents: *whistled*, *who whistled*, *the men who whistled*, *were*, and *were happy*. Boundary number 2 has four constituents coterminous; no other constituent boundary is coterminous for more than two constituents.

If there is some direct relationship between the perceptual segmentation of the speech signal and the syntax, in particular, the derived constituent structure, using the notion just defined we should be able to predict concentrations of click location responses. Of course, it is not clear whether this "segmental" function should be viewed as continuous (i.e., concentrations of responses graded by number of constituent boundaries) over group data, or if one should expect more uniformity (i.e., exclusivity of the effect of syntax in terms of some given minimal boundary strength or type of boundary).

In fact, if we reanalyze the results reported by Ladefoged and Broadbent with this in mind, we see that their results are compatible with the hypothesis that clicks will tend to be perceived as occurring at constituent boundaries. In their sentences the click tended to be located either in the boundary just preceding the click or just following the click, depending on which of the two has the larger number of constituents coterminous.

Several studies aimed at careful testing of this notion have been carried out and will be described subsequently. As indicated above, however, there were several factors to be considered before the relation of click location errors to syntax could be determined. The error in subjects' judgments of click positions is suggested as being, in part, a function of the structure of the strings, in particular, the syntactic structure of sentences. Further, it is presumed that these errors are perceptual errors--that is, that the structure is productive of errors rather than merely providing the niches into which errors of memory and attention are sorted. This is a double-barreled assertion, i.e., that the positioning of the error reflects the syntactic structure and that the presence of such structure causes errors in the judgment of click location to be made.

In order to support this view, three propositions need be established.

1. All other things being equal, sentences produce errors larger or more frequently than do random or unstructured strings.

2. Errors that are made in judgment of click positions show a tendency to fall into syntactic breaks.

3. There is no other plausible mechanism operating to produce errors that are then sorted into syntactic breaks (assuming #2 were supported).

The evidence for the first of these propositions is clear. It is consistently the case that structured materials produce larger error scores than do unstructured materials. In the Ladefoged and Broadbent data and in modified replications of their investigations (Garrett, 1965) , random strings showed with few exceptions smaller errors than sentences of comparable lengths. The second and third propositions are more difficult. Several relevant facts emerged from the replication of Ladefoged and Broadbent's work that require a brief description of the studies.

Two response conditions were employed. These varied the degree to which demands were made on the immediate memory of subjects. In one condition *Ss* were required to recall the stimulus string well enough to write it down. In a second condition the subjects were required only to recognize the string. In both conditions *Ss* indicated their judgment of the click position by making a slash mark (/) through either their own written version of the stimulus string or through a written version provided for them by the experimenter. The stimulus strings in each case were presented binaurally to *Ss* via headphones from tape-recorded materials. The stimulus materials were sentences of varied syntactic type: random strings of words taken from the lexical stock provided by the sentences, strings of nonsense syllables and strings of digits. All types were varied also in length.

Several of the results here are relevant to determining whether the processing of the sentences is one major part of the cause of errors. If some other factor or factors are responsible for the errors (setting aside the question of why such errors might migrate to syntactic breaks), what candidates are there for the causal agent or agents? There seem three possibilities. (1) Subjects make errors of memory--i.e., they correctly perceive the position of the click, but forget it. (2) Subjects make errors of attention--i.e., they were never aware of the click's position. (3) Subjects make perceptual error--they misperceive the click position but for reasons other than the syntax of the sentence.

If it is the case that the errors are primarily memory errors, it is necessary to account for the difference in errors between sentences and unstructured material with such an hypothesis. Two difficulties are encountered. First, the random strings were most difficult for subjects to recall, *not* the sentences. Subjects almost never made errors of recall and recognition in the sentences, but frequently did with random strings. It's clear that the unstructured material made greater demands on *Ss*' memory, yet these were the strings which showed smaller errors. Second, there is the matter of comparison of the two response conditions. In these two conditions of subject response memory requirements were varied (one required only recognition, while the other required recall). Over these two conditions the errors produced by sentential material remained roughtly equivalent. (In a runs test for differences between error scores for sentences in the two conditions, the null hypothesis *cannot* be rejected for $\alpha = .05$.) The error scores for unstructured strings, however, were significantly reduced. (The null hypothesis is rejected for $\alpha = .05$ for funs tests comparing digits, nonsense syllables and random words in the recall condition with those in the recognition condition.) This indicates that errors in sentences are less accountable through memory than are errors for unstructured strings.

The simplest interpretation of an assertion that attention is an important factor in the explanation of errors is that subjects do not attend to some strings, and having not heard the click or much of the string, recorded a guess. There are several objections to this.

Subjects indicated (in post-test interviews) confidence in the correctness of their judgments. Further, subjects in both response conditions were instructed to mark those sentences on which they had simply missed the click position. There were only seven such responses (distributed over four subjects). This would tend to eliminate some attentional errors.

Most telling, however, are the differences between sentences and random strings. Why should it be the case that subjects have more frequent lapses of attention for sentences than for unstructured strings? If anything, the reverse seems more likely. In order to retain the hypothesis that attention lapses are responsible for the errors, an additional assumption is required. If one assumes that with very little or with no information concerning the position of the click, subjects will tend to place their responses at syntactic breaks, then it might be possible that the larger error scores of sentences are the result of a greater dispersion of their error responses. That is, the syntax is assumed <u>not</u> to produce errors, but only to determine placement of the responses once errors due to attentional lapses have been made. If this interpretation were correct, one would expect that the proportion of correct responses would be the same for both sentences and random strings. This was not observed to be the case, however. For both response conditions the proportions or responses occurring at the objective position are lower ($p = .05$ for the runs test) for sentences than for random strings. That is, <u>more</u> errors were made for sentences as well as generally larger displacements from the objective position.

Errors are hypothesized as occurring some time prior to the assignment of a semantic interpretation for the whole sentence. It is in the sense of this processing that the errors are referred to as perceptual. To establish that click location errors are perceptual errors in this sense is most difficult. The evidence reviewed above provides reasonable grounds for supposing that memory and attention are not adequate to account for the observed results. There remains the possibility that for the recall condition the retention and reproduction of the sentences are the source of the effect; for the recognition condition the effect of these factors is reduced but not absent. If, in fact, these data are to be interpreted as evidence for a particular kind of perceptual segmentation of sentences, it is necessary to provide additional evidence that the errors are made at some stage of the processing of the sentences, and to determine what that stage is.

There is the further possibility that, even granting that errors in judging the click position are perceptual, the misperception is the result of something other than the processing of the sentence. Various suprasegmental features of utterances may play an important role in the way in which they are understood. The possibility that the errors in subjective location of clicks reflect no more than the intonation contour of sentences, for instance, cannot be ignored. We will return to both of these questions. At this point, however, we need to consider the evidence for the relation of concentration of error response at constituent boundaries.

The distributions of errors for unstructured strings were used to establish frequencies of error response at increasing distances from the objective click position. Frequencies of response at various points in the sentences were then compared to determine whether they significantly exceeded (.99 confidence interval) the expected frequencies determined from the unstructured strings. In the recall condition, eighteen of the twenty-five categories which exceeded the confidence interval were major constituent breaks. In the recognition condition, however, the distributions for unstructured strings were so tightly clustered around the objective position that any response more than three positions (a word or space between words were counted as response categories) removed from the objective position was significant. So, in spite of the fact that inspection of the frequency distributions showed major constituent boundaries to be the most frequent response categories, comparison with the distributions from unstructured strings did not distinguish these categories.

A study conducted by Fodor and Bever (1965), however, shows the effects of syntax on click location in a much less ambiguous fashion. The response conditions were the same as those for the recall condition described above--S̲s were required to write out the sentences. The respects in which the Fodor-Bever study differed most from theprevious studies were (1) dichotic presentation of the stimuli, and (2) the method of assessing the effect of structure. In dichotic presentation of the stimuli, subjects receive the sentence in one ear, the click in the other. This complicates the problem of associating click and sentence for the subject. Work by Broadbent (1958) et al. indicates that subjects attending to dichotically-presented material tend to shift attention from ear to ear rather than attempt to attend to both inputs simultaneously. This would tend to enhance the effect of click displacement as compared with binaurally-presented materials.

Fodor and Bever made nine copies of each of their thirty sentences and placed clicks in nine different positions for each member of the set of thirty. The objective locations of the clicks were balanced around the major constituent breaks in each sentence. That is, for every sentence, one copy had a click placed at the strongest constituent break, and on the remaining eight copies four clicks were placed at progressively farther distances on either side of the major break. The categories of placement were syllables and word boundaries--as, for example, (where *O* =major break, △=click, and # = word boundary):

```
                    #         #          #
syllable  syllable   syllable   syllable   syllable  syllable
    △          △    △    △    △     △    △     △         △
    8          7    6    5    0     1    2     3         4
```

Their measure of the effect of the break was the number of responses shifted from the objective position in the direction of the break. Of all responses, (36 S̲s) Fodor and Bever found 80 percent to be errors. For the thirty sentences the error responses were in the predicted direction 66 percent of the time. This percentage is significant for $\alpha = .01$. Further, the hypothesis that errors would be in the direction of the major constituent break was true for each of the sentences taken individually and for the responses of each of the subjects. A subsidiary prediction is that clicks objectively located in the major break should be more accurately located than other clicks. This proved to be true in the Fodor-Bever data. Clicks in the major break were correctly located significantly more often than clicks located elsewhere. These results taken with those of the Ladefoged and Broadbent and Garrett studies provide strong support for the view that subjective click location is affected by the structure of the stimulus strings and that some aspect of the syntax is descriptive of the relevant structures.

With respect to proposition (3) (that there be no alternative mechanisms for the production of click location errors), we discussed the possibility of memory or attention as major factors earlier. Fodor and Bever also provide evidence which militates against these. They had subjects provide confidence ratings for each of their click judgments. If attentional lapses or difficulty of recall were important in theproduction of errors, a correlation between errors and S̲s' confidence in their judgments would be expected. Fodor and Bever, however, found no significant correlation of errors and confidence scores.

There are, however, more serious contenders: (1) possible acoustic correlates of the syntactic structures and (2) retention and reproduction of the sentences. The latter of these, it can be argued, should exert some effect for the same reasons as we have suggested the decoding of the sentences should. That is, if the processing of the sentence during its reception influences S̲s' location of clicks, then processing of the sentence during S̲s' production of it should also be influential. There is evidence, however, which indicates that the processing of the sentences during their reception is a significant source of the error even where S̲s are required to write out the sentences. First, we have the subjective report of subjects. S̲s' confidence in their judgments is generally quite high, i.e., they believe they heard the click at the point which they indicate. Further, S̲s generally report that they make their decision about the click location before they write down the sentence.

As suggested earlier, the fact that location error in these experiments can be accounted for by the derived constituent structure is susceptible to two interpretations. It may be argued either that displacement of clicks is responsive to the recognition of some acoustic correlate of the constituent structure or to the active imposition of a syntactic analysis onto the speech signal. That is, click location errors may be viewed as responsive to some process which filters the acoustic characteristics of the speech signal, or to a process in which the perceiver provides an abstract characterization of the signal in terms of syntactic properties not marked in the signal.

That response to some acoustic correlates of structure is a plausible candidate for explication of click location errors is suggested by a finding in the study by Garrett (1965) cited earlier. Some of the digit strings in those experiments had relatively long pauses introduced (as if a twelve-digit string had been read and then two of the digits erased). The pause positions showed significant concentrations of error response. In the Fodor and Bever study, however, the relation of errors to pauses in the stimulus sentences was examined and displacement of clicks to constituent boundaries was found even when the boundary was <u>not</u> marked by an energy drop in the signal. It is not possible, however, to rule out the effect of other correlates of structure (intonation contour, stress, etc.) and their interaction with pausal phenomena.

Since, as has been indicated, this is a crucial question in the interpretation of click location errors, an experiment was performed in which the effect of immediate acoustic features of stimulus sentences was controlled for (Garrett, Bever and Fodor, 1966). This was done by the use of ambiguous strings whose constituent structures depended on preceding context.

Six pairs of sentences were constructed for which some string of lexical items was common to each pair. For example:

(a) In her
hope of marrying Anna was surely impractical.
(b) Your

Common portions of each pair were made acoustically identical by tape-splicing. For example, in the above sentences the portion underlined taken from a recording of (a) was spliced to the portion Your taken from a recording of (b). When this spliced version of (b) is paired with a copy of the original recording of (a), there are two sentences in which the acoustic material for the latter portion is identical but for which the constituent boundaries are different. In sentence (a) there is a deeper boundary before the word Anna than in sentence (b), and conversely, in (b) there is a deeper boundary following the word Anna than in (a).

On the second track of each recording of a stimulus string, a capacitor discharge click was recorded (the intensity was approximately equal to that of the loudest vowel sound in the sentences; the duration was approximately 20 milliseconds). Clicks were placed in the middle of the word or words around which the constituent boundaries were manipulated (in Anna, above) and in the first syllable following the second deep break (in was, above).

These operations resulted in expansion of the original six pairs of stimulus strings to 24 stimulus strings. These were assorted into four groups of six such that each group contained one member from each of the original six pairs of stimulus strings. The stimulus strings were assembled on tapes with a variety of other sentences which were designed to prevent a bias for any particular click position. These "padding sentences" were, with minor variations, the same in all four groups. The stimulus sentences for each group appeared in the same order and in the same serial position in each group (i.e., across groups, any two successive stimulus strings were separated by approximately the same number and type of padding strings).

The stimulus material was presented dichotically (click in one ear, sentence in the other) over shielded headphones. Subjects were required to write down the entire sentence and to indicate (by making a slash mark through their written version) the point in the sentence where they believed the click to have occurred.

The responses were scored as in the earlier experiments. The response categories recorded were established by words and the positions between words. Where a subject made errors in recall of the sentence which made positioning of his response relative to the correct version indeterminate, the response was discarded.

Relevant contrasts are between the two members of each acoustically-matched pair at the points of variation in their constituent structure. All responses for the same interpretation of each member were combined (i.e., responses for the two different click positions were combined) for this comparison. A graphic representation of one of the distributions obtained is given in Figure 3. All responses are tabulated and the points of comparison are indicated.

The frequency of response in each of the comparison categories is given in Table 11 for the six pairs of stimulus sentences.

Six pairs of stimulus sentences (click position indicated by △):

1. (a) In order to catch his train
   (b) The reporters assigned to
   Geo△rge dr△ove furiously to the station.

2. (a) In her
   (b) Your
   hope of marrying An△na w△as surely impractical.

3. (a) Because it was a... city
   (b) Only the... district of
   Ham△bourg w△as leveled by the war.

4. (a) ...because many were afraid to give
   (b) ...although a majority of people did
   support drin△king li△quor was made illegal.

5. (a) As a result of their invention's
   (b) The chairman whose methods still
   influence the com△pany w△as given an award.

6. (a) No matter how well trained these
   (b) Living near an airport where the
   new pilots are fly△ing planes c△an be dangerous.

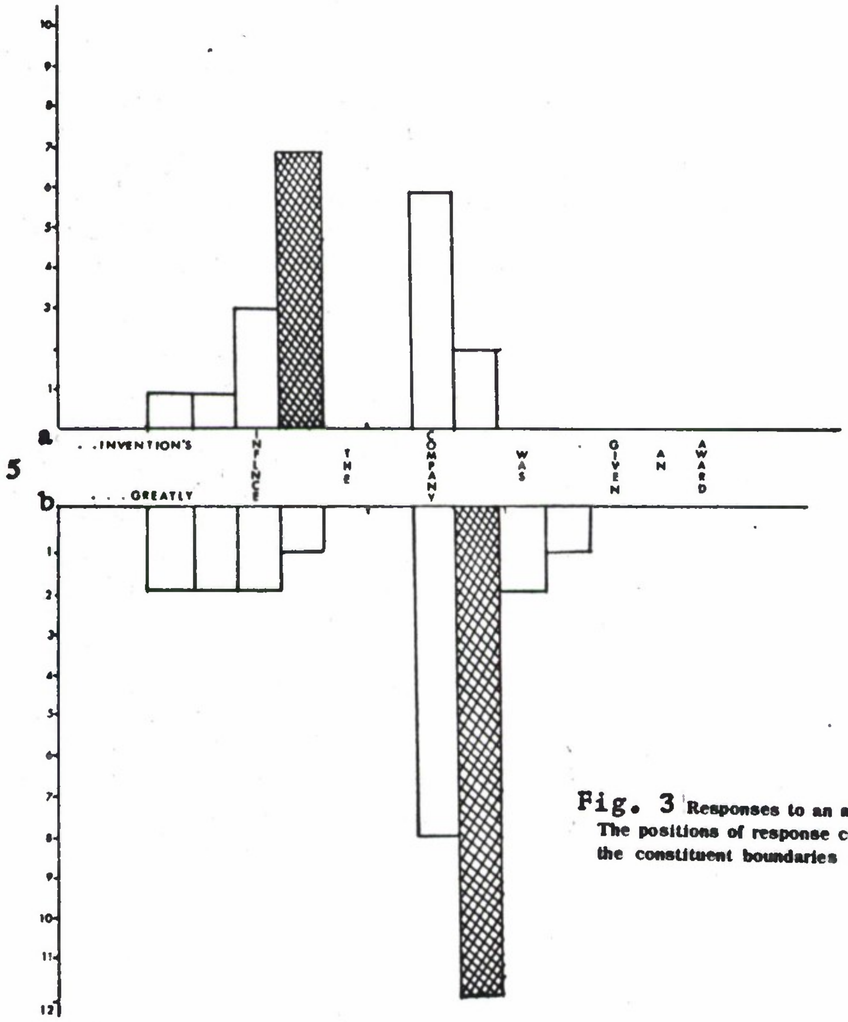


Fig. 3 Responses to an acoustically matched pair of sentences. The positions of response concentration predicted by variations in the constituent boundaries are cross-hatched areas in the graph.

Fisher's exact test was used to determine the probability of obtaining the observed frequencies of those more extreme.

Table 11

Frequencies of response at points of variation in constituent boundaries for six pairs of stimulus strings

(1)	"George"	a	9	0	p=.00005	(4)	"liquor"	a	6	0	p=.0022
		b	1	11				b	0	5	
(2)	"Anna"	a	13	2	p=.1130	(5)	"the co."	a	7	2	p=.0015
		b	5	4				b	1	12	
(3)	"Hamb".	a	11	0	p=.003	(6)	"planes"	a	7	3	p=.1850
		b	5	11				b	3	5	

As Table 11 shows, four of the six pairs are significantly different ($p < .01$). The remaining two pairs, although not significantly different, show shifts in the predicted direction. If one considers only which of the two comparison categories shows the larger number of responses, the pattern conforms to variation in constituent boundaries for eleven of the twelve instances in Table 11 ($p = .003$ for this pattern).

A second sort of comparison can be made. This contrasts positions within sentences rather than between sentences. For each click position the frequency of response at the predicted position can be contrasted with the frequency at all other boundaries adjacent to the click. The assumption here is that if the structure is not an effective determinant of the click's subjective location, there should be no tendency for the constituent breaks to show the greater frequencies. Such an additional comparison demonstrates that the results reported above were not specific to one of the two click positions.

Even under these relatively unfavorable conditions (i.e., two "non-break" positions sometimes weighed against a single constituent break position), the structure of the sentences is a markedly successful predictor of location errors. Of the twenty-four comparisons, twenty show greater frequencies in the categories predicted by syntactic description of the sentences (this pattern is significant, $p < .001$).

The assertion has been that the errors made under the conditions of these experiments are in part perceptual errors--errors introduced during the active processing of the sentences that is necessary to their understanding. Investigations to this point have been open to an alternative explanation, however. Since subjects have previously been required either to reproduce the sentences or to recall them well enough to recognize them, it might be argued that the displacement of the location of the click occurs some time during the interval after the sentence has been heard and understood before the subject records his judgment. This asserts that the errors are errors of recall or are produced by the encoding or by some combination of the two.

An experiment was done (Garrett, 1965) in which the role of memory was reduced in the process of judging the position of the click, and in which the subjects were freed of any requirement to reproduce the sentences. The intent was to require only that the subjects attend carefully to the meaning of stimulus sentences presented aurally--that they be required to provide a semantic interpretation for the sentences with only the auditory signal as input.

In order to accomplish this, the following method was employed. Subjects were presented with the same sentence twice in succession. Each sentence had one click embedded in it. <u>S</u>s were then asked to judge whether the clicks in the two sentences occurred at the same position or in different positions. <u>S</u>s were led to believe that either condition was possible by instruction and by inclusion of dummy pairs for which both clicks were in the same positions. For the test pairs the clicks were placed in different position. If the hypothesized effect on perception of the click position occurs, manipulation of the structure of the sentences should produce variation in the incidence of <u>same</u> judgments. The assumption, of course, is that two clicks will be more likely to converge (be heard at the same position) when a major boundary intervenes, as ("↓" indicates major constituent break, "△" indicates click position):

(a) The boy who delivers pap$\underset{\triangle}{\text{e}}$rs ↓ has gone home.

The boy who delivers papers ↓ h$\underset{\triangle}{\text{a}}$s gone home.

(b) The boy who deli$\underset{\triangle}{\text{v}}$ers papers ↓ has gone home.

The boy who delivers pap$\underset{\triangle}{\text{e}}$rs ↓ has gone home.

Other things being equal, the clicks in pair (a) should converge more often than those for pair (b) since a major boundary intervenes for (a), but not for (b). It is, of course, possible for the clicks in (b) to converge at the same point as those for (a), or in other positions as well, but it was assumed that this would happen less frequently for (a) than for (b). In this experiment acoustic correlates of structure (cf. Garret, Bever and Fodor, 1966) and the position of clicks in compared pairs of sentences were controlled for. An example of this is the following.

(c) Your hope of marrying Anna△↓was surely impractical.

Your hope of marrying△ Anna was surely impractical.

(d) In her hope of marrying△↓Anna was surely impractical.

In her hope of marrying Anna△ was surely impractical.

With clicks in the indicated positions, it was predicted that these sentences would show variation in the number of _same_ responses although (1) the clicks are in the same positions for both pairs, and (2) at the point where the clicks are placed, the acoustic signal is identical for both pairs. Subjects hearing pair (c) should give fewer same responses than those hearing pair (d); in pair (d) there is a deeper break between the positions of the two clicks than is the case for pair (c). Another way of expressing it is that for pair (d), the clicks have to cross fewer boundaries to achieve convergence at a major break in the sentence.

The stimulus strings used here were the same as those used in Garrett, Bever and Fodor (1966), six acoustically-matched sets of two sentences. Each set of two sentences yielded four _stimulus pairs_. Referring to the examples given above (pairs (c) and (d)), two stimulus pairs result from the repetition of each sentence from the acoustically-matched set, with particular click locations, i.e., each stimulus pair consists of the repetition of a sentence and a particular configuration of clicks (two sentences X two click configurations = four stimulus pairs). The two click configurations used are referred to as _position A_ and _position B._ Each acoustically-matched pair has a key word or words around which the structural variations are made. The word _Anna_ is the key word in the examples given earlier. Position A has clicks placed to test the effect of the break before the key word or words (in one of the interpretations). In the examples given (stimulus pairs (c) and (d)) earlier, the configuration of clicks is for position A. Position B has clicks placed to test the effect of the break after the key word or words (for example, in (c) and (d), this would be in _Anna_ and in _was_). The six sets of four stimulus pairs were assorted into four groups such that each group contained one member from each of the original acoustically-matched pairs of sentences. These four groups of stimulus strings were assembled into four tapes with a variety of other pairs designed to prevent bias for any particular click position, relation of clicks to structure, or relation of successive clicks (within a particular stimulus pair) to each other. The materials were presented dichotically over shielded headphones. _S_s were provided with a response booklet in which they recorded their judgment of the click positions as _same_ or _different_. In order to insure that _S_s were attending to the content of the sentences (i.e., were, in fact, attempting to provide an interpretation for it), Ss were required to state (after recording their judgment about the click positions) whether the two presented sentences were identical.

As indicated by the examples, the comparisons here were between stimulus pairs with clicks in the same objective positions; the contrast is between instances in which the clicks are both adjacent to a major break (as word$_{\Delta^1}$ ↓ word$_{\Delta^2}$ word) and instances in which only one of the clicks is adjacent to a major break (as word$_{\Delta_1}$ word$_{\Delta_2}$ ↓ word).

Since the four experimental groups were of equal size (n equaled 20), a direct comparison of the frequency of same response is given. Table 12 gives the frequencies for comparison for both click configurations; the results for stimulus strings derived from each acoustically-matched pair are reported together. Fisher's exact test

Table 12

Frequencies of same responses for variation in phrase boundaries in twelve pairs of sentences

		Position A		Position B	
"George"	train	*17	p=.348	2	p=.062
	reporters	15		* 7	
"Anna"	in her	* 9	p=.137	13	p=.239
	your	5		*16	
"the co."	invention	*12	p=.056	3	p=.224
	chairman	6		* 6	
"Hambourg"	city	* 7	p=.136	11	p=.137
	metro	3		*15	
"drinking"	afraid	*15	p=.640	3	p=.077
	majority	15		* 8	
"fly. pl."	no matter	* 4	p=.501	12	p=.739
	living	3		*11	

* larger number of same responses predicted.

was used to determine the significance of differences in frequency of same responses for the pairs with clicks in the same positions. The probabilities computed are given in the table with the frequencies. The first column in the table is for comparison of stimulus pairs with clicks positioned for testing the first deep break (position A). The second column is for comparison of stimulus pairs positioned for testing the effect of the second deep break (position B).

This experimental technique was intended to overcome two related difficulties with earlier methods of assessing the effect of sentence structure on the subjective location of clicks. The description of errors in click location as perceptual errors and, hence, as informative of decoding events, rests on ruling out memory and encoding variables as major determinants of the effect. This method required subjects to operate with only aural input, requires no delay in recording responses, and eliminates any requirement that the subject reproduce the sentences being tested.

If the assumptions about the method are accepted, the results here support the proposition that perceptual analysis of the sentences is a significant source of click location error. Although the changes in frequency of <u>same</u> responses under these conditions were not generally large, the two groups of scores defined by distance of clicks from a major boundary are clearly different. For position A, the prediction was for the interpretation with a break before the key word or words to show the largest number of <u>same</u> responses; five of the six pairs showed such a directional shift, with the sixth pair showing no change. For position B, the prediction was for the interpretation with a break after the key word or words to show the largest number of <u>same</u> responses; five of the six pairs did. For the twelve comparisons, ten of eleven cases in which there was a change showed a shift in the direction predicted by the structural variation in otherwise identical stimulus items; p=.006 for a sign test of this pattern of changes.

Examination of the comparisons made for individual stimulus pairs shows that none of the shifts in frequency of <u>same</u> response is significant for $\alpha = .05$. This is not surprising in light of the nature of comparisons here. Under these experimental conditions, the effect of a relatively strong syntactic break is contrasted with the effect of a weaker one (or a more distant strong one). Further, in view of the small sample size, the lack of significant changes for individual pairs should not be considered remarkable.

An unexpected result was the significant variation in the level of <u>same</u> responses among stimulus pairs for positions A and B (cf."George" sentences: the incidence of same responses for position A is high, for position B, low). This variation is due to differences in the <u>relative</u> positions of the two clicks in a stimulus pair. In half the stimulus pairs heard by any subject, the objective position of the click in the first presentation of a sentence was prior to the objective position of the click in the second presentation of the sentence; the relationship was reversed in the other half of the stimulus pairs in a group.

An analysis of variance shows that this effect is significant ($p < .001$) and independent of the structural manipulation (the test for interaction was not significant for $\alpha = .05$). *S*s' reports in post-test interviews indicate this relative click position effect was due to the strategy adopted by *S*s in the experimental situation. *S*s listened to the second presentation of a sentence in light of their decision about the position of the click in the first sentence; *S*s were "sensitized" to a particular position in the test sentence--their decision awaited the arrival of that position in its second presentation. Couched in terms of *S*s' expectations, where the second click occurred prior to the expected point (assuming either that the first click was correctly located or was perceived as occurring at the major constituent break), the level of *same* responses was *high*; where the second click occurred after the expected point, the level of *same* responses was *low*. The effect lends additional and independent support for the proposition that errors of click location do occur during the sentence's processing. That is, it is difficult to see how memory or encoding variables could plausibly account for the effect of relative click positions.

The research so far is conveniently discussed under two headings--the status of evidence that the immediate constituent analyses of sentences are related to subjective click locations and the status of evidence that such location errors are perceptual in some interesting way. Perhaps a more incisive way to express the difference in the two goals is to characterize the first as an effort to establish a relationship between two variables, the second as an effort to determine if it is a causal relation. The former is, of course, much the easier task.

Evidence of a relationship. The several experiments reviewed so far all provide strong evidence that errors in location of clicks are correlated with the structure of the material on which the clicks are superimposed. Three different response conditions were used: recall, recognition, and same-difference judgments. Four methods of measuring the relationship were used: tests of responses to structured material against distributions established for unstructured material, covariation of error responses and structural changes in sentences, direction of error shifts, and frequency of convergence of two separated clicks. The results in all these conditions have indicated a significant relationship between structure and subjective click location.

Evidence of causal factors. In attempting to establish the source of the relationship between subjective click locations and structural features of sentences, two considerations were relevant. First, there is the possibility that the effect is not perceptual at all, i.e., that errors do not occur during processing of sentences but at some later time as the result of memory failures or encoding processes.

The question of these factors has been discussed earlier. The second consideration is the possibility that the effect in perceptual, but in an uninteresting way. That is, the click location errors may occur during sentence processing but not as a result of that processing. Such errors might stem from either (1) some state(s) of the decoder quite unrelated to properties of the sentence (as attentional factors), or (2) some properties of the sentence which are partially correlated with its semantic and/or structural processing.

The effect of attentional factors, even if productive of some effect during processing, cannot account for differential effects among sentences. They can account only for some general dispositional tendency (as pre- or post-positional tendencies) and are, therefore, rejected as irrelevant here.

Much the more significant problem in assessing the source of click location errors is the possibility that they stem from properties of sentences that are correlated with structural features--in particular, acoustic cues such as pause and intonation contour. If the click effect were primarily responsive to such acoustic correlates of structure, it would have much reduced relevance to processing operations which have been assumed to be involved in sentence understanding. The effect would have to be regarded as a surface phenomenon--much less interesting than a response to stimulus properties not objectively marked. The difference lies in events which we can confidently assert demonstrate something about the organism rather than about the stimulus to which the organism responds. The findings in Bever (1965), in Garrett (1965), and in Garrett, Bever and Fodor (1966) demonstrate conclusively that click location errors cannot be accounted for in terms of such surface features.

It should perhaps be emphasized that finding errors of click location responsive to deeper syntactic features does not allow one to conclude that click placement cannot be affected by such things as pause and intonation. There is, for instance, no way of telling from the data of these experiments how much greater, if any, the contrast between the members of the acoustically-matched pairs would have been if each sentence were read with its normal intonation rather than a neutral reading. Further, it does not imply that such acoustic correlates of structure do not play a role in the sentence-understanding process. What the evidence does demonstrate is that, whatever the effect of intonation, etc., on click location errors, it is possible to investigate deeper syntactic processes with this technique. The exact extent to which these correlates of structure must be controlled in order to do this is something that remains to be determined.

Attention is turned now to the evidence relevant to the question of memory and encoding variables as an explanation of click location errors. As was indicated in the earlier discussion, explanation in terms of memory failure is suspect on several grounds: differences between errors for structured and unstructured strings, equivalence of error scores for sentences over variation in memory requirements, lack of correlation between subject confidence in judgment and errors, and inability to account for concentrations of error response at constituent boundaries. In short, there are grounds for doubting significant errors of recall and, in any event, memory failure alone cannot account for the data--some mechanism for concentrating memory-induced errors at word and phrase boundaries is required.

Encoding processes may be considered either as the required adjunct to memory errors, or, independently, as both the source of error and the determinant of click location. Of all the possible alternative explanations, this is the most likely in terms of the mechanisms proposed to account for a perceptual error during sentence decoding.

The evidence from the same-difference experiment, however, indicates that click location errors are attributable to decoding operations. Considering first the possible role of memory factors in that experiment, there is a much-reduced interval between presentation of the stimulus strings and recording of responses as compared with the earlier studies. The interval between the two presentations of a sentence was approximately 1.3 seconds (ten inches of tape at seven and one-half inches per second), and recording of the <u>same</u> or <u>different</u> response immediately followed completion of the second presentation of the sentence. Just as in the previous investigations, subjects regularly reported confidence in their judgments during post-test interviews.

In order that encoding errors might be assumed to be relevant to the errors in this experiment, covert rehearsal of the sentences is required. The interval between the two presentations of the sentence was not adequate for rehearsal of the entire sentence--even the speediest mumbler would have been hard pressed to squeeze in more than a repetition of the click's position (as "the click was at X"). Subjects regularly recorded their judgments immediately after the second presentation of the sentence. Although they were told they might have as much time to deliberate as they chose, subjects did not display any hesitancy. There was not time for rehearsal after presentation of the second sentence. Further, subjects indicated in the post-test interviews that they did not try to repeat the sentences during or immediately after their presentation. It seems clear that encoding could not have played an important role in the determination of click position here.

It is true, however, that the structural effect in the same-difference experiment was not as pronounced as that found (with comparable sample size) in Garrett, Bever and Fodor (1966). This suggests that the requirement of recalling sentences does enhance the effect of structure on the location errors.

The nature of effective structures

At this point it seems clear that errors in click location are related to perceptual analysis of sentences and that the constituent structure provides a characterization of relevant structural variables. Experiments by Bever, *et al.*, however, indicate that the derived constituent structure does not entirely represent the relevant *perceptual* segmentation of sentences. In the previous experiments the tacit assumption was that all constituent boundaries are effective in producing apparent shifts in click location--the more boundaries coterminous at a given point, the more likely a concentration of error responses. In two experiments using the same stimulus material but different response modes, however, Bever, Kirk and Lackner found in one instance no significant effect of minor constituent boundaries and in the other instance a weak effect. Though the differences in the outcome of these two experiments are equivocal, the relationship between the minor structure effect and the major structure effect is instructive.

The stimulus materials used in these experiments were 25 twelve-word sentences with constituent structure varied as follows: Each sentence had one strongest break occurring once at each of the positions following the 4th, 5th, 6th, 7th and 8th words for each of five types of sentences. The words preceding and following the higher break were monosyllables. The constituent relationships of the words immediately adjacent on either side of the major break were combinations of three types: right branching (R), left branching (L) and ternary (T). Figure 4 provides an example of these relationships.

On the left of the major break in the figure there is a right-branching structure, and on the right there is a left-branching structure. Of the nine logically possible configurations, five were used: R and R, L and L, T and T, R and L, L and R. There were five copies made of each of the 25 sentences, and as Figure 4 indicates, clicks were placed in the break and in the adjacent two words on either side. The click positions are hereafter referred to as positions A, B, C, D, and E (see Figure 4).

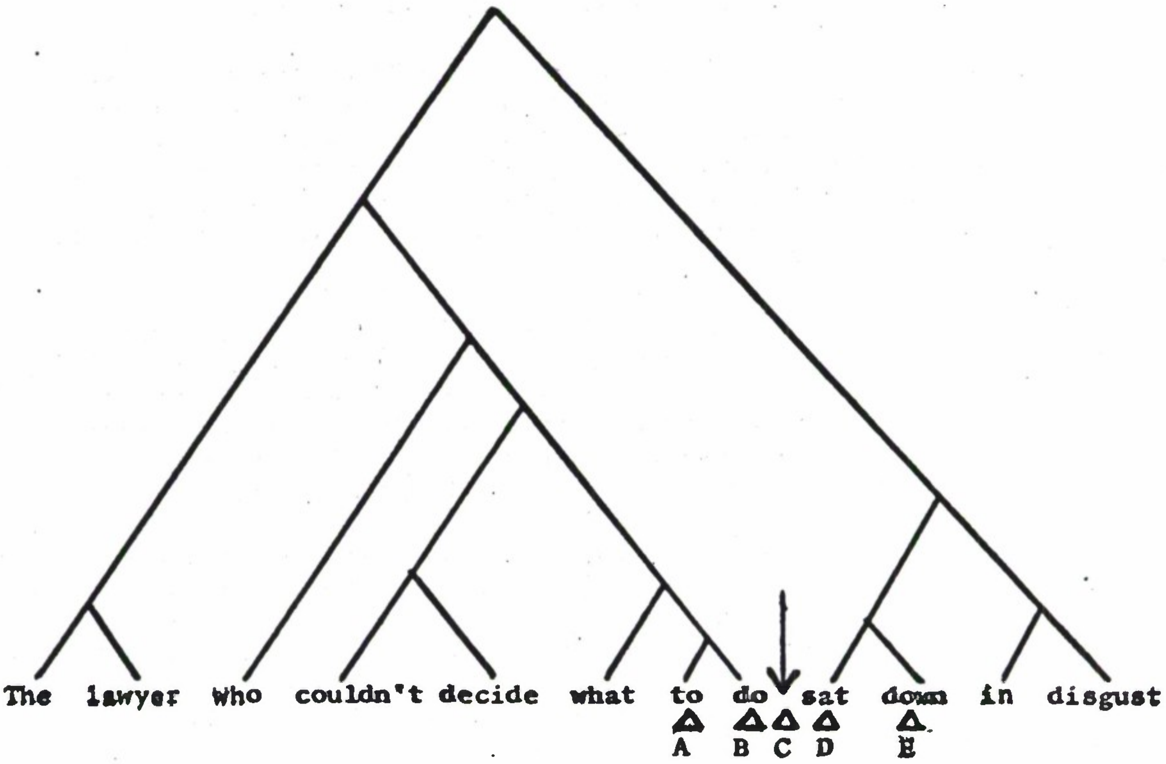


Figure 4. The derived constituent structure for a stimulus sentence used by Bever, Kirk and Lackner. ↓ indicates the major break and Δ indicates the positions of clicks in the five instances of the use of the sentence.

The dominant tendency under these circumstances should, of course, be for displacements into the major break. However, the strength of that tendency should be modified by the character of the minor structure for click positions A and E (see Figure 4). Where position A is in a right -branching structure there should be a reduced tendency for a shift to the major boundary as compared with sentences in which position A is in a left-branching structure. The minor break there favors movement away from the major break. Just the reverse is true for position E--a left-branching structure should reduce the frequency of shifts to the major break rather than a right-branching structure as is the case for position A. The ternary structures should produce neither of these asymmetries.

Clicks in position C (in the major break) should most often be correctly located. There is the further possible prediction that, whenever errors are made in the location of clicks at position C, the direction of the error will be influenced by the minor structure on either side: R:R should show post-positional error; L:L should show pre-positional errors and the remaining configurations should exert no biasing effect.

As indicated, the two experiments differed in their response conditions. In the first experiment (Bever, Kirk and Lackner, 1967) Ss responded verbally, first telling the location of the interfering stimulus, and then repeating the sentence. The experiment also differed in the use on a nonauditory interfering stimulus; a mild electric shock delivered to the wrist was used rather than a click.

Although under these changed conditions the major constituent breaks continued to show a significant effect, the minor constituent boundaries did not show any significant effects. For every click configuration and for all five minor structure types the major boundary was significantly more often named as the shock location ($p < .05$). The incidence of shifts to the major boundary from positions A and E was not significantly affected by the change from left-to right-branching structures. Nor did the variation in minor constituent boudnaries produce either a significant change or a discernible trend in the pre- or post-positional tendencies of shocks located at the major boundary.

In the second experiment using these stimulus materials (Bever, Lackner and Kirk, 1967) an auditory click was used. Subjects were required to write out the sentences and indicate the position of the click with a slash mark through the written version of the sentence.

In this experiment, as in the previous one using mild shock, the major constituent boundaries produced a significant effect. Table 13 contrasts the incidence of error shifts of four given dimensions where that shift is into a major break and where it is not. Each of the twenty-five sentences appears with five different click positions (see Figure 4). Hence, for example, there are twenty-five sentences for which a +3 (3 position post shift) error results in placement at a major constituent break and 100 sentences for which the same error does not result in placement at the major break; and, similarly, for each of the remaining four click positions and errors of +1, 0, -1, and -3.

Table 13

Frequency of specified errors for shifts into major constituent boundaries vs. shifts not into constituent boundaries

	FREQUENCY OF ERROR				
Error Size and Direction	Into Major Break		Not to a Major Break		
+3 (post-position shift of 3 places)	Pos A	92	Pos	B	1
				C	2
				D	9
				E	7
+1	Pos B	206	Pos	A	82
				C	34
				D	25
				E	40
0	Pos C	394	Pos	A	173
				B	379
				D	158
				E	168
-1	Pos D	343	Pos	A	72
				B	48
				C	213
				E	56
-3	Pos E	170	Pos	A	34
				B	23
				C	19
				D	21

In every instance, the frequency for a given error is greatest when it produces a shift to major constituent boundaries in the sentences. The differences between major boundary shifts and other shifts are significant ($p < .05$) for each of the 5 click positions.

If one looks for an effect of the minor structure variation on incidence of shifts to major boundaries, no trend is evident, just as was the case for the experiment using verbal report and these stimulus materials. However, the sample size for this experiment was much larger than that for the Bever *et al.* experiment (300 *Ss* vs. 25 *Ss*). Therefore, it was considered that an effect of the minor structure, independent of the effect of the major boundaries, might be discoverable by attention to single position shifts for click positions A and E. In these comparisons the effect of the major break is presumably balanced and only the variation introduced by the changes of left, right and ternary branching are relevant. Comparing the right-branching structure with left-branching at positions A and E, it was predicted that the right-branching structures should show a greater proportion of -1 (one position pre-shift_ errors than the left-branching structures while the left-branching structures should show greater proportion of +1 (one-position post-shift) errors than right-branching structures (see Figure 4). Ternary structures were expected to show proportions intermediate to those of the right- and left-branching structure.

The differences between the proportions of error in each of the eight comparisons of right- and left-branching were small and individually were not significant, but the changes were in the predicted direction in seven of the eight cases ($p < .035$). The ternary values did not conform to the expected pattern. Four changes were in the predicted direction; four were in the reverse direction.

The results for the ternary structures are puzzling and, coupled with the weakness of the effect found for the right- and left-branching structures, suggest a cautious interpretation. If, in fact, there is some effect of the minor constituent boundaries, it is certainly miniscule compared with the effect of the major boundaries in both the experiments. The disparity of the minor structure effect and major boundary effects suggests a difference in kind rather than of degree. In fact, one can examine the data from these two experiments in a somewhat different way in order to reveal the failure of earlier assumptions concerning the effect of constituent structure on click locations. The manipulation of minor structures was predicted to produce an effect on the basis of an assumption that the effectiveness of major boundaries is due to the tendency for large numbers of constituents to end and begin there as opposed to other positions in sentences. Only these positions were considered in all the previous experiments, of course, because the effects are so much more prominent there. However, because of the control of the distance of clicks from the major breaks in the sentences, the data from the two experiments by Bever, *et al*., can be examined for a correlation between the number of constituent boundaries coterminous at the various major breaks and the incidence of response at the major breaks. If the assumption concerning the *reasons* for effectiveness of the major breaks is correct, one would expect a significant correlation between number of boundaries and number of location errors into the major break. Three such correlations were computed using the data from Bever, *et al*.: (1) for clicks at positions A and B, the incidence of shifts to the major boundary was correlated with the number of constituents to the left of the major break which had boundaries coterminous with that break; (2) similarly, for clicks at positions D and E, shifts to the major break were correlated with the number of constituents to the right of the major break with boundaries coterminous with that break; (3) the third correlation is between clicks located accurately (when they occurred at position C) and the number of constituent boundaries *from structures both left and right which terminate at the major break*. None of these three correlations (Kendall τ) was significant and all were small ((1) $\tau = .086$; (2) $\tau = .113$; (3) $\tau = .170$). It appears from this that the extent to which the major break is a factor in the location of clicks is not affected in any serious way by the number of constituents the click interrupts or the number of constituent boundaries coterminous with the major break.

These findings coupled with a careful examination of results for individual sentences in several past experiments led Bever, Fodor and Garrett to suppose that a significant determinant of the extent to which a boundary affects the subjective location of clicks is the presence or absence of an S-node in the derived structure dominating a constituent, i.e., embedded versions of sentences produce the boundary effect for click location errors. Recent changes in linguistic theory made this a much more plausible notion than was the case when the initial investigations of click location errors were undertaken. That is, in the structural descriptions afforded by earlier versions of transformational grammar, one could find not only the desired sort of surface constituents dominated by S, but others as well. For instance, in the sentence *the little boy that he hit ran home crying*, the constituent *that he hit* is dominated by S in the surface structure; so, however, is the constituent *little*. This latter circumstance stems from the fact that in the grammar, *Adj + Noun* constructions are formed from an underlying structure *Noun which is Adj*. In the above example, the constituent *little boy* comes from *boy which is little*, and this structure is dominated by S. The grammar thus provided no motivated distinction between structures which seemed to satisfy the constraints on structures which appear to affect click location and those which did not. In order to save our notions about the role of S nodes in the determination of click locations by subjects, it was necessary to make a distinction between varieties of S nodes--those that branch and those which do not. That is, the first tentative move was to make a distinction among aspects of structural descriptions on the grounds of our experimental results alone.

A recent reformulation of the grammar (Ross, 1966) proposes a rule which mitigates the problem. In order to avoid having the grammar make counter-intuitive claims, he proposed that such nodes should be deleted from the structural description. The deletion of embedded S-nodes which do not dominate both NP and VP would be considered a condition on the well-formedness of structural descriptions.

To avoid having the grammar produce counter-intuitive results by the simple expedient of throwing away those parts of the description one doesn't want around is unsatisfactory for the linguist, however. He wants a motivated restriction on the grammar--that is, Ross' rule of "tree-pruning" (delection of S-nodes) must be imposed for causes internal to the grammatical system. Ross demonstrates by several examples how the S-node deletion rule interacts with other rules of the grammar to avoid several unsatisfactory results other than just those of "improper" S domination in the derived structure (see Ross, 1967, for an extended discussion of this and related considerations).

An experiment was carried out to test this revised notion of the source of click location errors by Bever, Fodor and Garrett (1966). In this experiment several more subtle distinctions based on the initial formulation of the tree-pruning rule of Ross were tested as well as the effectiveness of elements obviously dominated by S in the derived structure.

The stimulus materials used consisted of triples of sentences using so far as possible the same lexical items in all three. The triples were of three general sorts, as follows:

(1) I think that John likes $\left\{\begin{array}{l}\text{history very much but won't admit it}\\ \text{his mother very much but won't admit it.}\\ \text{his going very much but won't admit it}\end{array}\right.$

(2) The pedestrians watched the light $\left\{\begin{array}{l}\text{blue car}\\ \text{on the corner}\\ \text{turn green}\end{array}\right.$

(3) The colonel assumed $\left\{\begin{array}{l}\text{the command of the regiment}\\ \text{their command of the situation}\\ \text{the commands were self-evident}\end{array}\right.$

In the sentences of the first type of triple, there is a contrast between (a) structures which were not dominated by S at any point in the derivation (e.g., history in the example above; such sentences will be referred to as "no-S" sentences hereafter), (b) structures which were dominated by an S that the tree-pruning rule deletes (e.g., his in the example (1) above, referred to as S-deleted hereafter), and (c) structures which are still diminated by S (his going in the example above, referred to as S-dominated hereafter). In the sentences of the second type of triple, the structures tested were either S-deleted or S-dominated. These latter are closer approximations to well-formed sentences than are the S-dominated structures in the first type of triple. In the third type of triple, the contrast is between either two S-deleted structure and two S-dominated structures. Where there are two structures the same with respect to S-domination (deleted or undeleted), they differed in some other respect. Another example of a type (3) triple which has two S-dominated structures would be:

$$\text{The little old lady believed the speaker} \begin{cases} \text{on the radio (S-deleted)} \\ \text{to be truthful (S-dominated)} \\ \text{was truthful (S-dominated)} \end{cases}$$

In every instance of the type (3) triples, one of the S-dominated structures is one which would be well formed if it appeared in isolation.

For a given triple, clicks were placed in portions of the sentences which were identical for all three members and which were, in most instances, removed one response position from the break being tested. The members of the triples were assorted into three groups such that the various types were equally distributed among the groups. Subjects were required to write out the sentences and mark the position of the click with a slash mark as in most of the earlier experiments. This response mode was chosen since it gives maximal opportunity for differences to be manifest if they are present. Thirty paid volunteer <u>S</u>s were run (ten in each of the three groups).

There were eight triples of type (1); analysis of the results does not show any significant differences between the structures of this type which were assumed to be dominated by an S-node and those with deleted S or with no S in their derivational history. When S-dominated sentences in type (1) triples are compared with their S-deleted versions, four of the S-dominated types show larger proportions of error response at the test boundary and four show smaller proportions than the corresponding S-deleted types with no-D types show no significant differences. In both instances, however, the trend is for the structures with <u>S</u> in their derivational history to show greater effects than those with no <u>S</u>. It is possible that with larger subject groups the trend might prove to be statistically reliable. We will return to the question of the effectiveness of type (1) structures below.

There were six triples of type (2) and five of type (3). These are reported together since analysis separately does not affect the outcome in any important way. Of the twenty-two comparisons of S-dominated structures with S-deleted structures, seventeen of the sentences with S-dominated constituents showed larger concentrations of response at the test boundary than their S-deleted counterparts, while five showed smaller concentrations of response ($p < .005$, Wilcoxon test).

The primary concern in this experiment was to determine whether a difference in click location errors was effected by the presence or absence of an S-dominated constituent in the surface structure of sentences. On the basis of the clear cases from the type 2 and type 3 triples, it appears to be true that the S-dominated structures are significantly more effective in producing errors in click location than are S-deleted and no-S structures. The failure of the triples of type (1) to show significant differences among the three types of structures, however, confronts us with certain difficulties of interpretation.

There are several options. It may be that our analysis of the structures is wrong--i.e., that the structures (e.g., *his going* in *I don't like his going*) we have designated as S-dominated are *not* S-dominated or, that those designated as S-deleted (e.g., *him going* in *I don't like him going*) are, in fact, S-dominated. On the other hand, it amy be that the analysis is correct, but that the click location errors are not properly characterizable as sensitive to S-domination in the surface structure. That is, since the S-dominated structures in triples of type (2) and (3) *were* more effective than their S-deleted counterparts, while this was not true of triples of type (1), it may be that the click location errors are sensitive to some other property of the surface structure or the relation between surface and deep structure than presence or absence of an S.

From a study by Bever, Lackner and Kirk we can adduce some further evidence which bears on these problems. We have already considered the design of this study in connection with the effects of "fine structure" on click location errors (cf. page 88). Though the study was not designed to test for questions concerning the effect of deep structure-surface structure relationships or of S-domination of constituents, some of the stimulus sentences they employed can be looked at with these questions in mind.

It will be recalled that the manipulation of minor constituent structure in that experiment failed to produce significant effects on click location errors. It is possible, however, to subdivide the fine structure predictions into two categories: those in which the constituent break to be tested corresponds to the division between structures derived from different S-nodes in the deep structure and those in which the constituent break is between items dominated by the same S in the deep structure. In cases of the former kind, the "fine structure" prediction was borne out (20 of 23 cases, $p < .001$); in cases of the latter kind, the fine structure predictions were not borne out (19 confirming cases, 32 disconfirming or no difference). Although this does not bear directly on the question of S-domination in the surface structure, it does suggest a sensitivity of click location errors to deep structure properties of sentences. (It must be borne in mind, however, that the effects we are discussing at this point are much smaller than those which were found for the same sentences with respect to the major syntactic breaks.)

A subsequent study by Bever, Lackner and Kirk does contrast sentences in which surface structure relationships are held constant while deep structure relationships are changed. Consider the sentences:

1. (The police can't bear (the criminal to confess))

2. (The police can't bear (the criminal's) confessing)

3. (The police can't force (the criminal) to confess)

In the first type of sentence, the subject of the embedded complement clause (e.g.,"criminal") is not the logical object of the main verb (e.g., "bear"), so the deep structure sentence corresponding to "the criminal confesses" is directly reflected in the surface order of the constituents. In the second sentence there are two possible deep structure--one identical with that of (1) or one in which the object of the main verb is the gerundive itself (e.g., "confessing). In this second structural analysis, the surface order of constituents does not reflect the deep structure order. Finally, in sentence (3), the subject of the embedded complement sentence is simultaneously the object of the main verb of the matrix sentence. Thus, in this sentence type the surface order represents two distinct deep structure sentences which overlap in the surface structure.

The expectation was that the location of clicks for the noun complement cases like (3) will be much less often between the verb and its following noun than in the verb complement cases like (1). Cases like (2) were expected to be intermediate in frequency of click location errors to that boundary.

Clicks were placed either in the main verb or in the following word. Possible differences in pronunciation of the three versions was controlled by tape-splicing (cf. Garrett, Bever and Fodor, 1966; Bever, Lackner and Stolz, 1967). Six triples of the sort described above were prepared.

Sixty-five percent of subjects' responses to these sentences were location errors. Scoring was just as described for the first Bever, Lackner, Kirk experiment.

The order of the sentence types was as expected. Sentences with only the verb complement analysis attracted more clicks to the boundary between verb and following noun phrase; those sentences with only the noun complement analysis attracted the fewst; those sentences with two analyses attracted a number intermediate to the noun complement and verb complement cases. The analysis of the results was significant for $\alpha = .05$ both for sentence-by-sentence comparisons and for comparisons of the performance of subjects on the three types (e.g., significantly more subjects showed larger proportions of intrusions of clicks for the verb complement cases they heard than for the noun complement cases).

The research on the role of S-domination of surface structure constituents and of the effect of deep structure mapping onto surface structure order is still incomplete. The results from the experiments so far performed, however, suggest the following conclusions.

1. The concentrations of error responses in the various click location tasks are predictable in terms of the syntactic structure of sentences. Though such factors as prosodic features, short-term memory and rehearsal strategies, etc., cannot be claimed to have no effect on click location errors, when these factors are controlled, there remains a strong and significant effect of the syntax during the processing of the sentences.

2. A substantial part of the tendency for error responses to be concentrated at major constituent boundaries can be attributed to the subjects' tendency to treat relatively undistorted embedded sentences as perceptual units. That is, the results from the experiments on the effect of fine constituent structure and those evaluating the effect of S-domination in surface and deep structure indicate that subjects treat as perceptual units those sequences of elements which are directly traceable to a single S-node in the deep structure. Another way to put this is that <u>S</u>s tended to segment the input in terms of elements which could be treated as subject-predicate structures.

Bibliography

Abrams, K., Bever, T. and Garrett, M. "Syntactic structure modifies attention during speech perception." Submitted to Quar. J. Exp. Psyc., Nov., 1967.

Bever, T. G. "Associations to Stimulus-Response Theories of Language." In: Verbal Behavior and General Behavior Theory (Dixon and Horton, eds.), Prentice Hall, 1967 (in press).

Bever, T. G., Kirk, R. and Lackner, J. "An effect of syntactic structure on GSR response." MIT, 1967 (in preparation).

Bever, T. G., Kirk, R. and Lackner, J. "The perceptural kernelization of speech." (in preparation) MIT, 1967.

Bever, T., Lackner, J. and Stolz, L. "Transitional probability is not the general mechanism for the segmentation of speech." Submitted to Quar. J. Exp. Psyc., Nov., 1967.

Bever, T., Fodor, J. and Garrett, M. "Speech perception: some experimental results for sentences." Paper presented at the International Congress of Psychology, Moscow, Russia, 1966.

Broadbent, D. Perception and Communication. New York: Pergamon, 1958.

Chomsky, N. Syntactic Structures. The Hague, Netherlands: Mouton and Co., 1957.

Chomsky, N. "On the notion 'Rule of Grammar'." Proceedings of Symposia in Applied Mathematics, XII, American Mathematical Society, 1961.

Chomsky, N. "Current issues in linguistic theory." In: The Structure of Language (Fodor and Katz, eds.), Englewood Cliffs, N.J.: Prentice Hall, 1964.

Chomsky, N. Aspects of the Theory of Syntax, Cambridge, Mass.: MIT Press, 1965.

Fodor, J. and Bever, T. "The psychological reality of linguistic segments." Jour. Verb. Learn. Verb. Behav., 4, 1965, 414-420.

Fodor, J., Bever, T. and Garrett, M. Language Structure and Verbal Behavior. (in preparation) MIT, 1967.

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

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

Fodor, J., Garrett, M. and Bever, T. "Some syntactic determinants of sentential complexity, II: Verb structure." *Perception and Psychophysics*, (forthcoming).

Fodor, J., Jenkins, J. and Saporta, S. "Some tests on implications from Transformational Grammar." Unpublished paper. Center for Advanced Study, Palo Alto, Calif.

Garrett, M. Syntactic Structures and Judgments of Auditory Events. Doctoral Dissertation (Unpublished), University of Illinois, 1965.

Garrett, M. and Bever, T. "The perceptual segmentation of sentences." In: *The Structure and Psychology of Language*, (T. Bever and W. Weksel, eds.), Holt, Rinehart and Winston, (in press).

Garrett, M., Bever, T. and Fodor, J. "The active use of grammar in speech perception." *Perception and Psycholphysics*, 1, 1966, 30-32.

Garrett, M. and Fodor, J. "Psychological Theories and Linguistic Constructs." In: *Verbal Behavior and General Behavior Theory* (Dixon and Horton, eds.). Prentice Hall, 1967 (in press).

Gliedman, J., Fodor, J., and Garrett, M. "The effect of verb structure on the recognition of sentences under noise." (in preparation) MIT, 1967.

Halle, M. and Stevens, K. "Speech recognition, a model and a program for research." In: *The Structure of Language* (Fodor and Katz, eds.) New Jersey: Prentice Hall, 1964.

Katz, J. and Fodor, J. "The structure of a semantic theory." *Language*, *39*:170-210.

Katz, J. and Postal, P. *An Integrated Theory of Linguistic Descriptions*. Cambridge, Mass.: MIT Press, 1964.

Ladefoged, P. and Broadbent, D. "Perception of sequence in auditory events." *Q. Jour. Exp. Psyc.*, *12*, 1960, 162-170.

Marshall, J. "Behavioral Concomitants of Linguistic Complexity." MRC Psycholinguistics Research Unit, Institute of Exp. Psyc., Oxford Univ.

Matthews, G. H. Analysis by Synthesis of Sentences in Natural Language. In: *Proceedings of International Congress on Machine Translation and Applied Language Analysis*. London: H.M.S.O., 1962.

McMahon, L. Grammatical Analysis as Part of Understanding a Sentence. Doctoral Dissertation (Unpublished), Harvard University, 1963.

Mehler, J. How Some Sentences Are Remembered. Doctoral Dissertation (Unpublished), Harvard University, 1964.

Miller, G., Galanter, E., and Pribram, K. Plans and the Structure of Behavior. New York: Holt-Dryden, 1960.

Miller, G. and McKean, K. "A Chronometric Study of Some Relations Between Sentences." Q. Jour. Exp. Psychol., 16, 1964, 297-308.

Ross, J. R. Constraints on Variables in Syntax. Doctoral Dissertation (Unpublished), Massachusetts Institute of Technology, 1967.

Savin, H. and Perchonock, E. Grammatical Structures and the Immediate Recall of English Sentences. Jour. Verb. Learn. Verb. Behav., 4, 1964, 348-353.

Sapir, E. "The Psychological Reality of Phonemes." In: eds. Mandelbaum, D. Selected Writings of Edward Sapir. Berkeley, University of California Press, 1958.

Schlesinger, I. M. Sentence structure and the reading process. Doctoral dissertation (Unpublished), The Hebrew University of Jerusalem, 1966.

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*

Massachusetts Institute of Technology
Department of Psychology
Cambridge, Massachusetts

2a. REPORT SECURITY CLASSIFICATION

UNCLASSIFIED

2b. GROUP

N/A

3. REPORT TITLE

THE DEVELOPMENT OF PSYCHOLOGICAL MODELS FOR SPEECH RECOGNITION

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Final Report

5. AUTHOR(S) *(First name, middle initial, last name)*

J. A. Fodor
T. G. Bever
M. Garrett

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
January 1968	85	36

8a. CONTRACT OR GRANT NO.

AF 19(628)-5705

b. PROJECT NO.

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

ESD-TR-67-633

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

Deputy for Command Systems, Electronic
Systems Division, AFSC, USAF,
L G Hanscom Field, Bedford, Mass. 01730

13. ABSTRACT

This report has two principal sections. In Section I the theoretical issues relating to the development of syntax recognition routines based on psychological models of human speech recognition are discussed and the relevant psychological literature reviewed. The research reported in this section deals with attempts to relate various syntactic variables to measures of the perceptual complexity of sentences. The results of the research indicate that (1) analysis by synthesis routines are probably not appropriate as models of the system employed by human speakers for speech recognition, (2) the complexity of sentences is not related in any direct way to the number of operations required by a grammar to produce them, and (3) both the lexical structure of verbs and the relation of certain other lexical formatives to transformational operations of the grammar are significantly related to the ease of understanding sentences. Section II of the report deals with research concerning the perceptual segmentation strategies employed by speakers in their analysis of continuous speech. The report details the development of a particular investigative technique and its application to the determination of segmentation strategies. The research reported indicates that speech signals are initially analyzed into segments that approximate the clause. The experiments reviewed esplore the relationship between the constituent structure of sentences and this clause-like segmentation.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
1. Generative grammar						
2. Sentence complexity						
3. Sentence recognition						
4. Syntax						
5. Transformations						
6. Verb structure						