Chapter 9

SENTENCE COMPREHENSION: A CASE STUDY IN THE RELATION OF KNOWLEDGE AND PERCEPTION

J. M. CARROLL AND T. G. BEVER

The perceiver . . . responds to the stimuli according to some organization that he imposes upon them.

—G. A. MILLER, 1951, p. 79.

I. INTRODUCTION

For several decades psychologists have been investigating the problem of speech perception as defined by G. A. Miller in the early 1950s. Miller and his colleagues showed that listening to speech involves simultaneous reference to a number of linguistic "levels," including the levels of the "sentence" and "meaning." A sample demonstration of this is the fact that a sequence of random words is harder to hear than the same words ordered into a sentence (Miller, 1951). The most striking aspect of this phenomenon is the fact that the words themselves seem to be acoustically clearer as a function of their role in the sentence. This suggested that the sentential structure is providing information that can guide the acoustic analysis of the signal. The question was, *How does this kind of interaction in speech perception occur?*

That question is with us today, still unanswered. However, it has stimulated a considerable body of research, which has advanced our understanding of how to answer it. There are several principles that guide virtually every kind of research program on speech perception:

1. The amount of information that can be stored in memory in a single form is extremely limited.
2. The structure of language specifies a number of forms in which speech information can be simultaneously represented. These forms are structurally ordered in a hierarchy of "levels."
3. The way to study the perception of a representation at a particular linguistic level is to vary the stimulus at another level and observe the changes in reports about the representation.

The linguistic levels applicable to an utterance include (at least) the following (ordered roughly according to a traditional and intuitive notion of increasing "abstractness"):

Psycho-acoustics: A physiological specification of the speech waveform as transmitted by the ear.

Phones: A segmentation of the signal into discerte categories of speech sounds, using universal features

Phonemes: A segmentation in terms of the categories of sounds and features used in the particular language.

Syllables: A segmentation of the sequence in terms of canonical acoustic forms that could be uttered in isolation (e.g., CVC, CV, VC).

Morphemes: A segmentation in terms of the (memorized) meaning-bearing units of the language.

Words: The minimal meaning-bearing units that can be uttered in isolation in universal categories (e.g., noun, verb).

Phrases: Hierarchical groupings of adjacent words into universal categories (e.g., noun phrase, verb phrase).

Clauses: Groupings of adjacent phrases in terms of the canonical external relations they bear to each other (e.g., subject, verb, object).

Propositions: Groupings of the phrases (not necessarily adjacent) in terms of the canonical internal grammatical relations they bear to each other (e.g., agent, action, object, modifier).

Semantics: Interpretation (usually of a sentoid) in terms of relations to other utterances with which it is synonymous, contradictory, etc.
Speech acts: Analysis of the utterance in terms of the act it performs (e.g., promise, request, inform).
Intentions: An analysis of what the utterance indicates to be the speaker's state of mind (often in universal terms, e.g., supportive, critical, assertive, submissive).

This list is by no means complete, nor is it the case that every school of linguistics claims that all of these levels are properly included within the domain of a formal "grammar." The fact remains that we can analyze even the simplest utterance simultaneously in terms of a large number of different kinds of knowledge. For example,

(1) *Can you take out the garbage tonight dear?*

can be represented in ways varying from an acoustic specification to a (possible) description of an act by a speaker who intends to indicate that he/she feels unwell.

In a logical sense, it might appear that each level of representation *must* be present before the next more abstract level can be fully specified: for example, how can one isolate the phonetic segments without first having fully analyzed the acoustic structure?

The answer is that *some* acoustic analysis must be available, not necessarily a complete one, but one which renders those specific parameters that are critical for phonetic analysis. Furthermore, the number of choices at a "lower" level are often restricted by the representations at "higher" levels of analysis. For example, lexical and sentential information renders almost completely predictable the phonetic segment following the fragment in

(2) *Can you take out the garba--.*

Almost *any* acoustic parameter of /ǰ/ will be sufficient for its perception at that point. Thus, it is probably the case that perception at each level can facilitate perception at the other levels. The availability of such a large number of interacting kinds of simultaneous representation may effectively circumvent the limitations on any single mode of storage. Speech perception is at least not magic. But how does it in fact occur?

The studies of this process characteristically manipulate the stimulus at one level and observe the changes in the reported percept at another level. Most intense effort has been devoted to the role of acoustic parameters in syllable perception; to the role of acoustic, phonetic, and syllabic parameters in word perception; and to the role of word and phrase representations in the perception of sentences. The reasons for these choices are not accidental or arbitrary. Each of these perceptual levels is a relatively natural one for listeners to discuss in a laboratory setting—the syllable is the minimal pronounceable unit of sound, the word is the minimal memorized unit of meaning, and the sentence is the minimal unit of word combinations that can ordinarily stand alone. Despite the natural reasons to study these levels, there is no doubt that our understanding of many of the unconscious acts of perception has been obscured by the primary focus on these relatively conscious units of language.

The perception of sentence construction, for example, is only part of what listeners do when they understand speech. However, a detailed understanding of this part of the comprehension process will aid our understanding of comprehension in general. Linguistic analysis aids this process by offering a structural analysis of the sentence. Studying sentence comprehension this way is like starting the study of visual object recognition by close examination of the processes used in mapping three-dimensional percepts onto simple two-dimensional line drawings. Although this is only a part of normal visual perception, it has played an important role in the development of visual theories.

II. GENERAL BACKGROUND

The claim that the role of linguistic structure in sentence perception is a viable research problem has been assailed from two general points of view about psychological theory. The first of these, represented by Skinner (1957), Mowrer (1960), and others, derives from the position that language is describable with the mechanisms of stimulus–response (S–R) psychology. The second view, which we return to in Section III, asserts that sentence perception is exhaustively modeled by the theoretical apparatus of formal linguistics.

Much of the psycholinguistic research of the early 1960s spoke to the first of these viewpoints. The overwhelming conclusion of this work was that the sentence is a complex and abstract psychological object. In particular, syntactic variables became recognized as relevant behavioral variables (see, Epstein, 1961; Glanzer, 1962; Mandler & Mandler, 1964; Miller, 1962; Miller & Isard, 1963). The effect of these empirical findings was enhanced by a growing theoretical dissatisfaction in psychology with the S–R model of human knowledge and a commensurate development of interest in formally richer "cognitive" models (e.g., Miller, Galanter, & Pribram, 1960).

The finding that linguistic knowledge cannot be subsumed under the behavioristic learning theory model had two important consequences for psycholinguistics. First, it opened wide the door for experimental investigation of the organization of language behavior. Free of the methodological strictures and theoretical impoverishment of the S–R paradigm, psychologists could at last begin to broach these issues. The second consequence of the debunking of the S–R approach to the psychology of language was to leave psycholinguistics without a theory. Just at that time, however, a particularly exciting theoretical revolution was taking place in linguistic science. Chomsky's (1957) theory of transformational generative grammar was displacing the taxonomic approach of American descriptivism.

Chomsky rejected the taxonomic linguists' method of constructing tiered hierarchies of descriptive levels and their operationalist goal of automatic "discovery procedures" (see, Fodor, Bever, & Garrett, 1974, Chapter 2). He argued that the goal of linguistic theory was the construction of a formal grammar that would "generate" all and only the sentences of the language.

Chomsky elaborated the "transformational" model of sentence structure initially developed by Harris (1965). In this model actual sentences are generated from a set of kernel sentence structures by means of a set of transformations as blocked out in

(3) $\boxed{\text{Kernel Component}} \rightarrow \boxed{\text{Kernel Sentence Structure}} \rightarrow \boxed{\text{Transformational Rules}} \rightarrow \boxed{\text{Sentences}}$

The kernel sentence structures are close to simple declarative sentences like

(4) John past warn the boy

These structures can be transformed into other structures by optional transformations like

(5) *Passive:* $NP_1 + \text{tense} + V + NP_2 \rightarrow NP_2 + \text{tense} + \text{be} + \text{past participle} + V + \text{by} + PN_1$

or

(6) *Negative:* $\ldots V \ldots \rightarrow \ldots \text{neg} + V \ldots$.*

A final set of obligatory transformations convert the abstract structures to morphologically correct sentences:

(7) *Do-support:* $\text{tense} + \text{neg} \rightarrow \text{tense} + \text{Do} + \text{neg}$

(8) *Affix movemment:* $\text{affix} \ldots V \rightarrow V + \text{affix}$

* NP = noun phrase; V = verb; neg = not.

Each sentence is assigned a "derivation," consisting of its kernel structure and the transformations that apply to yield its surface appearance. Sample derivations are given in*

(9) a. John past warn the boy
John warn + past the boy (affix movement)
John warned the boy (morphological replacement)
b. John past warn the boy
John past neg warn the boy (negative)
John past Do neg warn the boy (Do-support)
John Do + past neg warn the boy (affix movement)
John did not warn the boy (morphological replacement)
c. John past warn the boy
The boy past be past participle warn by John (passive)
The boy be + past warn + past participle by John (affix movement)
The boy was warned by John (morphological replacement)
d. John past warn the boy
The boy past be past participle warn by John (passive)
The boy past be neg past participle warn by John (negative)
The boy be + past neg warn + past particple by John (affix movement)
The boy was not warned by John (morphological replacement)

This model of language is quite simple. Nevertheless, it captures a number of regularities about sentences. For example, it represents the fact that the sentences in

(10)	a. *John warned the boy*	Active (K)
	b. *John didn't warn the boy*	Negative (N)
	c. *The boy was warned by John*	Passive (P)
	d. *The boy wasn't warned by John*	Negative Passive (PN)

are all related by assigning them the same basic kernel structure. The model also proposes that the formal *source* for every sentence is an *abstract* kernel structure **which is not itself an actual sentence.** This claim, that sentences have abstract structures underlying them, was the focus of much controversy in linguistics and psychology; it was difficult to see how children

* The linguistic rules and derivations we present in this paper are included for purely expositional purposes, although they are similar to those proposed by Chomsky. The same is true of the three diagrams and bracketings presented in what follows. For further discussions of transformational grammar, see Chomsky (1957, 1965) and many other sources.

could "learn" that such abstract structures exist. This led Chomsky and others to postulate that the knowledge of such structures must be innate, which is a fundamental claim about human knowledge. Of course this claim is of interest only if the grammar is in some sense psychologically "real."

Indeed, Miller, Galanter, and Pribram (1960) take as pretheoretical the position that a formal grammatical derivation is literally a part of the "plan" involved in speaking. This assumption became routine in the psycholinguistic research of the early 1960s—research that was designed to establish the "psychological reality" of transformational grammar. It identified the rules of linguistic grammar directly with psychological processes.

III. THE DERIVATIONAL THEORY OF COMPLEXITY— EXPLICIT TRANSFORMATIONS ARE REAL

Consider one such study, which is paradigmatic, by Miller and McKean (1964) (a modified replication of a study by Miller, McKean, and Slobin as reported in Miller, 1962). In this study, subjects were tachistoscopically presented with a sentence. The sentence represented one of the four transformational types exemplified in (10). The subjects' first task was to change his mental representation of the presented sentence into one of the other transformational types. When he had done so, he pressed a button which tachistoscopically displated a search list of new sentences all of that type. The subject then searched the list for the transformed version of the originally presented sentence and again pressed the button when he had found it. For example, if presented with (10b) and told to apply the passive, the subject would first change it to (10d) and then press the button in order to search for (10d) in the search list.

The response times for the first task (i.e., the transforming task) suggested that transformations create additive increments of complexity in sentences. Response times for sentences that require two transformations are longer than those for sentences that require only one transformation. Response times for transforming K sentences into PN sentences were approximately equal to the sum of the response times for transforming K sentences into P sentences and K sentences into N sentences. Moreover, response times do not appear to be sensitive to the direction of the transformation. Thus, transforming K sentences into P sentences takes approximately as long as transforming P sentences into K sentences.

The Miller and McKean results appeared to confirm the view that psycholinguistic processes correspond to grammatical operations. In Chomsky's (1957) transformational grammar the P and N versions of a sentence are obtained by the application of transformational rules to the kernel sentence structure underlying the K version. What Miller and McKean found suggests that the application of transformational rules is a psychological process that can be measured. Even the distinguished status reserved for the K sentences in Chomsky's theory seems to be borne out. Sentence pairs that included a K sentence yielded especially rapid response times.

The hypothesis that the psychological complexity of a sentence is related to the number of transformations in its derivation has been called the *derivational theory of complexity,* or DTC. In its strongest form this view would hold that the greater the number of rules in a derivation, the greater the complexity of the sentence corresponding to that derivation. A series of studies followed Miller and McKean (1964) and provided results that converged on the DTC.

For example, McMahon (1963) presented subjects with the same four sentence types used by Miller and McKean. He found that when the subject is asked to decide whether a given sentence is true or false, passives and negatives require a longer time than do active affirmatives (K sentences). McMahon (see also Gough, 1965) obtained similar results in a task that required the subject to decide whether a given picture made a sentence true or false. In both studies, the passive took longer to verify than the corresponding active and the negative took longer than the corresponding affirmative. (For other relevant results and discussion see Compton, 1967; Gough, 1966; Mehler, 1963; Mehler & Miller, 1964; Savin & Perchonock, 1965.) Thus, it appeared initially to be the case that a transformational grammar is literally a psychological model. This stimulated the hope that linguistic and psychological research could become directly linked, so that new linguistic hypotheses could be subjected to direct experimental verification.

However, several qualifications are required of the initial results supporting DTC. First, there are theoretical problems with the original model. Second, there are confounding variables built into most of the DTC studies that call into question their conclusions. Finally, there are a variety of intuitive and empirical facts that are in conflict with DTC. We discuss these in order.

Recall that for Miller and McKean the translation of N sentences into P sentences is more complex (yields longer response times) than the translation (mapping) of N sentences into NP sentences. On their view, this is because for the first pair, the negative transformation must be undone and then the passive transformation applied (two transformations), whereas for the second pair, only the passive need be applied in order to complete the mapping (one transformation). However, in the grammar the passive transformation is ordered before the negative transformation. In fact, if passive applies after negative, ungrammatical strings like

(11) **The boy didn't be warned by John.*

will be generated by the grammar. Coming back to our example, *three* transformations, and not one, will be necessary for the mapping of an N version into an PN version (i.e., first undo negative, then apply passive, then apply negative). This reverses the prediction for Miller and McKean's results on DTC and hence converts their support for DTC into evidence against. It is important to note that ordering of transformations is not unique to this particular case, but rather a typical and essential property of the transformational grammars that these workers were using.

Furthermore, the sentence types studied in the DTC experiments have properties that confound the syntactic variables at issue. For example, the common result that negative sentences are more complex than their corresponding affirmatives could reasonably be ascribed to the semantic complexity of negation, and not necessarily to the addition of a transformational rule in the sentence derivation (see Wason, 1961). Besides semantic confounding, the DTC experiments confounded sentence length with syntactic structure. The result that passive sentences are more complex than their corresponding actives could be due in part to the fact that passive sentences are two words longer than corresponding active sentences.

The initial experimental demonstrations of DTC rested on a small number of types of constructions. Intuitive consideration of other types of sentences suggests that DTC is not true in general. For example,

(12) a. *The nice red wooden box fell.*

is clearly more complex transformationally than

(12) b. *The box that was nice and that was red and that was wooden fell.*

from whose structure it is derived. Yet, it is clearly *not* more complex psychologically. There is a large number of such examples.

There are also several experimental disconfirmations of the DTC. For example, Bever and Mehler (1967) found that sentences like

(13) a. *Slowly the strongman picked the girl up.*

are recalled better than sentences like

(13) b. *The strongman slowly picked up the girl.*

in a short-term memory task, even though these derivations include two additional transformations that move out the particle and adverb, respectively. This result, with sentence length and meaning controlled, directly contradicts DTC. Bever, Fodor, Garrett, and Mehler (1966) studied a variety of cases in which increased derivational complexity did not involve increased sentence length. They asked subjects to decide which one of four tones hear before a sentence was repeated immediately after the sentence. In this task no performance differences reliably correlated with transformational complexity of the stimuli.

The support for DTC rests, as we have seen, on results based almost exclusively on the negative and the passive transformations. Furthermore, the analysis of the results for these two rules is liable to a number of problems. When other rules are investigated, as in the Bever and Mehler study of particle shift and adverb postposing, predicted effects are not obtained. Jenkins, Fodor, and Saporta (1965) also failed to find the predicted results for the rule of comparative deletion. They measured tachistoscopic threshold for sentences like

(14) a. *Harry eats more than Alice eats.*
b. *Harry eats more than Alice.*
c. *Harry eats more than Alice does.*

but found that the threshold was highest for (14a), which is transformationally the least complex.

What we have seen is that an initial surprising convergence between linguistic and psycholinguistic theories proved to be specious. Nevertheless, many studies following that of Miller and McKean suggest that people *are* sensitive to transformational relations when they are explicitly marked. Experiments by Clifton, Kurcz, and Jenkins (1965), Clifton and Odom (1966), Koplin and Davis (1966), and Smith (1965) support the psychological validity of transformational "sentence families" like those in (10). In these experiments subjects characteristically rated different sentences in the same "family" as being more similar than paraphrastically related sentences not in the same transformational family.

One conclusion drawn from the psycholinguistic research on the psychological reality of grammar by 1965 was that shared underlying structures are perceived as the basis for sentence similarity. However, there was no consistent evidence that the linguistically defined transformations corresponded to psychological operations. This focused research onto the behavioral basis of the different linguistically defined structures in a derivation and away from the formal operations that manipulate these structures.

This represents an important turning point in psycholinguistic theory, although in a sense it was a reflection of a revision in linguistic theory. In the original model (see Chomsky, 1957) kernel sentences ("simple declaratives") are close to the apparent source of all sentence constructions;

they are deformed and combined by transformations. The revised linguistic model (Chomsky, 1965), the *standard theory,* emphasizes a special prominence for two levels of linguistic description: *deep structure* and *surface structure.* Deep structure represents the underlying logical form of all sentences, whereas surface structure is the constituent analysis of the sentences' manifest form. On this model, "transformational families" like (10) can be described in terms of the similarity or identity of their deep structures and need not make reference to specific transformations that have applied in their derivations.

This revision in transformational theory had special complications for the representation of complex sentences (sentences with more than one underlying proposition). In the earlier model such sentences were described as the transformational combination of separate kernel sentence structures. For example

(15) a. *The ball dropped*
b. *The ball is red*

would be generated as separate kernel structures and then combined into

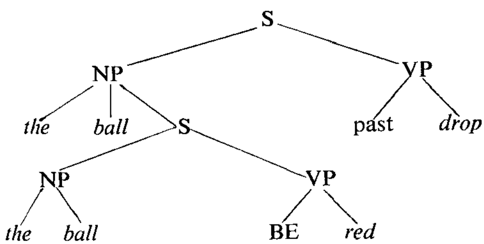
c. *The ball that is red dropped*

or

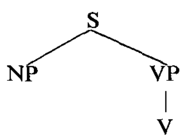
d. *The red ball dropped*

by a special transformation. In the later model, a sentence like (15c) derives from a single deep structure like

(16)

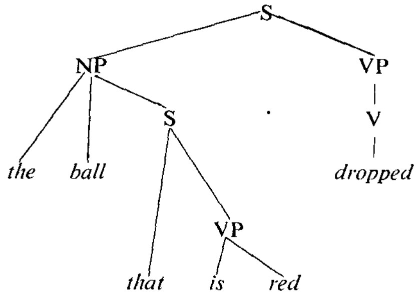


This form of analysis articulates two notions of "clause." At the deep-structure level, the clause structure represents the basic grammatical relations in each proposition (actor, action, object, modifier). That is, the actor of a predicate is the NP that is dominated by the S that dominates the VP that dominates the predicate (V); this complicated verbal description is simply represented by the following hierarchical tree:

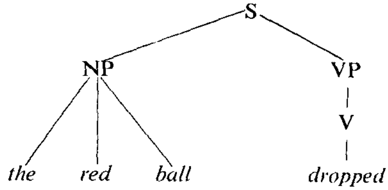


At the surface level, the clause structure represents the hierarchical groupings of adjacent phrases, as shown in

(17) a.



b.



Accordingly, the surface structure represents the hierarchical groupings of constituents that are created by the operation of the transformations on the deep-structure configurations.

The proposal that speaker–hearers are sensitive to both deep and surface structure but not necessarily to particular rules that relate these levels in formal linguistic theory is intuitively appealing. Deep structure in linguistic theory is the level at which all information relevant to semantic interpretation is represented (e.g., grammatical relations like "subject of," "predicate," and "object,"). If listeners are to understand sentences, it seems reasonable to assume that the computation of some such structure is involved. Indeed, at some level we understand a grammatical difference between sentences with superficially homonymous structures like

(18) a. *John is eager to please.*
b. *John is easy to please.*

Only in (18a) is *John* the logical (i.e., deep) subject of the verb. Similarly, we understand that

(19) *Murdering Cossacks can be horrifying*

expresses two different logical propositions even though this ambiguity cannot be represented in its surface structure.

IV. SURFACE AND DEEP STRUCTURES ARE PSYCHOLOGICAL OBJECTS

A. Surface Structures

A number of experiments have demonstrated the psychological pertinence of surface clause representations of sentences. Epstein (1961) found that subjects could learn strings of grammatically inflected nonsense syllables more easily than they could learn strings of uninflected nonsense syllables—despite the greater length of the inflected strings. Glanzer (1962) showed that mixed sequences of nonsense words and English words are learned more efficaciously when they are interpretable as surface constituent frames. For example, (20a) and (20c) are more easily learned than (20b) and (20d).

(20)
a. YIG FOOD
b. KEX AND
c. WOJ AND KEX
d. YIG FOOD SEB

Anglin and Miller (1968) found that printed prose passages are more easily learned when the end of each line corresponds to a surface constituent boundary than when it does not. Suci (1967) and R. E. Johnson (1970) report similar effects of facilitation of learning with surface constituent segmentation.

N. F. Johnson (1965) studied the patterns of errors made during the learning of lists of sentences. He devised a measure called *transitional error probability* (TEP). The value of TEP for a pair of words is the probability that the second word is recalled incorrectly, given that the first word is recalled correctly. Johnson found that the TEP for pairs of words adjacent in surface structure increases sharply for the first word in a major constituent (e.g., noun phrase, verb phrase), and that a significant increase obtains for the initial words of minor constituents as well. Johnson's central finding is that surface clause structure can define the psychological units for learning sentences in lists.

A similar paradigm was used by Suci, Ammon, and Gamlin (1967). In this study subjects were presented with a "probe" word immediately following a sentence. The probe had occurred in the sentence, and the subject's task was to supply the next word from his recollection of the sentence. Greatest response latencies were obtained for probe words that were final words in a major constituent [e.g., "red" in (15c)]. Similar results have been obtained by Kennedy and Wilkes (1968) and Ammon (1968).

Like Johnson's work, they suggest that the surface constituent is a psychologically real basis for organizing a sentence.

Mehler and Carey (1967) report surface structure effects with another paradigm. They presented subjects with sets of 11 sentences in noise. The first 10 sentences in the list had identical constituent structures, but in one condition, the last sentence had a different structure [contrast (21a) with (21b)].

(21) a. *They are performing monkeys.*
b. *They are fixing benches.*

In this condition, subjects did not understand the final sentence in the list as well as in the condition in which it had a surface structure consonant with the preceding 10 sentences. Mehler and Carey attributed this result to a perceptual "set" for surface structure configuration induced by the first 10 sentences.

Fodor and Bever (1965) and Garrett (1965) studied surface structure effects using a paradigm derived from work by Ladefoged and Broadbent (1960). Ladefoged and Broadbent had noticed that subjects were less accurate in reporting the location of short bursts of noise (of about 30 msec) superimposed on recorded sentences than they were in reporting the location of these interruptions in recorded digit strings. They concluded that the perceptual units for digit strings were smaller than the units of perceptual analysis used for sentences. Fodor and Bever found that errors in click location were, in fact, sensitive to constituent structure. They found a greater location error for clicks not objectively located at constituent boundaries than for clicks objectively located in boundaries. Moreover, they found a reliable tendency for mislocated clicks to be placed in constituent boundaries.

Garrett, Bever, and Fodor (1966) controlled for acoustic artifact by constructing stimulus pair sentences like

(22) a. *In her hope of marrying Anna is surely impractical.*
b. *Your constant hope of marrying Anna is surely impractical.*

By cross-recording, the acoustic properties of the final six words are held constant while their syntactic configuration is manipulated. In (22a) the constituent structure break occurs between *marrying* and *Anna,* whereas in (22b) it occurs between *Anna* and *is*. Abrams and Bever (1969) constructed stimulus sentences by splicing together words that had been recorded in random word lists. These materials had no cues of stress, intonation, or pausing to reveal constituent structure. In both experiments the pattern of results from Fodor and Bever's study was replicated. (See also Berry, 1970; Holmes & Forster, 1970; Seitz, 1972; Scholes, 1969).

Stewart and Gough (1967) used a task in which a two-word probe was presented to the subject directly after a stimulus sentence. They measured the reaction time for a subject to decide whether or not the two-words had occurred in the sentence. Using materials like those used by Garrett *et al.*, they contrasted response latencies for cases in which the probe sequence "straddled" a constituent boundary with those cases for which it did not. For example, in (22b) *marrying Anna* does not straddle a boundary but in (22a) it does. Stewart and Gough found that reaction times for the straddling trials are greater than those for probes belonging to a single constituent. This again is consistent with the view that surface constituents (in this case, the surface clause) have a perceptual integrity.

As a final example of the reality of surface structure, consider a study by Levelt (1970). Levelt presented subjects with a sentence and then with a list of word pairs derived from the sentence. The subject rated the "relatedness" of each pair and the data were scaled by a "hierarchical clustering" analysis. Scaling structures obtained by using this procedure are highly similar to the surface analyses of the sentences further supporting the "reality" of surface structures.

B. Deep Structures

A number of studies also support the psychological pertinance of deep structure relations. For example, in some cases of Levelt's experiment, the scaling structures do not yield well-formed hierarchies. In these cases, Levelt found that subjects' judgments of relatedness were sensitive to deep grammatical relations. Thus, the chaotic scaling structure for

(23) a. *Carla takes the book and goes to school.*

becomes very orderly when imposed on

(23) b. *Carla takes the book and Carla goes to school.*

But, in fact, (23b) is similar to the deep structure organization of (23a) (i.e., before the linguistic rule of conjunction reduction that deletes the second *Carla*).

The same point holds for Stewart and Gough's results. Some of the reaction time asymmetries they found were not readily explainable in terms of surface constituent structure. In a subsequent study, Walker, Gough and Wall (1968) contrasted the effects of surface structure adjacency with those of deep structure relation. They used the two-word probe technique with sentences like

(24) *The scouts the Indians saw killed a buffalo.*

Of the four probe types: *scouts saw, scouts killed, Indians saw,* and *Indians killed;* they found the fastest response latency for *scouts killed.* In general, probe words that are related in deep structure yielded faster response times than those which are not.

Davidson (1969) used the TEP measure to contrast full passives like

(25) a. *Gloves were made by tailors.*

with truncated passives like

(25) b. *Gloves were made by hand.*

He found no difference in TEP for words adjacent in surface structure (e.g., *made tailors* versus *made hands*). However, he found that the TEP for pairs related in deep structure like *tailors made* was lower than TEP scores for pairs not so related, like *hand made,* even though they had comparable surface structure positions. In sentences like (25a), the noun of the by-phrase is the deep structure subject of the verb; whereas in sentences like (25b), it is not.

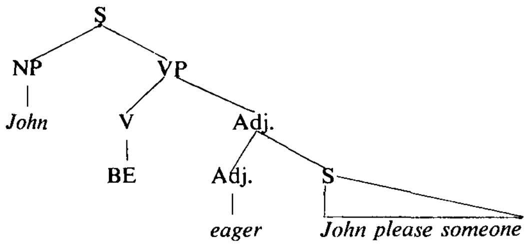
A series of studies by Blumenthal, Boakes, and Wanner investigated the psychological reality of deep structure relations in prompted recall tasks. Blumenthal (1967) used sentences similar to those used by Davidson, finding that *tailors* was a significantly better prompt word than *hand.* This result, consistent with Davidson's, suggests the conclusion that deep subjects, like *tailors,* are more salient in the stored psychological representation of a sentence than are nouns in adverbial phrases, like *hand.* Blumenthal attributed this salience to the dominating position (or height) of subject nouns in linguistic deep structures.

To test the generality of the result, Blumenthal and Boakes (1967) tested other sentence structure types, like

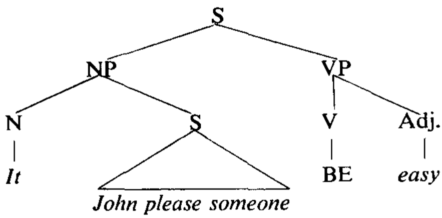
(26) a. *John is eager to please.*
b. *John is easy to please.*

in the same prompted recall paradigm. *John* was found to be a better prompt for sentences like (26a) than for sentences like (26b). This result is compatible with two interpretations, however. In (26a), *John* is the deep subject, whereas in (26b), it is not. Therefore, *John* is higher in the deep structure of (26a) than it is in the deep structure of (26b). But *John* also occurs twice in the deep structure of (26a) but only once in the deep structure of (26b). The two deep structures are glossed as

(27) a. *John is eager for John to please someone.*



b. *It is easy for John to please someone.*



The Blumenthal and Boakes result cannot distinguish these two explanations of the differential prompt efficacy of *John*.

Wanner (1968) again using prompted recall sought to distinguish these two accounts. He used sentences like

(28) a. *The governor asked the detective to prevent drinking.*
b. *The governor asked the detective to cease drinking.*

and contrasted the prompt efficacy of *detective*. For both sentences, the noun has a position of equal height; however, in (28b) the noun occurs three times in the deep structure, while in (28a) it appears only twice. The two deep structures may be glossed as

(29) a. [*The governor ask the detective* [*the detective prevent* [*someone drink*]$_S$]$_S$]$_S$
b. [*The governor ask the detective* [*the detective cease* [*the detective drink*]$_S$]$_S$]$_S$

Wanner's results are persuasive support for the claim that the structures of the grammatical formalism are effective in the speaker–hearer's psychological representation of linguistic information. Wanner concluded that the number of deep structure propositions, or sentoids, that an item occurs in determines its value as a prompt for the matrix sentence. This, in turn, is consonant with the view that the psychological units recruited in a paradigm like the one used by Blumenthal, Boakes, and Wanner are deep structure sentoids. We will return to this position in the next section, when we discuss clausal analysis.

V. CLAUSAL ANALYSIS—THE SEARCH FOR SEMANTIC RELATIONS

Up to this point, we have reviewed evidence supporting the "psychological reality" of deep and surface structures but generally disconfirming the view that the explicit formal transformations of linguistic theory are psychologically real, at least for perception. The latter conclusion is somewhat equivocal, since any test of the derivational theory of complexity (DTC) must be intimately tied to some specific assumptions about grammar. In order to refute DTC decisively, one would need to demonstrate that no empirically adequate grammar could be formulated that would make correct predictions vis-à-vis the psychological reality of formal linguistic operations (see Watt, 1970).

We could attempt to recast the grammatical rules in order to accommodate them in a way consistent with differential behavioral complexity of sentences. However, even if such an attempt could succeed in specific cases, there is no known principled way to accommodate the full set of rules to the full set of complexity facts. The available evidence makes the prospects for such a grammar very dubious. Thus, each accommodation would be descriptive, but not explanatory.

In fact, there is reason to believe that no attempt to find such a principle could succeed. Transformations, in linguistic theory, operate on hierarchical phrase structures, not on strings of words or acoustic waveforms. The derivational theory of complexity must assume that the surface phrase marker of a sentence is specified in order for it to be detransformed into the corresponding deep structure. However the derivational theory of complexity gives no procedure for this. If such a procedure were part of the entire recognition process, it would presumably be a substantial contributor to complexity variation. Thus, DTC would be less subject to direct test: Even if the number of transformations *does* evoke a corresponding number of perceptual operations, their effect might be entirely masked by the perceptual processes involved in constructing a surface constituent structure.

Moreover, even given a surface phrase structure, how is the deep structure to be recovered? A random trial-and-error detransformation process is implausible: If it were true, it would also be inconsistent with a rigorous DTC. First, if the derivation of actual sentences involves a fraction of the number of transformations stated in the grammar (about 10^2), then a trial-and-error procedure would swamp the complexity difference between a sentence derivation with 10 transformations and one with 9. It cannot be replied that the transformations all apply virtually effortlessly and simultaneously, because then DTC itself could not make any empirical predictions.

The surface structure could be detransformed to the deep structure by way of an ordered set of inverse transformations. Indeed, one such computer model has been developed (see, Petrick, 1965). This model, however, is not directly consistent with DTC, because there is no particular number of inverse transformations corresponding to every linguistic transformation. Thus, the recognition of some sentences with 10 linguistic transformations could involve 10 inverse transformations, whereas the recognition of others could involve much more computation.

The considerations up to this point have established the following:

1. There is a variety of evidence for the psychological reality of linguistically defined surface and deep structure representations of sentences.
2. There is no consistent evidence for the perceptual reality of transformations as perceptual processes. Furthermore, there are formal arguments suggesting that no evidence could be found.

The question now is, given the "general reality" of surface and deep structures and the perceptual "unreality" of transformations, how do listeners in fact recover deep structures when they hear surface structures?

The simple diagram in (30) sketches the framework that has guided much research on this problem.

(30) $\left.\begin{array}{l}\text{Uttered}\\ \text{Sequence}\end{array}\right\} \rightarrow \text{Store} \rightarrow \begin{array}{l}\text{Segmentation,}\\ \text{Recoding,}\\ \text{Labeling,}\end{array} \rightarrow \begin{array}{l}\text{Deep}\\ \text{Structures}\end{array}$

Basically, deep structures are constructed by active processes. Associated with the perceptual construction of each deep structure clause is a set of simultaneous behavioral phenomena; the perceptual segmentation of the corresponding surface sequence; the recoding of the sequence into a more abstract representation; and the labeling of the grammatical relations. These distinguishable behavioral phenomena should not be viewed as essentially independent from each other—rather, each reflects the active processes of constructing deep structures out of surface sequences. However, we shall discuss the behavioral phenomena separately for clarity of exposition.

A. Segmentation

One way to approach the problem of sentence perception is to ask: What is the primary unit of sentence perception? This question focuses on the clause as an obvious candidate—but the further question emerges immediately: Is it the clause at the surface or deep structure level that serves as a primary perceptual unit?

The deep structure clause, of course, is the more important of the two, in that it represents the information essential for the recovery of meaning (information about the relations between a sentence's logical propositions, grammatical relations, etc.). An interesting hypothesis, then, would be that deep structure propositions, or "sentoids," are the primary units of sentence perception. Such a view would predict that an early process in sentence perception includes grouping together surface structure items that belong to a common deep structure sentoid. It would predict perceptual effects of surface clause boundaries, since they always reflect deep structure clause boundaries [as in (15c)]. But it would also predict perceptual effects at positions in surface structure that, although they represent deep structure sentoid boundaries, are not marked by surface clause boundaries [as in (15d)]. We now review some studies that bear on this hypothesis.

Earlier, we discussed the click-location paradigm. Fodor and Bever (1965) had originally argued that all constituents in surface structure can function as perceptual units and can determine click mislocation. A reanalysis of their data (Fodor, Bever, & Garrett, 1974), however, shows that the experimental support for this claim comes from clicks placed at or near a major constituent (i.e., clause) boundary. In fact, Bever, Lackner, and Stolz (1969) and Bever, Lackner, and Kirk (1969) found no differential effects for minor constituent boundaries of various types. Moreover, the latter investigators found that much of the variance in the Fodor and Bever data is due to a tendency for clicks to be mislocated into surface positions corresponding to boundaries between deep sentoids.

These results are not consistent with the claim that it is the boundaries of surface structure clauses that are effective in perceptual segmentation. Note that although every surface structure clause boundary corresponds to a deep sentoid juncture, the converse is not always true. Bever, Lackner, and Kirk (1969) compared cases in which surface clause boundaries coincide with sentoid boundaries with those in which they do not. Compare

(31) a. *John defied Bill to resign.*

and

(31) b. *John desired Bill to resign.*

The two sentences have similar surface structures but different deep structures. In particular, the position between *desire* and *Bill* marks a sentoid boundary in (31b) whereas the corresponding position in (31a) does not mark any deep boundary. Bever *et al.* found the click effect to be reliably greater following the main verb in sentences like (31b) than in sentences like (31a). Of course, the interest of this particular result depends on the surface structure identity of sentences of the type of (31a) and (31b) but at present there are no conclusive linguistic arguments to the contrary. This result was replicated and extended by Fodor, Fodor, Garrett, and Lackner (1974).

The theoretical interest of the preceding findings rests on the methodological status of the click-location paradigm. Several investigators have suggested that the frequency of click location in the clause break is not a perceptual response to syntactic structure, but is actually due to a response bias. Reber and Anderson (1970) suggested that the mislocation patterns reflect a serial position bias for the middle of the sentence. Indeed, a serial position bias was reported in the original study by Ladefoged and Broadbent (1960). However, Bever, Lackner, and Kirk (1969) found no exaggeration of click mislocation into the major clause boundary when it coincided with the medial sentence position. Moreover, serial position has typically been controlled in the click location experiments we are reviewing.

A second possibility is that clause breaks are perceptually prepotent primarily because the redundancy of successive words increases (i.e., TEP decreases) in approaching a clause boundary but then decreases markedly in crossing the boundary. On this view, the clause acts as a perceptual unit because it is an island of relatively high transitional probability surrounded by boundaries of relatively low transitional probability. Bever, Lackner, and Stolz (1969) manipulated information redundancy, however, and showed that click mislocations do not occur into points of low transitional probability, when those positions do not coincide with a clause break.

A third possible alternative to the hypothesis that the click effect is truly perceptual response to syntactic structure is the claim that the location errors are due to a response bias in conjunction with recall error. Fodor and Bever (1965) had subjects recall stimulus sentences and then mark the perceived location of clicks. In fact, Ladefoged (1967) reported that when subjects are asked to guess the position of a subliminal (but in fact nonexistent) click, they manifest some predilection for locating it in a major constituent break. The issue, of course, is not whether the click location effect is a response bias for sentoidal boundaries, for this is just what all the results show. The question is whether it measures a truly perceptual effect.

A variety of data suggests that the effect cannot be attributed to recall error. First, when a recognition task is substituted for the recall task, there is no attenuation of the effect of the clause boundary on click location even though memory load is presumably decreased. Garrett (1965) and Holmes (1970) gave the subject a written script of the sentences after their auditory presentation and found the usual pattern of results. The conclusive study (to date) was carried out by Bever, Hurtig, and Handle (1976). They presented brief tones (rather than clicks) in a background of white noise during sentences. Subjects had to be prepared to write out the sentence, although on critical trials a script was given them after hearing the stimulus. On the preprepared script, the subject was given a response "window" four syllables wide, as shown in

(32) a. *After the dry summer of /that year all crops/ were destroyed.*
b. *After the dry summer /of that year all/ crops were destroyed.*
c. *After the dry summer of that /year all crops were/ destroyed.*

The subject was told that the tone occurred somewhere in the window. The response window always straddled the location of the tone. [Hence, the three windows in (32) each correspond to a different tone location.] However, in certain cases *there was no tone present* (the purpose of embedding tones in noise was to make their apparent absence on certain trials less jarring to the subjects). Responses within the different response windows [as shown in (32a–c)] when tones were absent, gave a picture of the perception-free response bias subjects use. Bever *et al.* found that the response bias pattern does not account for the tendency to locate tones in the clause boundary. In brief, the click-location response patterns are shaped after listening and before responding a few seconds later (even when no writing is necessary).

Studies using other paradigms have also tended to converge on the conclusion that deep structure sentoids are perceptual units. For example, Forster (1970) and Forster and Ryder (1971) used the Rapid Serial Visual Presentation (RSVP) technique in which sentences are displaced word by word on successive frames of a film strip. In this technique, error rate can be manipulated by altering the rate of presentation. Forster finds that when length and semantic plausibility are controlled, sentences with only one deep sentoid obtain lower error rates than sentences having two sentoids in their deep structures. This follows from the view that sentences with more perceptual units ought to be more complicated than sentences with fewer of them.

B. Semantic Recoding

Although the studies we have just reviewed suggest that the clausal structure of sentences is effective in their perceptual segmentation, they do not address the behavioral role of clausal units during perception. A number of other studies shed some light on this question. Abrams and Bever (1969) found that reaction times to clicks located objectively at the ends of clauses are much slower than reaction times to clicks located immediately after a clause boundary. Bever (1968) varied click amplitude and found that, just before a break, detectability of clicks is lowest (replicated in Bever & Hurtig, 1975). These results support the conclusion that processing load in the perceptual system is high at the ends of clauses. Apparently, listeners are relatively preoccupied at the end of a clause. We have discussed the evidence that this preoccupation is reflected in perceptual segmentation. We turn now to evidence that at clause boundaries the acoustic information is recoded and integrated into more abstract levels of representation.

Jarvella (1971, 1973) had subjects listen to short stories that were interrupted without warning. The subject's task was to recall as much as he could of the story. Jarvella finds that rote recall is best for words in the clause immediately preceding the interruption, but that it falls off sharply for words in earlier clauses. There was no reliable serial position effect for words within clauses. However, when Jarvella tested for recall of meaning by questioning subjects on the content of the stories, he found virtually no difference as a function of clause position. These results, like those of Bever and of Abrams and Bever, are consonant with the view that sentences are perceptually integrated clause by clause. When a full clause has been parsed, it is recoded into some more abstract semantic store—its meaning is available but the exact form is not (see also Sachs, 1967).

A further result consistent with this analysis is reported by Caplan (1972). Caplan measured response time to decide whether a probe word presented after a stimulus sentence did or did not occur in that sentence. His sentences were constructed in the manner of Garrett *et al.* (1966) so that the clause in which the probe word occurred could be manipulated independently of intonational cues or distance from the end of the sentence. Response latencies for sentences with the probe word in the first clause were greater than those for sentences with the probe word in the second clause. This is the expected result on the view that as clauses are completed they are represented in a more abstract store. The same finding is reported by Shedletsky (1974), who systematically varied whether or not the first clause is a main clause or a subordinate clause.

With this viewpoint in mind, we now turn to a consideration of the case of ambiguous sentences. This topic has been of particular interest in psycholinguistics, since listeners must contend with many potentially ambiguous sentences every day. We will suggest that the clausal analysis view may be able to reconcile an apparent paradox in the literature about ambiguities. On the face of it, there seem to be two differentiable hypotheses about how ambiguous sentences are perceptually processed. First, one could hold that given a number of possible interpretations the perceptual system computes them all in parallel. On the other hand, it could be the case that only one interpretation is selected serially at a time for perceptual computation. For the case of unresolved ambiguities, the former model predicts that the perceptual complexity of a sentence ought to increase with the number of possible interpretations. In the latter model, the serial process model, ambiguity should only increase perceptual complexity when the interpretation selected first is incompatible with subsequent information. There are separate experiments supporting each of these models. We shall suggest that both models are correct, but at different times: During a clause listeners entertain a number of possible meanings in parallel; after the clause is completed only one meaning is recoded, and hence, only one meaning is retained.

Consider three studies that support the parallel model. MacKay (1966) presented subjects with sentence fragments and measured the time required to supply a completion. He found that ambiguous fragments took longer to complete than unambiguous fragments, even when the subject did not notice the ambiguity. This result supports the parallel view insofar as response time is assumed to be due to response interference.

Foss (1970) measured reaction times to prespecified phonemes in recorded sentence strings. He found slower reaction times following ambiguous material. Since "phoneme-monitor" times are taken as indications of computational load (see Foss, 1969; Hakes, 1971), this result is also amenable to the explanation that multiple readings are computed in parallel for ambiguous sequences.

Finally, Lackner and Garrett (1973) presented subjects simultaneously with an ambiguous sentence in one ear and disambiguating context in the other. Subjects were instructed to attend only to the ear to which the ambiguous material was presented. Following the stimulus presentation, subjects provided paraphrases of the attended material. Although both signals were easily understood in isolation, subjects were unable to report anything about the unattended (disambiguating) signal, except that it was speech. In particular, they did not notice the ambiguities of the attended signal. Nevertheless, Lackner and Garrett found a strong tendency for the paraphrase results to covary with the unattended biasing context material. The result seems to support the claim that both interpretations of an ambiguity are computed and available; otherwise, there would be no way that the biasing context could have an effect.

In contrast with the preceding findings, there are experiments that support the hypothesis that ambiguities are treated serially. Foss, Bever, and Silver (1968) presented subjects with sentences followed immediately by pictures. They measured the response time for the subject to decide whether the picture was true or false vis-à-vis the sentence. Foss *et al.* compared latencies for ambiguous and unambiguous sentences, finding that response time was greater for ambiguous sentences only if the picture favored the less frequent interpretation of the ambiguity. They concluded that subjects assign only one immediate interpretation to an ambiguous sentence (i.e., the more frequent). If that meaning is not appropriate to the situation then they may go back and reanalyze the sentence for another meaning. Supporting results have been reported by Cairns (1970).

Carey, Mehler, and Bever (1970) used the same paradigm as Foss *et al.* but pre-set subjects to expect a sentence with one of two phrase structures [see Mehler and Carey (1967), and sentences (21a) and (21b)]. Subjects then heard a sentence that was ambiguous between the two phrase structure interpretations. Subjects who later reported not having noticed the ambiguity and who heard the ambiguous sentence as having the phrase structure for which they had been "set" gave response times that were indistinguishable from those of subjects dealing with analogous but objectively unambiguous sentences. Subjects who noticed the ambiguity despite their "set" and heard the sentence as having the structure for which they had not been set, gave longer response times than subjects who heard the reading for which they had been set. Both of these results are incompatible with the parallel model. They suggest that the induced set can suppress the processing of one interpretation of an ambiguity.

There is a possibility that the conflicting results just reviewed can be reconciled by the clausal analysis view sketched earlier. On this view we might expect that many interpretations of an ambiguity are computed within a clause, but that when the clause is completed and recoded and dismissed from short-term storage, a decision is made as to the univocal semantic interpretation of the clause. This position has two sources of support.

First, it reconciles virtually all of the conflicts reviewed so far. A typical characteristic of the studies that support the parallel model is that they test for processing effects *within* a clause. On the other hand, the studies that tend to show support for the serial model typically measure an effect *after* the end of the sentence. In addition, Bever (1968) reanalyzed MacKay's data and showed that the effect of ambiguity reported by MacKay obtains only in sentence fragments that are incomplete clauses.

A second piece of evidence is a result by Bever, Garrett, and Hurtig (1973). Bever *et al.* used the sentence completion paradigm of MacKay but varied systematically the clausal properties of the fragments. One-third of their test fragments were incomplete clauses, one-third were clauses and one-third were completed clauses plus part of a following (unambiguous) clause. Only for the incomplete clause fragments were completion times reliably slower than for unambiguous control fragments. This result agrees with Bever's reanalysis of MacKay's data and supports the view that the computational complexity of ambiguous sentences may indeed be limited by clause boundaries. (See Foss & Jenkins, 1973, for further relevant results and discussion.)

C. Perceptual Strategies

We have reviewed some evidence that supports the claim that listeners are sensitive to both deep and surface linguistic structure. We have suggested that the deep sentoid, and therefore its surface realization, the clause, is the primary perceptual unit of sentence perception. We have explored the ways in which the clause may function as a perceptual unit and examined some of the literature on the particular issue of ambiguous sentences. We now turn to a consideration of the ways in which surface elements belonging to a common sentoid might become grouped together in perceptual analysis.

The canonical form of deep sentoids is subject noun followed by verb followed optionally by object noun. This is not to say that subject–verb–object (SVO) is universally preferred in all languages—the most frequent word orders are SVO and SOV. It is an interesting fact about English that very few surface sequences of the form noun–verb–(noun) do not represent the subject–verb–(object) of a common deep sentoid. It is reasonable to suppose that listeners might have internalized this fact and that they routinely employ what may be called the canonical-sentoid strategy, interpreting surface noun–verb–(noun) sequences as subject, verb and object of a common sentoid. The difficulty of sentences like

(33) a. *The bomb rolled past the tent exploded.*
b. *The pitcher tossed the ball tossed the ball.*

may be attributed to misapplication of the canonical-sentoid strategy. In these cases, the verb of a relative clause which has been reduced by the transformational rule of WHIZ Deletion is frequently misanalyzed as being the main verb.

This proposal also finds some empirical support. The first group of studies we will review deal with doubly self-embedded structures like

(34) a. *The man the girl the boy liked hated laughed.*
b. *The editor the authors the newspapers hired like laughed.*
c. *The editor authors the newspapers hired liked laughed.*

d. *The water the fish the man caught swam in was polluted.*
e. *The reporter everyone I met trusts said Thieu will regain power someday.*

These constructions are rare in actual speech behavior. Like rare visually complex examples (e.g., the Kopferman cube, the "impossible pitchfork") they are important to consider because of what they reveal about the nature of speech perception and the relation between the grammar and behavior. It is an important achievement of a perceptual theory to explain such examples. Indeed, Miller and Chomsky (1963) recognized explicitly that center embedded constructions must be included in the set of grammatical sentences—there is no natural way of allowing single embeddings which are perfectly acceptable as in

(35) a. *The girl the boy liked hated the man.*
b. *The man the girl hated laughed.*

but blocking them from combining into unacceptable multiple embeddings. They suggested that such sequences tax short-term memory beyond its ability to store noun phrases at different phrase structure levels when the noun phrases are not yet organized in relation to verb phrases. Later research has been consistent with this general view. In particular, it has focused attention on listener's attempts to isolate appropriate sentoids with their grammatical relations.

Blumenthal (1966) analyzed the errors of paraphrase responses to sentences like (34a) and found a strong tendency for subjects to treat them as noun–verb sequences in which both the noun and the verb were conjunctions (i.e., *The man, the girl and the boy liked, hated, and laughed.*). Bever (1968) reported a significant tendency for subjects to misanalyze sentences like (34c), assigning the deep relations subject–verb–object to the sequence *The editor authors the newspapers.* This tendency was resistant to learning effects. Finally, Schlesinger (1968) found that the difficulty of forms like (34a) is mitigated when predictable semantic relations obtain between corresponding nouns and verbs, as in (34d). Sentences like (34d) are sufficient to demonstrate that the center embedded construction is syntactically grammatical, independent of Miller and Chomsky's formal point that they cannot be formally differentiated from singly embedded sentences. This is further supported by the ease of comprehending such constructions when the three initial nounphrases are from different lexical classes, as in (34e). We return later to an explanation of *why* this is true. It is sufficient here to note that such cases indicate that we must accept center embedded sentences as grammatical, both for formal and empirical reasons. Why, then, are so many of them unacceptable?

If we adopt the clausal view of sentence processing, the inherent difficulty of self-embedded forms is explicable. On this view, listeners are searching for sentoids consisting of actor–action–object clusters. In the preceding sections, we discussed the phenomena of segmentation and abstract recoding associated with assigning a sentoid analysis. It is clear that there must also be a set of mapping operations that assigns the specific internal relations that obtain between lexical items and phrases. It is only as such relations are assigned that the sequences are recoded. To put it another way, surface items cannot be dismissed from short-term memory until they have been assigned to a deep structure sentoid. Self-embedded forms, then, place a special burden on short-term storage because they present the listener with three sentoids, in which only the most embedded has its members adjacent. Accordingly, it is true that the most embedded clause in these structures is typically the one most easily understood, since that is the only sentoid whose surface elements are juxtaposed. Blumenthal's and Bever's results show that with fewer cues or with misleading cues, listeners will attempt to impose a canonical order on the material, subject–verb–object. Thus, these considerations demonstrate that there are three factors in the assignment of sentoid grammatical relations—are the assigned phrases adjacent, are they semantically constrained to be related in only one way, and are they in a canonical SVO order?

The existence of such different factors is further supported by studies contrasting reversible (36a) and irreversible (36b) passives by Slobin (1966) and by Walker, Gough, and Wall (1968).

(36) a. *The woman was chased by the dog.*
b. *The car was repaired by the mechanic.*

These studies demonstrated the result that passives are more complex than their active counterparts. However, the result did not obtain when the passives were semantically irreversible. This outcome is consistent with the interpretation that passive versions are more difficult primarily because they deform the canonical order (i.e., reverse deep subject and object). When, as in the case of irreversible passives, the deep sentoid relations are univocally determined by semantic selection, the passive–active difference is not observed.

Experiments by Wanner and Maratsos (1971) and by Walker (1969) also marshall support for the canonical-sentoid strategy. Both of these studies found that sentences with subject relative clauses, like

(37) a. *The man who hit John is a Rotarian.*

are easier than sentences with object relative clauses, like

(37) b. *The man who John hit is a Rotarian.*

Only in the case of subject relatives can the canonical-sentoid strategy apply to the embedded clause (i.e., to the relative clause). Finally, recall that Bever and Mehler (1967) found a tendency for verb particles and adverbs to be displaced from their deep structure positions in an immediate recall task. This result is consonant with the view that preanalysis favors a sequencing of material to which the canonical SVO sentoid strategy can apply.*

As is clear from the results on semantically constrained self-embedded sentences and irreversible passives, the canonical-sentoid strategy cannot be the only heuristic device available to listeners for clausal analysis. We now review some of the ways listeners can use different kinds of lexical information in perceptually processing sentences. Our main point is that listeners are extremely sensitive to subtle differences in lexical surface form when those differences are important signals to sentoid relations.

Moore (1972) presented subjects with sentence frames from which one content word had been deleted. He then displayed a word to the subject and measured the time required to decide whether or not inserting the word in the blank made the sentence grammatical. The results showed that the decision times for verbs were greater than those for nouns. This is consistent with the claim that verbs play a greater role in determining the deep structure organization of a sentence than nouns do. This hypothesis is supported by Forster (1969) who found that subjects have greater difficulty in completing left-deleted fragments like

(38) a. ———*the man come running through the trees.*

than they do with fragments like

(38) b. ———*saw the man come running through the trees.*

In this case, the verb governs the neutralization of tense in the subordinate clause.

Fodor, Garrett, and Bever (1968) tested the hypothesis that the number of deep structure configurations with which a verb is potentially compatible is related to the difficulty in understanding sentences containing the verb.

* See also, Kaplan (1972) for an attempt to formalize the perceptual strategies in the style of a general purpose computer program originally developed by Bobrow and Fraser (1969) (an augmented transition network, ATN). Kaplan argues that modeling the strategies produces precise and testable claims about the comprehension process; indeed, Wanner and Maratsos (1975) have tested certain features of the particular ATN developed by Kaplan. The weakness of this approach is that the computer program can be arranged to simulate any conceivable result: Thus it cannot reveal or explain any internal properties of perception. It can only model them. Furthermore, all the empirical tests thus far are consistent with the canonical order hypothesis (e.g., Wanner and Maratsos, discussed above).

They contrasted verbs like *slap,* which can only occur with a direct object, with verbs like *like,* which occur with direct objects and a variety of complement types. They found that subjects were less accurate in supplying immediate paraphrases of self-embedded sentences with *like* type verbs than for those with *slap* type verbs. Similarly, subjects were less successful with *like* type verbs in a sentence anagram task (i.e., ordering a random array of words into a sentence).

Other investigators have also found much the same sort of results, although not all paradigms are sensitive to verb complexity. Hakes (1971) replicated Fodor, Garrett, and Bever's paraphrase result but failed to find an effect of verb complexity in the phoneme-monitor task (listening for the first instance of a particular sound). Holmes and Forster (1972) used the RSVP technique to show that one-clause sentences with complex verbs are more difficult than one-clause sentences with simple verbs. They also found some differences consistent with the other verb complexity results for several kinds of two clause sentences. Finally, the results of Bever, Lackner, and Kirk's (1969) click study can be shown to be almost entirely due to mislocations of clicks objectively in the verb. In the sentence types studied, the verb's properties are important cues to the probable location of sentoid boundaries in the surface sequence. It is possible to argue that since the main verbs of sentences like (31b) are more complex (i.e., compatible with more deep structure geometries) than the main verbs of sentences like (31a) subjects process the two verb types differently. The assumption that subjects use these cues will explain the Bever *et al.* data.

The canonical sentoid strategy and the use of lexical heuristics cannot even be the whole story. In those cases in which a surface item is amenable to more than one deep-structure analysis, the sentence surface structure must provide information to decide between alternatives. Knowing that sentences with *expect* type main verbs like (39) below are compatible with several different complement types will not in iteself tell the listener which complement type has occurred in a given sentence. Clearly, it is to the listener's advantage to be sensitive to cues in the surface configuration which can narrow down the range of hypotheses about the deep structure of a sentence he is trying to understand.

Consider sentences like (39a–f).

(39)
a. *The woman expects that John will retreat.*
b. *The woman expects John will retreat.*
c. *The woman expects he will retreat.*
d. *The woman expects him to retreat.*
e. *The woman expects John to retreat.*
f. *The woman expects John's retreating.*

These examples illustrate several surface structure cues which identify complement types. For example, in English, *that*-complements are the only complement types that both govern *tensed* subordinate clauses and permit free deletion of the complementizer *that*. Thus, the deletion of the complementizer and the tense marking in the subordinate clause uniquely identify the complement type in (39b). Furthermore, as in (39c), *that*-complements take subject pronouns in the subjective case. In contrast, notice that infinitival (*for–to*) complements take their subject pronouns in the objective case, as in (39d). In addition, these complements have the surface item *to*, which is a cue to their deep structure identity. Finally, note that *Poss–ing* complements may be recognized by the surface configuration *'s . . . ing*.

If such surface structure cues are effective in sentence perception, one would expect that the elimination or confusion of them would tend to increase perceptual complexity. Indeed, Fodor and Garrett (1967) found that when the relative pronouns of center-embedded sentences are deleted as in

(40) a. *The cow the horse the boy rode chased likes grass.*
b. *The cow that the horse that the boy rode chased likes grass.*

subject's performance is impaired in a paraphrase task. This result has been replicated by Hakes and Cairns (1970). Hakes and Cairns and Hakes and Foss (1970) also found that phoneme-monitor times (the latency to identify a particular speech sound) are sensitive to this relative-pronoun deletion.

VI. CENTER EMBEDDED SENTENCES AGAIN

The preceding considerations suggest that listeners use perceptual strategies to map the grammatical relations among elements in surface sequences. We can schematize such strategies as shown in

(41) $\ldots X\,P\,Y\,Q\,Z \ldots$ ———— $P(R_i)Q$.
That is, "In a surface sequence in which P precedes Q, assign the grammatical relation R_i to $P \cdots Q$. (Note that X, Y, Z may be null, or may be specified, depending on the strategy.)

For example, the tendency to listen for the "canonical order" of grammatical relations can be formalized as in

(42) NP V ———— NP (subject of) V.
V NP ———— V (object of) NP.

It is important at the present state of our understanding of these strategies not to take their formalization literally—it is not at all clear what the correct formalization is. Nevertheless, even an approximate formalization can express the fact that any NVN sequence (not already assigned other relations) is susceptible to being interpreted as SVO. This can lead to perceptual illusion in cases where the NVN in fact is not an SVO sequence. Consider

(43) **The horse raced past the barn fell.*

When presented with this sequence in isolation, most listeners reinterpret it either as a "sloppy speech" version of (44a), rarely (44b),

(44) a. *The horse raced past the barn and fell.*
b. *The horse raced past. The barn fell.*

or as ungrammatical nonsense. If, however, (43) is preceded by either

(45) a. *The horse that was raced past the barn fell.*

or

(45) b. *The horse ridden past the barn fell.*

then it is immediately understood and accepted as grammatical. The difficulty of (43) is explained by the strategies in (42); since the past participle; *raced,* is homonomous with the simple past, listeners first treat the initial portion of the sequence as an SVO. it is striking that even when the word *fell* indicates that this analysis is incorrect, most listeners cannot break the SVO set and find the correct interpretation.

Now that the theory of speech perception has been somewhat more developed, we can examine its application to center-embedded sentences in detail. The action of another perceptual mapping rule does partially explain the initial complexity of double-embeddings. Consider the perceptual mapping rule that assigns the functions of underlying subject and object to noun phrases with relative clauses

(46) In a surface sequence ". . . NP_1 NP_2 ($\neq$ who) (NP^*) V ($\neq$ Be) . . ." NP_1 is the object of a sentoid of which NP_2 is the subject.

This rule capitalizes on the fact that in clause initial (or postverbal) position two adjacent noun phrases followed by a verb other than BE (with other noun phrases optionally intervening) are uniquely related such that the first noun phrase is the object (direct or indirect) of an underlying clause of which the second noun phrase is the subject. For example, the initial sequence of noun phrases in (34a) would be assigned the appropriate relations by (46) as shown

(47)	*The girl*	*the boy*	*liked. . . .*	
	Object_i	Subject_i		

and

(48)	*The man*	*the girl*	*the boy*	*liked. . . .*
	Object_i	Subject_i		
		Object_j	Subject_j	

Clearly (46) must apply twice to double-embeddings to mark the middle noun phrase as both a subject and an object. This double marking by the same perceptual rule lies at the heart of the difficulty of center-embedded sentences.

There is a logical restriction on the utilization of any conceptual dimension, given in the following principle:

(49) A stimulus may not be perceived as simultaneously having two positions on the same classificatory relation.

This interacts with the double application of rules like (46) to double-embeddings. Principle (49) articulates the tautology that a stimulus cannot be perceived in two incompatible ways at the same time. For example, a noun phrase in the surface sequence cannot simultaneously be subject and object of the same nonreflexive verb. This principle, when applied according to the view of speech perception as a direct mapping of external sequences onto internal structures, will predict the difficulty of any sequence in which a phrase as a "double function" in such a mapping operation.

The general double-function hypothesis for speech perception following from (49) is this:

(50) In a sequence of constituents x, y, z, if x has an internal asymmetric relation R_i to y and y has the same internal relation to z, and x, y, and z are superficially identical in construction type, then the stimulus is relatively complex.

This is due to y's double function in the perceptual mapping rule r_i in which y is both a p and a q. (Note that $R_i \neq$ conjunction)

$$(r_i) \quad pq \longrightarrow pR_iq$$

Thus, doubly embedded sentences are complex because they involve the second noun phrase in a double function with respect to the perceptual mapping rule (46).

It should be emphasized that the difficulty implied by (50) does not refer to all cases in which a single phrase has more than one role in the deep structure. For example, in

(51) *The girl liked the sleeping boy Sam gave the sandwich to.*

the boy is simultaneously the underlying *subject* of *sleep,* the *object* of *like,* and the *indirect object* of *give.* Yet, the sentence is quite simple. Principle (50) applies only to those cases in which exactly the same perceptual mapping rule is used to assign different functions to the same phrase. That is, it is not the "double relation" of the middle noun phrase in the underlying structure which makes such sentences complex. It is, rather, the double application to it of the perceptual rule (46). In applying the rule the listener first categorizes the middle phrase as the right-hand member of a pair related by rule (46) and then must characterize it immediately as the left-hand member of the same kind of pair defined by the same rule. In the terms used in (50) the middle noun phrase of (34a) is simultaneously a p and q of the same mapping rule.

There are other examples of complex constructions explained by principle (50). Consider the relative complexity of sentences (52) and (53).

(52) ?*Maxine did not ask Harvey not to say he would not go.*

(53) *Maxine did not ask Harvey to say he would not go.*

Sentence (52) is an example of triple negation, which has often been recognized as extremely complex, if acceptable at all. Like singly embedded sentences, sentences with only two negation markers are perfectly comprehensible and acceptable [as in (53)]. Principle (50) predicts the relative difficulty of sentences with three negations. The perceptual mapping operation corresponding to the negative marks the second *not* in (52) as *simultaneously* the scope of the first negation and the operator on the third negation. Following principle (50), any sequence containing such a double perceptual function is perceptually complex.

Principle (50) also explains some examples of perceptually complex constructions that are intuitively of the same sort as the preceding examples. Consider the sentences (54) and (55).

(54) ?*They were tired of discussing* **considering** *producing toys.*

(55) ?*They were tired of the discussion of* **the consideration** *of the production of toys.*

In each case, the sentences are extremely difficult to understand, if they are acceptable at all. As in double-embedding and triple-negation sentences, the complexity of these sentences is a function of the presence of three superficially identical phrases in which the second phrase is modified by the first phrase in the same way that the second phrase modifies the third phrase (in the underlying structure). Consider the relative perceptual ease of such constructions if only two phrases occur:

(56) *They were tired of discussing producing toys.*

(57) *They were tired of the discussion of the production of toys.*

Sentences like those in (54) and (55) are much easier to understand if the internal relations among the three critical phrases are varied as in

(58) *They were tired of discussing ceiling producing toys.*

and

(59) *They were tired of the discussion of the evolution of the production of toys.*

The explanation is that the middle phrase no longer has a double function, since *different* perceptual mapping rules relate the first two and second two verb phrases. [Note that in (58) the midddle phrase (*ceiling*) is the underlying structure object of the following phrase (*producing*), whereas the first (*discussing*) and second phrase are not directly related. In (59) the middle phrase (*evolution*) is the action carried out by the third phrase (*production*), but the object of the first phrase (*discussion*).]

Finally, (54) and (55) become perceptually simpler if the superficial form of the critical phrases is varied, even though the internal relations are held constant. The explanation for the relative ease of

(60) *They were tired of discussing the consideration of producing toys.*

(61) *They were tired of the discussion of considering the production of toys.*

is that the middle phrase does not have a double function with respect to the *same* exact perceptual mapping rule. The superficial difference in the middle phrase allows the listener to use a different mapping rule for relating the first two phrases than for relating the second two phrases.

It is striking that such a subtle difference in superficial form can have major effects on the perceivability of sentences, although it is just what the double function principle predicts. This raises the possibility that if we change the superficial form of the noun phrases in sentences like (34a) so that they differ from each other, then the sentences should become more comprehensible. Sentence (34e) is an example of this. We pointed out that most listeners find it immediately acceptable (although it is a center-embedded construction). Each of the initial noun phrases is of a different surface lexical type and thus differentiates the operation of strategy (46) into different rules. In fact, many listeners find

(62) *The pictures a reporter everyone I met trusts took showed that arson caused the fire.*

to be acceptable though it is triply embedded;

(63) *The banker your portrait an art dealer everyone I met trusts appraised belonged to knows nothing about art.*

a quadruply embedded sentence, may be acceptable because of the differences between the lexical form of the initial noun phrases.

The acceptability of (62) and (63) is not due to the fact that certain subject–object–verb relations are semantically facilitated (e.g., *reporter . . . take pictures; art dealer . . . "appraise . . . portrait"*). In the structurally parallel sentences

(64) *The cricket the mouse the cat the dog barked at miawed at squeaked at chirped at midnight.*

(65) *The hearth the cricket the mouse the cat the dog barked at miawed at squeaked at chirped at warmed the room.*

almost every subject–verb–object relation is semantically constrained but the sequences are not acceptable. This is due to the fact that the noun phrases are not superficially distinguished in the lexical class look-up phrase of perception and a number of them carry perceptual double functions.

This analysis modifies the Miller and Chomsky proposal that center-embedding is unacceptable due to the difficulty of extracting and holding two noun phrases in memory before they are attached to a verb in a sentoid. The present argument is that the memory difficulty is itself due to the fact that the middle noun phrase must be contradictorily coded by the perceptual mapping rules. This analysis, in turn, predicts new cases and even isolates critical examples of multiple embeddings that are comprehensible. Accordingly, the general model we have proposed for speech perception accounts for existing facts and predicts new facts as well. In this sense, our ability to deal with center-embedded sequences confirms the perceptual model.

VII. WHAT ARE THE PROCESSES THAT CREATE THE PERCEPTUAL ORGANIZATION?

It should be clear that most attention has been given to a *structural* definition of the sequences that are treated as points of perceptual segmentation, recoding, and functional labeling. The ultimate question for a psychological model concerns the processes that bring these organizational phenomena about. The view that we have outlined is roughly the following: During clauses, listeners use local lexical and semantic information to develop hypotheses about the possible sentoid mappings onto elementary grammatical relations. When a sufficiently clear mapping is possible, listeners recode the acoustic material into a more abstract form. This empties the immediate memory store for further speech material. In brief, listeners are continually playing "catch-up" with the speech stimulus—recoding the speech to a more abstract level whenever possible.

Several questions are suggested by this model. First, how will it treat degraded clauses (i.e., sequences whose surface realization does not explicitly include a full set of grammatical relations but whose underlying syntactic structure corresponds to a sentoid)? For example, the sentence

(66) a. *After leaving, John regretted his harsh words.*

corresponds to the deep structure bracketing sketched in

(66) b. [[*After John leave*]$_S$ *John regret John's harsh word*]$_S$
c. *After John left, he regretted his harsh words*

After leaving does not explicitly mention the subject noun. The question is whether the initial sequence in (66a) will stimulate segmentation in the way that the clausal processing model predicts the initial sequence in (66c) will do so.

A second question concerns surface structure properties. We have reviewed evidence that surface structure cues can condition perceptual complexity (pages 328–329). Fodor, Fodor, Garrett, and Lackner (1974) tested pairs like

(67) a. *This time of year anyone/who's rich enough takes a vacation.*
b. *This time of year anyone/rich enough takes a vacation*

in a click location paradigm. They found that boundary effects (boundaries are indicated by a slash) were reliably stronger in the unreduced (67a) version. This result tends to complicate the sentence perception picture. It does not disconfirm the psychological reality of linguistic deep structure but it does eliminate the hypothesis that *only* deep structure variables are relevant in sentence perception. The linguistic difference between (67a) and (67b) is the reduction of a surface relative clause in (67a) to a postnominal modifier in (67b). Since (67a) and (67b) differ in surface structure but not in deep structure, the result suggests that although deep structure sentoid junctures determine potential points of perceptual segmentation, whether such a juncture is used as a perceptual unit boundary may depend on surface structure factors.

Another aspect of surface structure that may act as a segmentation variable is the clarity with which a clause is marked as an independent unit. Just as the grammatical relations within a clause must be integrated into a whole, so the various segmentation units must be integrated into a logical structure corresponding to the entire sentence. The listener's processing must be sensitive to the range of grammatical devices which mark subordination relations between clauses. Thus, in

(68) a. **After Max ate steak,** *he became very tall.*
b. **Max's eating steak** *made him very tall.*
c. **Max eats more steak** *than he really ought to.*
d. **Max eats a lot of steak** *and it made him beefy.*

surface markers signal the listener that the initial clause will have to be integrated with the following main clause: In (68a) the adverb *after,* in (68b) the sequence *'s . . . -ing* and in (68c) the comparative *more.* Subordinated clauses, like (68a), (68b), and (68c), cannot be as fully recoded as independent clauses, like (68d), because they must, at some subsequent point in processing, be integrated with their main clauses. The relative difficulty of the clause order subordinate–main versus the clause order main–subordinate seems to be consistent with these speculations (see Weksel and Bever, 1966; Clark & Clark, 1968; and Holmes, 1973). Moreover, noninitial markers [like those in (68b) and (68c)] might be less effective than initial markers [like that of (68a)] in cuing a suspended recoding strategy. Hence, the placement as well as the presence of surface structure marking may in part determine the likelihood that a given clausal sequence will stimulate segmentation.

The length (in words, syllables, phones, etc.) of clauses also seems to be a factor which could be expected to play a role in determining suitability as a segmentation unit. For example, a very short sentence-initial subordinate clause might just be parsed with its following main clause (i.e., it might not stimulate segmentation at all). Bever, Lackner, and Kirk (1969) suggest that even a proper noun (grammatically a single item) could serve as a segmentation unit if it were long enough. Sentences like

(69) a. *John told the man that the girl knew the answer.*
b. *John told the obese and slovenly man the girl knew the answer.*

are ambiguous in that they may be interpreted as having a structure with a relative clause (i.e., *The man that the girl knew was told the answer* *by John.*) or with a complement clause (i.e., *That man was told by John that the girl knew the answer.*). Perhaps when short-term storage is overtaxed as in (69b) versus (69a), listeners will try to segment perceptually earlier. If this is true, then one might expect that sentences like (69b) will tend to be interpreted on the complement-clause reading more often than sentences like (69a). The relative-clause reading requires that *that the girl knew* be grouped into the initial clause as part of the indirect object of the main verb. Tanenhaus and Carroll (1975) have found that length does play a role in determining the usability of noun phrases as segmentation units (longer noun phrases occasion stronger effects of closure).

Finally, semantic and lexical differences should be expected to play some role in determining how much linguistic material a listener will parse per segmentation unit. The grammatical difference between (70a) and (70b) is a trivial lexical alternation.

(70) a. *Who who I know wants Sally to be at the party?*
b. *Who that I know wants Sally to be at the party?*

However, (70b) seems much easier to understand (see, Bever, Carroll, & Hurtig, 1976). Similarly, recall the embedded sentences studied by Schlesinger (1968) [example (34d)]. In these sentences, difficulty seems to be related to the perspicuity of the pairings of nouns and verbs vis-à-vis semantic constraints. Finally, we have seen that the complexity and structural characteristics of individual lexical items (specifically, verbs) can play an organizing role in sentence perception (pages 327–328). It is likely that more factors must be taken into account, to say nothing of the need for better specifying those factors that have been identified.

Tanenhaus and Carroll (1975) have attempted initially to integrate factors like grammatical completeness, surface marking clause, length, and complexity in developing the notion of a *functional clause*. Using these factors, they have proposed a *hierarchy* of clause types, which scales the suitability of clausal structures as segmentation units. So, for example, they predict that main clauses are better functional clauses than are subordinates, which in turn are better than relative clauses, which are better than sentential subjects, which are better than nominalizations, etc.

Carroll and Tanenhaus (1975) argued that grammatical completeness (or in their term *functional completeness*) is the most important factor in defining the functional clause. They conducted two experiments using the "click location" technique. In their first experiment, they contrasted sentence-initial main clauses and subordinate clauses [(71a) and (71b), respectively] with headless nominalizations and noun phrases [(71c) and (71d), respectively].

(71) a. *Harry insisted on his point / but nobody was swayed.*
b. *Whenever I watch the news / I get depressed.*
c. *Running down a city street / is a good way to get killed.*
d. *The very old and worn shirt / was on sale.*

Main clauses and subordinate clauses are functionally complete in that they contain a complete and explicit set of grammatical relations. Headless nominalizations and noun phrases are functionally incomplete: Headless nominalizations lack an explicit subject noun and noun phrases lack an explicit verb.

Carroll and Tanenhaus found that brief tones superimposed on recorded sentences tend to be mislocated into or towards the clause-boundary [indicated in (71) by a slash] more often for functionally complete clauses [like (71a) and (71b)] than for functionally incomplete clauses [like (71c) and (71d)]. Following Fodor and Bever (1965) and Bever *et al.* (1969), they interpreted this as demonstrating that functionally complete clauses are better segmentation units than functionally incomplete clauses. The deep structure sentoid clausal processing model would not predict a difference between the clause types in (71) since *all* of them are derived from deep structure sentoids.

In their second experiment, Carroll and Tanenhaus contrasted relative clauses with noun phrases and headed nominalizations with headless nominalizations. Relative clauses contain a full set of grammatical relations [i.e., subject, copulative verb and predicate adjective, as in (72a)], but noun phrases with adjectives do not (72b).

(72) a. *The man who is tall / played basketball.*
b. *The tall man / played basketball.*
c. *The army's destruction of the city / annoyed the senator.*
d. *The destruction of the city / annoyed the senator.*

However, both are derived from the same deep structure configuration. Headed (72c) and headless (72d) nominalizations provide a critical test: at both surface and deep levels of syntactic description they have identical representations. They differ only in functional completeness.

Carroll and Tanenhaus found that relative clauses are better segmentation units than noun phrases with adjectives and that headed nominalizations are better than headless nominalizations. Their research with the click-location technique suggests that functional completeness and not deep and surface grammatical structure predicts the suitability of a linguistic sequence as a segmentation unit. And this, in turn, suggests that purely structural theories of segementation make the correct predictions only when they are in accord with functional theories.

VIII. CONCLUSION

The study of sentence perception can serve as a model for the study of the relationship between knowledge and behavior.

We mentioned in the introduction that treating the sentence as a perceptual object is like starting the study of three-dimensional visual perception with the perception of simple geometric forms. The same distinction arises between the formalization of the mapping between two dimensions and three and the actual visual processes that perceivers use to recognize three-dimensional objects. To take the grammatical rules as perceptual processes would be like taking the laws of perspective as direct models of psychological processes; to claim that the different levels of linguistic structure are perceived objects would be like claiming that a two-dimensional structure is analyzed independently of the three-dimensional object it represents. In both cases, the manifest stimulus is analyzed only insofar as the analysis supports a usable interpretation. Miller and his colleagues demonstrated that the sentence is a relevant unit in speech perception. Since there is an arbitrarily large number of sentences this required a theory that would describe the set of *possible* sentences: A generative linguistic grammar provided such a theory, and thereby became a hypothesis about the structure of linguistic knowledge. The grammar provides a structural description of each sentence, independent of perception. Such an independent description of what people *know* about a sentence sets the goal of what they must *discover* about a sentence when they comprehend it. This raised the question of how the grammar that describes linguistic knowledge is employed during sentence comprehension. The answer to this question offered by research has evolved in three phases. The first phase was to take the *grammatical transformations* literally as psychological processes. When this proposal broke down, the next phase was to argue that the linguistically defined *structures* are psychologically "real." Further research shows that structures are pertinent behaviorally only as a result of perceptual operations. Accordingly, our current focus must be on structures only insofar as they reveal properties of the psychological operations themselves.

References

Abrams, K., & Bever, T. G. Syntactic structure modifies attention during speech perception and recognition. *Quarterly Journal of Experimental Psychology*, 1969, **21**, 280–290.

Ammon, P. The perception of grammatical relations in sentences: A methodological exploration. *Journal of Verbal Learning and Verbal Behavior*, 1968, **7**, 869–875.

Anglin, J. M., & Miller, G. A. The role of phrase structure in the recall of meaningful verbal material. *Psychonomic Science,* 1968, **10,** 343–344.

Berry, R. A. A critical review of noise location during simultaneously presented sentences. Unpublished doctoral dissertation, University of Illinois, 1970.

Bever, T. G. A survey of some recent work in psycholinguistics. In W. J. Plath (Ed.), *Specification and utilization of a transformational grammar: Scientific report number three.* Yorktown Heights, New York: Thomas J. Watson Research Center, International Business Machines Corporation, 1968.

Bever, T. G., Carroll, J. M., & Hurtig, R. Analogy: or ungrammatical sequences that are utterable and comprehensible are the origins of new grammars in language acquisition and linguistic evolution. In T. Bever, J. Katz, & T. Langendoen (Eds.), *Integrated theories of linguistic ability.* New York: T. Y. Crowell Press, 1976.

Bever, T. G., Fodor, J. A., Garrett, M. F., & Mehler, J. Transformational operations and stimulus complexity. Unpublished, M.I.T., 1966.

Bever, T. G., Garrett, M. F., & Hurtig, R. Ambiguity increases complexity of perceptually incomplete clauses. *Memory and Cognition,* 1973, **1,** 279–286.

Bever, T. G., & Hurtig, R. Detection of a nonlinguistic stimulus is poorest at the end of a clause. *Journal of Psycholinguistic Research,* 1975 **4,** 1, 1–7.

Bever, T. G., Hurtig, R., & Handel, A. Response biases do not account for the effect of clause structure on the perception of nonlinguistic stimuli. In preparation, 1976.

Bever, T. G., Lackner, J. R., & Kirk, R. The underlying structures of sentences are the primary units of immediate speech processing. *Perception and Psychophysics,* 1969, **5,** 225–231.

Bever T. G., Lackner, J. R., & Stolz, W. Transitional probability is not a general mechanism for the segmentation of speech. *Journal of Experimental Psychology,* 1969, **79,** 387–394.

Bever, T. G., & Mehler, J. The coding hypothesis and short-term memory. *AF Technical Report,* Cambridge, Massachusetts: Harvard Center for Cognitive Studies, 1967.

Blumenthal, A. L. Observations with self-embedded sentences. *Psychonomic Science,* 1966, **6,** 453–454.

Blumenthal, A. L. Prompted recall of sentences. *Journal of Verbal Learning and Verbal Behavior,* 1967, **6,** 203–206.

Blumenthal, A. L., & Boakes, R. Prompted recall of sentences, a further study. *Journal of Verbal Learning and Verbal Behavior,* 1967, **6,** 674–676.

Bobrow, D., & Fraser, B. An augmented state transition network analysis procedure. In N. Walker & L. Norton (Eds.), *Proceedings of the International Joint Conference on Artificial Intelligence.* Washington, D.C., 1969. 557–568.

Cairns, H. S. Ambiguous sentence processing. Unpublished doctoral dissertation, The University of Texas at Austin, 1970.

Caplan, D. Clause boundaries and recognition latencies for words in sentences. *Perception and Psychophysics,* 1972, **12,** 73–76.

Carey, P. W., Mehler, J., & Bever, T. G. Judging the veracity of ambiguous sentences. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9,** 243–254.

Carroll, J. M. & Tanenhaus, M. K. Functional clauses and sentence segmentation. Unpublished manuscript, 1975.

Chomsky, N. *Syntactic structures.* The Hague: Mouton, 1957.

Chomsky, N. Review of Skinner's *Verbal Behavior. Language,* 1959, **35,** 26–58.

Chomsky, N. *Aspects of the theory of syntax.* Cambridge, Massachusetts: M.I.T. Press, 1965.

Clark, H. H., & Clark, E. V. Semantic distinctions and memory for complex sentences. *Quarterly Journal of Experimental Psychology,* 1968, **20,** 129–138.

Clifton, C., Kurcz, I., & Jenkins, J. J. Grammatical relations as determinants of sentence similarity. *Journal of Verbal Learning and Verbal Behavior,* 1965, **4,** 112–117.

Clifton, C., & Odom, P. Similarity relations among certain English sentence constructions. *Psychological Monographs,* 1966, **80**(Whole No. 613).

Compton, A. C. Aural perception of different syntactic structures and lengths. *Language and Speech,* 1967, **10,** 81–87.

Davidson, R. E. Transitional errors and deep structure differences. *Psychonomic Science,* 1969, **14,** 293.

Epstein, W. The influence of syntactical structure on learning. *American Journal of Psychology,* 1961, **74,** 80–85.

Fodor, J. A., & Bever, T. G. The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior,* 1965, **4,** 414–420.

Fodor, J. A., Bever, T. G., & Garrett, M. F. *The psychology of language.* New York: McGraw-Hill, 1974.

Fodor, J. A., Fodor, J. D., Garrett, M. F., & Lackner, J. R. Effects of surface and underlying clausal structure on click location. *Quarterly Progress Report No. 113,* Research Laboratory of Electronics, M.I.T., 1974.

Fodor, J. A., & Garrett, M. F. Competence and Performance. In J. Lyons & R. Wales (Eds.), *Psycholinguistic papers.* Edinburgh: University of Edinburgh Press, 1966.

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

Fodor, J. A., Garrett, M. F., & Bever, T. G. Some syntactic determinants of sentential complexity, II: Verb structure. *Perception and Psychophysics,* 1968, **3,** 453–461.

Forster, K. I. Visual perception of rapidly presented word sequences of varying complexity. *Perception and Psychophysics,* 1970, **8,** 215–221.

Forster, K. I. Linguistic structure and sentence production. Unpublished manuscript, 1969.

Forster, K. I., & Ryder, L. A. Perceiving the structure and meaning of sentences. *Journal of Verbal Learning and Verbal Behavior,* 1971, **10,** 285–296.

Foss, D. J. Decision processes during sentence comprehension: Effects of lexical item difficulty and position upon decision times. *Journal of Verbal Learning and Verbal Behavior,* 1969, **8,** 457–462.

Foss, D. J. Some effects of ambiguity upon sentence comprehension. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9,** 699–706.

Foss, D. J., Bever, T. G., & Silver, M. The comprehension and verification of ambiguous sentences. *Perception and Psychophysics,* 1968, **4,** 304–306.

Foss, D. J., & Jenkins, C. M. Some effects of context on the comprehension of ambiguous sentences. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 577–589.

Garrett, M. F. Syntactic structures and judgments of auditory events. Unpublished doctoral dissertation, University of Illinois, 1965.

Garrett, M. F., Bever, T. G., & Fodor, J. A. The active use of grammar in speech perception. *Perception and Psychophysics,* 1966, **1,** 30–32.

Glanzer, M. Grammatical category: A rote learning and word association analysis. *Journal of Verbal Learning and Verbal Behavior,* 1962, **1,** 31–41.

Gough, P. B. Grammatical transformations and speed of understanding. *Journal of Verbal Learning and Verbal Behavior,* 1965, **4,** 104–111.

Gough, P. B. The verification of sentences. The effects of delay of evidence and sentence length. *Journal of Verbal Learning and Verbal Behavior,* 1966, **5,** 492–496.

Hakes, D. T. Does verb structure affect sentence comprehension? *Perception and Psychophysics,* 1971, **10,** 229–232.

Hakes, D. T., & Cairns, H. S. Sentence comprehension and relative pronouns. *Perception and Psychophysics,* 1970, **8,** 5–8.

Hakes, D. T., & Foss, D. J. Decision processes during sentence comprehension: Effects of surface structure reconsidered. *Perception and Psychophysics,* 1970, **8,** 413–416.

Harris, Z. Transformation theory. *Language,* 1965, **41,** 363–401.

Holmes, V. M. Some effects of syntactic structure on sentence recognition. Unpublished doctoral dissertation, University of Melbourne, 1970.

Holmes, V. M. Order of main and subordinate clauses in sentence perception. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 285–293.

Holmes, V. M., & Forster, K. I. Detection of extraneous signals during sentence recognition. *Perception and Psychophysics,* 1970, **7,** 297–301.

Holmes, V. M., & Forster, K. I. Perceptual complexity and underlying sentence structure. *Journal of Verbal Learning and Verbal Behavior,* 1972, **11,** 148–156.

Jarvella, R. Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior,* 1971, **10,** 409–416.

Jarvella, R. Coreference and short-term memory for discourse. *Journal of Experimental Psychology,* 1973, **98,** 426–428.

Jenkins, J. J., Fodor, J. A., & Saporta, S. An introduction to psycholinguistic theory. Unpublished, 1965.

Johnson, N. F. The psychological reality of phrase structure rules. *Journal of Verbal Learning and Verbal Behavior,* 1965, **4,** 469–475.

Johnson, R. E. Recall of prose as a function of the structural importance of the linguistic units. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9,** 12–20.

Kaplan, R. Augmented transition networks as psychological models of sentence comprehension. *Artificial Intelligence,* 1972, **3,** 77–100.

Kennedy, A., & Wilkes, A. Response times at different positions with a sentence. *Quarterly Journal of Experimental Psychology,* 1968, **20,** 390–394.

Koplin, J. H., & Davis, J. Grammatical transformations and recognition memory of sentences. *Psychonomic Science,* 1966, **6,** 257–258.

Lackner, J. R., & Garrett, M. F. Resolving ambiguity: Effects of biasing context in the unattended ear. *Cognition,* 1973, **1,** 359–372.

Ladefoged, P. *Three areas of experimental phonetics.* London: Oxford University Press, 1967.

Ladefoged, P., & Broadbent, D. E. Perception of sequence in auditory events. *Quarterly Journal of Experimental Psychology,* 1960, **13,** 162–170.

Levelt, W. J. M. A scaling approach to the study of syntactic relations. In G. B. Flores d'Arcais & W. J. M. Levelt (Eds.), *Advances in psycholinguistics.* New York: American Elsevier, 1970.

MacKay, D. G. To end ambiguous sentences. *Perception and Psychophysics,* 1966, **1,** 426–436.

Mandler, G., & Mandler, J. Serial position effects in sentences. *Journal of Verbal Learning and Verbal Behavior,* 1964, **3,** 195–202.
McMahon, L. Grammatical analysis as part of understanding a sentence. Unpublished doctoral dissertation, Harvard University, 1963.
Mehler, J. Some effects of grammatical transformations on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior,* 1963, **2,** 346–351.
Mehler, J., & Carey, P. Role of surface and base structure in the perception of sentences. *Journal of Verbal Learning and Verbal Behavior,* 1967, **6,** 335–338.
Mehler, J., & Miller, G. A. Retroactive interference in the recall of simple sentences. *British Journal of Psychology,* 1964, **55,** 3, 295–301.
Miller, G. A. *Language and communication.* New York: McGraw-Hill, 1951.
Miller, G. A. Some psychological studies of grammar. *American Psychologist,* 1962, **17,** 748–762.
Miller, G. A., & Chomsky, N. A. Finitary models of language users. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology,* Vol. II. New York: Wiley, 1963.
Miller, G. A., Galanter, E., & Pribram, K. H. *Plans and the structure of behavior.* New York: Holt, 1960.
Miller, G. A., & Isard, S. Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior,* 1963, **2,** 217–228.
Miller, G. A., & McKean, K. A chronometric study of some relations between sentences. *Quarterly Journal of Experimental Psychology,* 1964, **16,** 297–308.
Moore, T. M. Speeded recognition of ungrammaticality. *Journal of Verbal Learning and Verbal Behavior,* 1972, **11,** 550–560.
Mowrer, O. H. *Learning theory and the symbolic processes.* New York: Wiley, 1960.
Petrick, S. R. A recognition procedure for transformational grammars. Unpublished doctoral dissertation, M.I.T., 1965.
Reber, A. S., & Anderson, J. R. The perception of clicks in linguistic and nonlinguistic messages. *Perception and Psychophysics,* 1970, **8,** 81–89.
Savin, H., & Perchonock, E. Grammatical structure and the immediate recall of English sentences. *Journal of Verbal Learning and Verbal Behavior,* 1965, **4,** 348–353.
Sachs, J. Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics,* 1967, **2,** 437–442.
Schlesinger, I. M. *Sentence structure and the reading process.* The Hague: Mouton, 1968.
Scholes, R. J. Click location judgments. *Quarterly Report, Department of Speech, University of Florida,* 1969, **7**(1), 33–38.
Shedletsky, L. Effects of some clause variables on memory scanning. Unpublished doctoral dissertation, University of Illinois, 1974.
Seitz, M. AER and the perception of speech. Unpublished doctoral dissertation, University of Washington, 1972.
Skinner, B. F. *Verbal behavior.* New York: Appleton, 1957.
Slobin, D. I. Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of Verbal Learing and Verbal Behavior,* 1966, **5,** 219–227.
Smith, F. Reversal of meaning as a variable in the transformation of grammatical sentences. *Journal of Verbal Learning and Verbal Behavior,* 1965, **4,** 39–43.
Stewart, C., & Gough, P. Constituent search in immediate memory for sentences. *Proceedings of the Midwestern Psychology Association,* 1967.
Suci, G. The validity of pause as an index of units in language. *Journal of Verbal Learning and Verbal Behavior,* 1967, **6,** 26–32.

Suci, G., Ammon, P., & Gamlin, P. The validity of the probe-latency technique for assessing structure in language. *Language and Speech,* 1967, **10,** 69-80.

Tanenhaus, M. K., & Carroll, J. M. The clausal processing hierarchy . . . and nouniness. In R. Grossman *et al.* (Eds.), *Papers from the parasession on functionalism.* Chicago: University of Chicago, 1975.

Walker, E. Grammatical relations in sentence memory. Unpublished doctoral dissertation, Indiana University, 1969.

Walker, E., Gough, P., & Wall, R. Grammatical relations and the search of sentences in immediate memory. *Proceedings of the Midwestern Psychological Association,* 1968.

Wanner, E. On remembering, forgetting, and understanding sentences: A study of the deep structure hypothesis. Unpublished doctoral dissertation, Harvard University, 1968. See also (same title) The Hague: Mouton, forthcoming.

Wanner, E., & Maratsos, M. An augmented transition network model of relative clause comprehension. Unpublished, 1975.

Wanner, E., & Maratsos, M. On understanding relative clauses. Unpublished paper, Harvard University, 1971.

Wason, P. C. Response to affirmative and negative binary statements. *British Journal of Psychology,* 1961, **52,** 133-142.

Watt, J. R. Two hypotheses concerning psycholinguistics. In J. R. Hayes (Ed.), *Cognition and development of language.* New York: Wiley, 1970.

Weksel, W., & Bever, T. G. Harvard Cognitive Studies Progress Report, 1966.

Woods, W. Transition network grammars for natural language analysis. *Communications of the A.C.M.,* 1970, **13,** 591-602.