

# The cognitive basis for linguistic structures<sup>1</sup>

THOMAS G. BEVER

The Rockefeller University (at present, The University of Arizona)

Originally published in R. Hayes (ed.), *Cognition and Language Development* (New York: Wiley & Sons, 1970), 279–362.

## 1 Introduction

We can all agree that the capacity to symbolize and communicate in language has powerful effects on how we think and behave, but how does the way we think affect the structure of our language? This article explores the ways in which specific properties of language structure and speech behavior reflect certain general cognitive laws.

Recent investigations of language have made an important simplifying assumption: the primary subject for linguistic description is linguistic *knowledge*, as opposed to linguistic *behavior*. This heuristic strategy has facilitated progress in the formal description of the “abstract” structures of language, but has left open the question of how such structures are learned by children and utilized by adults. Previous attempts to integrate linguistic structure and speech behavior have been efforts to find direct mappings of abstract linguistic structures onto language learning, speech perception, and speech production. Although the details of these mappings differ, the basic postulate is the same: actual speech behavior is some regular function of the abstract linguistic structure originally isolated in linguistic investigations. That is, grammar rests at the epicenter of all language behavior, with different functions of

<sup>1</sup> This research was supported by A. F. 19(628)-5705 and ARPA SD-J87. I am particularly indebted to G. A. Miller and H. L. Teuber for their logistic and intellectual support during much of the research reported in this paper.

grammatical structure accounting for different kinds of behaviors (for example, talking, listening, memorizing, and the like).

This paper begins an exploration of an alternative approach to the study of language: as a conceptual and communicative system which recruits various kinds of human behavior, but which is not exhaustively manifested in any particular form of language behavior. That is, the concept of “language” is like those of “species” or “organ” as they are used in biological science. Consider, for example, the problem of accounting for the concept “cow” by an exhaustive description of its physiological and behavioral structures (including genetic material).

Such an inventory would leave untouched the species concept itself; it would merely specify the interaction among the descriptively isolable components of the organism. The fact that there is a distinct and distinguishable bovine species that organizes these components is taken as a given before the descriptive science begins. No aspect of the cow’s physiology could be pinpointed as *the* “physical bovine essence.” This is true even for the genetic material, which at first might be taken as *the essential physical manifestation of every species*. However, the genetic material in a fertilized cow egg “is” a cow only *given* the bovine processes of uterine nourishment, embryological maturation, postnatal behavioral development, and so on; these all *assume* various interactions which the genetic material itself would be called upon to explain without a prior concept of the species.

Analogously, in the study of language, we cannot aspire to “explain” the presence and structure of language as a composite function of various descriptively isolable language behaviors, nor is there any aspect of language that can be specified as the uniquely central aspect. This is true even of linguistic grammars, which at first have been taken as *the essential structure inherent in language behavior*. Yet grammatical structure “is” the language only given the child’s intellectual environment provided by a society and the processes of physiological and cognitive development, which are the basis for language behavior.

Language has various manifestations, each of which draws on and contributes to structural constraints on the language created by every other subsystem. During the first decade of his life a child simultaneously learns all these systems—primarily, how to talk in sentences, how to understand sentences, and how to predict new sentences in his language. These and other cognitive skills can mutually influence each other as the child acquires them and as they are integrated in adult language behavior. Consequently, our first task in the study of a particular structure implicit in adult language behavior is to

ascertain its source rather than immediately assuming that it is *grammatically* relevant.

Recent linguistic “grammars” are intended to describe our knowledge of a language (as revealed by our intuitions about potential sentences) rather than how we ordinarily use that knowledge. Accordingly, in recent considerations of language learning, attention has been focused on the acquisition of linguistic competence as isolated in linguistic investigations of grammar. This view of language “structure” and its development in children, isolated from the rest of cognitive development, has been too narrow. Many aspects of adult language derive from the interaction of grammar with the child’s processes of learning and using language. Certain ostensibly grammatical structures may develop out of other behavioral systems rather than being inherent in grammar. That is, linguistic structure is itself partially determined by the learning and behavioral processes that are involved in acquiring and implementing that structure.

The example of such an interaction discussed in this paper is the effect of perceptual strategies in the child and adult on linguistic structure and the relative acceptability of certain kinds of sentences. Since adult linguistic ability includes the ability to talk in sentences, to listen to sentences, and to produce intuitions about sentences, the child must simultaneously acquire “concrete” *behavioral* systems for actually talking and listening as well as an “abstract” appreciation of linguistic structure itself. It is clear that the child tends not to learn sentence constructions that are hard to understand or say. Thus, some formally possible linguistic structures will never appear in any language because no child can use them. In this way the child’s systems for talking and listening partially determine the form of linguistic structure even as the structure is being learned and used by the child. In other words, the way we use language as we learn it can determine the manifest structure of language once we know it.

The demonstration of this thesis proceeds in several steps. First, I outline the relationship between perceptual mechanisms and conceptual structure in the child and adult exemplified in the capacity to judge relative numerosity of visual arrays. Adults can make such judgments using three different kinds of mechanisms, basic perceptual mechanisms (for example, numerosity detectors), superficial perceptual strategies, or an internalization of the structure of integers. In acquiring these capacities, children pass through three stages, each of which is successively dominated by one of the modes of adult functioning (in the order given above).

Second, I show that language is processed by adults in the same three modes, and that in acquiring language, children are dependent on each of

the three modes successively. In particular, speech comprehension in the child from two to four is relatively dependent on behavioral strategies rather than on a primitive mechanism or sophisticated structural knowledge.

Third, I suggest that the properties of the system of speech perception affect adult linguistic structure, since the young child may learn linguistic structures only after he acquires many perceptual mechanisms. Finally, I show that, in fact, many linguistic structures in adult language are clearly determined by behavioral systems like those that characterize speech perception in the young child and adult.

It would be tempting to argue that these investigations reduce the extent to which language can be viewed as peculiarly innate, since they show that certain aspects of linguistic structure are direct reflections in language of our general cognitive structure and its development. However, this would be like arguing that the physiological structure of joints and of reciprocal muscular inhibition *explains* the fact that we can walk upright. While it is true that our walking capacity depends on certain anatomical structures, it is also true that the presence of the structures themselves does not explain why we walk, nor does it explain how the anatomical structures are recruited by walking behavior. In each case in which we discover neurophysiological substrata involved in specific behavior systems, the problem is merely made more precise: how do the behavioral systems recruit and organize such neurophysiological capacities? The argument in this chapter, that language structure and behavior are the joint product of both linguistic and psychological structures, leaves us with the analogous question: how does the instinct to communicate integrate the distinct components of perception, cognition, and motor behavior into human language?

## 2 Judgments of relative numerosity

In our research we have distinguished three aspects of cognition for separate study: basic capacities, behavioral strategies, and epistemological structures (see Mehler and Bever 1968b; Bever, Mehler, and Epstein 1968). First we investigate the *basic capacities* that appear in young children without obvious specific environmental training. Consider, for example, the two-year-old's capacity to judge numerical inequalities (Mehler and Bever 1968; Bever, Mehler, and Epstein 1968), or his ability to predicate actions with verbs in speech (Bever, Mehler, Valian, Epstein, and Morissey in preparation). Second, in both perceptual and productive behavior, children and adults utilize many systems of *behavioral strategies* to shortcut the internal structure implied by the regularities in their behavior. For example, to make relative judgments of

large numbers, we may suspend our knowledge of integers and counting, and simply use the perceptual rule that an array that “looks” larger has more components; or if we hear a series of words with only one reasonable semantic connection (for example, ‘dog bite cracker’) then we suspend any further perceptual analysis of the speech signal and assume that the sentence follows the usual semantic constraints on ‘dog,’ ‘bite,’ and ‘cracker.’ Finally, as adults, we have a set of *epistemological structures*—systematic generalizations of our intuitions about the regularities in our own behavior. Consider, for example, the theoretical concept of an integer and counting which we can use in justifying our judgments of quantities; or the intuition of relative “grammaticality” that a parent uses to guide a child’s speech and a linguist depends on for the isolation of linguistically relevant data.

All three aspects of cognitive behavior are simultaneously present in adults. Consider first the case of number: when presented with an array such as that in (3) and asked to judge which row has more circles in it, we can either count the circles in each row and see which number is larger; use the generalization that a row that looks larger has more components in it; or depend on elementary “subitizing” strategies to come to an immediate decision (in those cases in which the absolute size of the stimuli is less than seven). Which of these strategies we deploy will depend on the clarity and duration of the stimuli, the penalty for an incorrect answer, and so on (cf. Kaufman et al. 1949).

- (1) o o o o o o  
       oooo

Children do not appear to have the choice among these three strategies that adults have. Our research into the acquisition of the ability to judge numerical inequalities in arrays like (1) shows that, at different periods, the child’s behavior appears to be dependent on one or another of the strategies. For example, child of 2.0 to 3.0 makes correct judgments as to which row has more circles in (2), although they do not count the number in each row. It does not appear that they are choosing on the basis of density in (2a) and (2b), since at the same age they have no tendency to pick the denser row in a situation like (3) as having “more.”

- (2) o o o o        o o o o        o o o o  
       oooooo        o o oo o o        o o o o o o  
           (a)                (b)                (c)
- (3) o o o o  
       oooo

The children are not merely choosing the longer row as having “more,” since they choose correctly in (2a) as well as (2c). Thus we must conclude that young children have a primitive capacity to appreciate the relative numerosity of small arrays even though they cannot count and do not have an (explicit) notion of integers. Six-year-olds, on the other hand, also perform correctly on judging the more numerous row in the arrays in (2), but they generally count the two rows in each case, or they perform a 1:1 matching operation to see which row has some circles left over after the matching. Both of these operations depend on a psychological notion of integers and their relation to external objects (note that it is not necessary to claim that these children understand the concept of integer in any deep mathematical sense).

Between the ages of three and five the child appears to depend on the generalization that larger arrays have “more” components. For example, their performance is worse on (2a) than that of younger and older children (Figure 1), while their tendency to choose the longer row as having more in (3) goes up from 60 percent at age two to 100 percent at age three.

It is an open question at the moment as to whether such perceptual strategies are learned as inductions across experience (since it is probably true that many things that “look” larger do have more parts) or whether the strategies appear as a result of maturation.

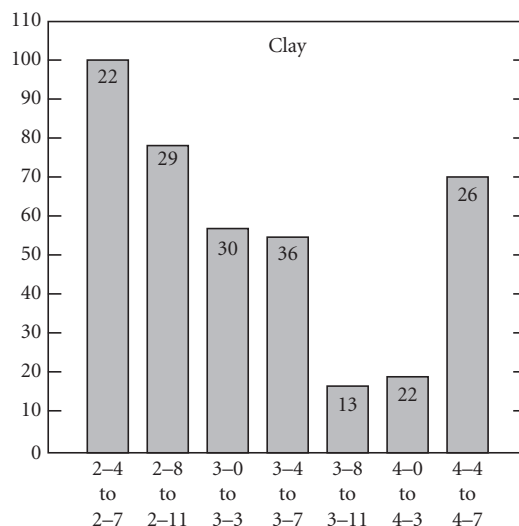


FIGURE 1. The proportion by age of responses choosing the row with more members in (2a) (numbers inside bar indicate the number of children interviewed at that age) (see text; taken from Figure 2 in Mehler and Bever 1968a)

On either interpretation the child successively displays dependence on the three types of thinking: use of basic mechanisms, behavioral generalizations, and epistemological systems. Thus the study of the child at different ages can isolate and lay bare the operation of the different types of thought processes that are often integrated inextricably in adult thought.

### 3 Three aspects of language behavior

We have been able to pursue the same sort of combined study of adult behavior and its development in children with respect to language. In this section I first explore the interaction of three linguistic systems analogous to those for the adult's perception of numerosity, and then demonstrate that the child's behavior depends primarily on each of three systems at different points in his development.

#### 3.1 *Basic linguistic capacities and epistemological structures*

In this discussion I shall largely take for granted the fact that *basic linguistic capacities* and *linguistic intuitions* are relatively clear concepts, while *perceptual strategies* of speech require more careful exposition. In the first place, it is clear that adults have the *basic capacity* to understand that words refer to objects and actions, and to understand the basic relational concepts of "actor, action, object." It is (almost as) clear that adults can make sophisticated phenomenological judgments about the sentences in their languages. Thus, adults can recognize that (4a) is a sentence and (4b) is not; that (4c–e) are ambiguous but in characteristically different ways; that, while (4f) and (4g) mean the same, they are not related in the way that (4g) and (4h) are; that the unacceptability of (4k) is caused by what it says, but the unacceptability of (4l) is caused by how it says it; and so on.

- (4) a. He kicked the can.
- b. \*Can the kicked he.
- c. The file is on the floor.
- d. He read carefully prepared papers.
- e. The missionary is ready to eat.
- f. The cockroach died from the poison.
- g. The poison killed the cockroach.
- h. The cockroach was killed by the poison.
- i. It's the poison that killed the cockroach.

- j. Did the poison kill the cockroach?
- k. ?The cockroach then wrote me an unwritten letter complaining about the poison.
- l. \*I be replying politely please for not to bug me.

All of these judgments require that the adult have the concept of language as a system independent of any actual use, that he be able to consider the potential usability of a sentence, and the relations among potential sentences. It is in this sense that the capacity to have linguistically relevant intuitions is an example of the kinds of epistemological systems we can construct to capture and extrapolate on the regularities implied by our own actual behavior.

A generative grammar attempts to provide a description of the structural basis for intuitions about sentences like the above. The basic intuition of *sentencehood* (e.g., that (4a) is part of the language and (4b) is not) is accounted for if the grammar provides a description only for those sequences that are accepted as sentences. The structural descriptions are correct insofar as they provide the basis for intuitions about the relations among sentences, such as the different types of ambiguity in (4c–e) or relations between specific classes of sentences, such as those represented by (4g) (“active”), (4h) (“passive”), (4i) (“cleft sentence”), or (4j) (“question”).

Current transformational grammars represent the traditional notion that sentences have two structural levels of description, the basic internal relations among phrases, “actor, action, object, modifier” and the explicit relations among adjacent phrases in the actual appearance of the sentence. For example, sentences (4g–j) all have the basic relations outlined in (5), while the superficial relations are obviously different. A transformational grammar represents the relations between the internal and external form of a sentence with a set of rules (“transformations”) that map abstract internal structures such as that represented in (5) onto actual sequences.

- (5) actor = the poison  
 action = kill  
 object = the cockroach



For example, a passive transformation applies to (5) to place the internal object at the front of the actual sentence and the actor at the end of the actual



sentence; a question transformation inserts a form of the auxiliary verb ‘do’ at the beginning of the sentence (see Brown and Hanlon 1970 for a detailed exemplification of the formal operation of transformations). Thus, the variety of transformations can account for the fact that a variety of actual sequences (for example, (4g–j)) can share the same internal structure (for example, (5))).

The fact that every sentence has an internal and external structure is maintained by all linguistic theories—although the theories may differ as to the role the internal structure plays within the linguistic description. Thus, talking involves actively mapping internal structures onto external sequences, and understanding others involves mapping external sequences onto internal structures.

### 3.2 *Strategies of speech perception*

In addition to basic linguistic capacities and systematic sets of structural intuitions, adult language behavior also appears to depend on behavioral inductions involved in these mapping operations. Our most intensive research has been devoted to exploring the role of these inductions in speech perception. In a recent paper, Fodor and Garrett (1966) reviewed the experimental evidence in favor of the working hypothesis that the perceptual operations that map external sequences onto internal structures are themselves directly related to the grammatical transformations specified within a grammar; that is, the view that for every linguistic transformation involved in the linguistic analysis of the relation between the internal and external structure there corresponds one perceptual “decoding” operation. Fodor and Garrett argue that this view leads to an empirical predictor that the perceptual complexity of a sentence is proportional (or at least monotonically related) to the number of transformations involved in the grammatical description of that sentence. According to this view, the passive construction is harder to understand than the active because one more rule is used in the grammatical derivation of the passive sequence than the active sequence. Fodor and Garrett review the evidence for the general claim that the relative number of rules predicts perceptual complexity and conclude that the evidence is unconvincing. The cases in which added transformations do not involve added behavioral complexity are of two types. First, transformational rules that delete internal structures do not necessarily involve added complexity; for example (7a) is not more complex than (7b); in fact, (7a) is less complex psychologically, although more complex grammatically.

- (7) a. The dog was called.
- b. The dog was called by someone.

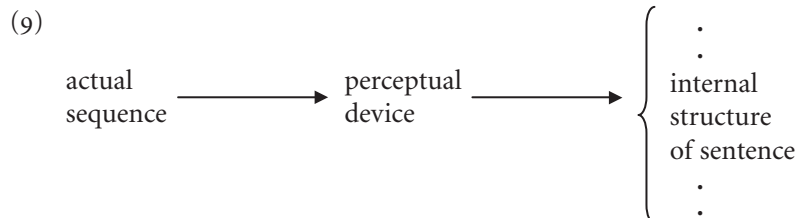
The second type of failure of added grammatical transformations to predict added psychological complexity is in certain reordering transformations. Thus (8a) is obviously not more complex to understand than (8b), (8e) is not more complex to understand than (8d) and (8e) is not more complex than (8f).

- (8) a. The small cat is on the grass mat.  
 b. The cat that is small is on the mat that is made of grass.  
 c. The operator looked the address up.  
 d. The operator looked up the address.  
 e. It amazed Bill that John left the party angrily.  
 f. That John left the party angrily amazed Bill.

Fodor and Garrett conclude from such examples and their review of the experimental literature that the relation between grammatical rules and perceptual operations is “abstract” rather than direct. This negative point has clarified many issues for us but has left open what the actual nature of the operations of speech perception is. In the following section, I outline the positive evidence that bears on the processes of speech perception and the role of perceptual strategies in mapping external sequences onto internal structures.

### 3.2.1 Segmentation strategies

Recently, a great deal of attention has been given to the “psychological reality” of the structures and rules postulated in transformational grammars. The most notable success has been to show both by experiment and appeal to intuition that the form in which sentences are understood and memorized corresponds closely to their internal syntactic structure (see Miller 1962; Mehler 1963; Mehler and Bever 1968b). Thus, any model for speech perception proposed in this tradition includes a device that isolates the internal structure corresponding to each lexical sequence.



For such a perceptual device to operate efficiently, the actual sequence of words in a speech utterance must be segmented into those subsequences that

correspond to a full sentence at the internal structure level. For example, if one hears the sequence represented phonetically in (10), one must decide that it has two distinct sentences corresponding to it at the underlying structure level, and not more or less.

(10)  $\delta\text{əboyləy ksgəlgəlgəzləyboyz}$  (that is, *the boy likes girls girls love boys*)

Failure to separate the correct basic segmentation into sequences that do correspond to underlying structure sentences could seriously degrade comprehension. For example, suppose that a listener assumed that the second instance of 'girls' above was actually a spurious repetition; then he would be faced with finding an underlying structure for the following: *The boy likes girls love boys*. The problem is that this sequence has no single underlying syntactic structure.

There is no known automatic procedure that insures the proper segmentation of actual sequences. In cases like the above, however, pronunciation often provides many cues that indicate where the segmentation into basic sentences should occur. The operation of this segmentation strategy to separate sentences in a discourse like (10) can utilize many situational, semantic, and pronunciation cues. The segmentation problem is much more complex, however, for sentences embedded within other sentences. Consider (11a), for example:

(11) a. When he left, everybody grew sad

This has two deep structure sentences, each one corresponding to one of the “clauses” in the apparent sequence: (‘When he left, everybody grew sad’). I shall represent this structure division as clauses at the surface structure level with parentheses, ( ), and the corresponding internal structure segmentation with brackets, [ ]; for example, (11b).

(11) b. ([when he left]) ([everybody grew sad])

If the wrong perceptual segmentation were attempted, then further perceptual analysis of the sentence would be impossible. For example, the listener might initially segment the first four words into a sequence bound together by an internal structure (that is, 'When he left everybody...'), but would then have two words left over ('grew sad') with no possible segmentation deriving from another internal structure sentence.

A recent series of experiments has given initial support to the existence of a perceptual strategy of isolating lexical sequences that correspond directly to underlying structure sentences (Fodor and Bever 1965; Garrett, Bever, and Fodor 1966; Bever, Fodor, and Garrett 1966). These investigations have studied the perception of nonspeech interruptions in sentences with two clauses. The basic finding is that subjects report the location of a single click in a sentence as having occurred toward the point between the clauses from its objective location. For example, Fodor and Bever found that in sentence (12), a click objectively located in ‘yesterday’ or in ‘the’ was most often reported as having occurred between those two words. Fodor and Bever argued that the systematic displacement of the click towards the point between clauses showed that the clause has relatively high psychological coherence, since it “resists” interruption by the click.

(12) because it rained yesterday the picnic will be cancelled

Several experiments have shown that this systematic effect of the syntactic segmentation is not due to any actual pauses or cues in the pronunciation of the sentence. First, Garrett, Bever, and Fodor used materials in which the exactly identical acoustic sequence was assigned different clause structures depending on what preceded. Consider the sequence ‘... eagerness to win the horse is quite immature.’ If it is preceded by ‘your...’, then the clause break immediately follows ‘horse.’ But if that sequence is preceded by ‘In its...’, then the clause break immediately follows ‘win.’ The authors cross-recorded one initial sequence or the other and tested subjects on their ability to locate clicks in the different sentences. The results showed that the clause structure assigned each sequence “attracted” the subjective location of the clicks. Abrams and Bever (1969) found similar results with sentences constructed by splicing words from a random list.

Scattered through the materials in these experiments were sentences that did not consist of two entirely separate clauses in the external structure, but which had one clause embedded within another. For example, in the sentences (13a) and (13b), there are two sentences at the level of internal structure, but they are not literally reflected in an organization into distinct uninterrupted sequence in the actual sentence.

- (13) a. ([the man ([who nobody likes]) is leaving soon])  
 b. ([nobody likes the man ([who is leaving soon])])

Nevertheless, Fodor and Bever found that the points at the extremes of the embedded clauses are as effective in attracting the subjective location of clicks as they are in sentences with two entirely separate clauses.

In certain cases in the previous experiments, two internal structure sentences corresponded to a sequence in which the division into two clauses was even less obvious in the actual structure. Consider (14a):

- (14) a. ([the reporters assigned to George] drove to the airport])  
 b. ([the reporters ([who were assigned to George]) drove to the airport])

The sequence ‘... assigned to George ...’ does not have the same distinctiveness as a clause in the surface structure of (14a) as in (14b). Nevertheless, sentences in which the surface structure does not obviously reflect the underlying structure, like (14a), were found to affect the subjective location of clicks (for example, clicks were displaced perceptually to the point following ‘George’).

These data suggest that an early step in the perceptual organization of a string of words is the isolation of those adjacent phrases in the surface order, which together could correspond to sentences at the level of internal structure (Strategy A). This perceptual strategy would generate the experimental prediction (15) for the location of clicks.

### **Strategy A**

Segment together any sequence X ... Y, in which the members could be related by primary internal structural relations, “actor action object ... modifier.”

- (15) Errors in location of clicks presented during sentences are towards those points that are external reflections of (potential) divisions between internal structure sentences.

Various further experiments indicate that (15) is correct. First, some negative experiments indicate that within-clause minor phrase structure divisions do not affect perceived click location. Bever, Lackner, and Kirk (1969) used the same technique of click location in which they systematically varied the within-clause surface phrase structure of 25 sentences. They found no tendency for the number of errors into a break to be correlated with the relative depth of that break in the surface phrase structure. Bever, Lackner, and Stolz (1969) found no difference in the effect on click location of three kinds of within-clause structures: adjective–noun (‘red ball’), verb–object (‘hit ball’), and subject–verb (‘ball hit’). Finally Bever, Fodor, and Garrett (1966) investigated the relative effectiveness of pairs of surface structure transitions which were superficially quite similar, but which differed by having

or not having an “S” node in the surface phrase structure tree. Consider the two sentences in (16):

- (16) a. ([they watched [the light turn green]])  
 b. ([they watched [the light green car]])

The relevant difference between the two structures just after the verb is the presence of an S-node in the surface structure of the first sentence. Bever et al. (1966) found that this difference of a single node had a profound effect on the pattern of errors in click placement. Together with the negative results from the other experiments, this finding supports the following initial conclusion: a relative increase in the number of surface structure nodes defines a perceptual unit only if the increase is due to an explicitly marked sentence node in the surface structure.

Although negative experimental findings are always inconclusive, these experiments do indicate at least that within-clause surface structure has far less effect on click location than breaks between clauses. Several other experiments indicate that points in the surface that correspond to underlying structure divisions do attract clicks, even in the absence of major division between apparent clauses. Bever, Lackner, and Kirk found several instances among their 25 sentences in which a within-clause phrase structure break corresponded to a division between sentences in the internal structure. These breaks did attract the subjective location of clicks. In a second experiment they found that subjects locate clicks subjectively between a verb and its complement object significantly more for “noun phrase” verbs ((17a)) than for “verb phrase” verbs ((17b)). This corresponds to the fact that in “noun phrase complements” the break following the verb corresponds uniquely to the beginning of a new internal sentence.

- (17) a. [<sub>1</sub>they desired [<sub>2</sub>the general to fight]<sub>2 1</sub>]  
 b. [<sub>1</sub>they defied [<sub>2</sub>the general]<sub>1</sub> to fight]<sub>2</sub>

These results demonstrate that Strategy A is correct, that perceptual segmentation proceeds primarily in terms of internal structure organization into sentences. However, there are various difficulties with each of the experiments we have reviewed and further work is necessary. Furthermore, it is not clear whether every internal structure division has an effect on perceptual segmentation, or whether this effect is limited to those internal structure sentences whose order is literally reflected in the surface structure (or which are marked by a sentence node in the surface structure); for example, it is not clear that both the first and second underlined sequences below will be treated as a perceptual

unit—they both derive from the same internal sentences (*my steak is rare, my steak is tender*), but only the first preserves this order in the surface structure.

- (18) a. I like *my steak rare and tender*.  
 b. I like *my rare and tender steak*.

Whatever the outcome of further experimentation, it is clear that the internal logical relations are a major determiner of perceptual segmentation in speech processing. As we hear a sentence, we organize it perceptually in terms of internal structure sentence units with subjects, verbs, objects, and modifiers.

Consider now the relation between Strategy A and a transformational grammar. Clearly Strategy A presupposes the distinction between internal and external structural relations. But there is no obvious way in which the grammatical transformations may themselves be transmuted into subcomponents of Strategy A. Rather, Strategy A is implemented on the basis of knowledge of the possibility that a particular external form class sequence could correspond to an internal structure; the possible external sequences are, of course, enumerated by the transformations, but not necessarily in a way that can be directly utilized in the process of perception. In fact, several experiments demonstrate that the application of the segmentation Strategy A is directly sensitive to knowledge of the potential internal–external relations that individual lexical items can have, as opposed to their actual deployment in a particular structure.

First, Kaplan and Kaplan showed recently (personal communication of pilot work) that subjects respond to the interruption of a sentence following an adjective like ‘hard’ ((19a)) faster than they responded to the interruption of a sentence ending with an adjective like ‘eager’ ((19b)).

- (19) a. The old general was hard ...  
 b. The old general was eager ...

I suggest that this is because immediately following ‘eager’ the listener is processing an internal structure unit, while following ‘hard’ he is not. This is presumptively due to the fact that the listener makes immediate use of his knowledge that ‘eager’ must terminate an internal structure sentence, while ‘hard’ may not terminate such a unit. Consider the examples of the possible continuations of the sentence fragments in (19) (internal structural sentence boundaries are marked with brackets).<sup>2</sup>

<sup>2</sup> There are many facts that motivate the distinction between the internal analyses of (20a) and (20c). For example, consider the fact that (20a) is related to ‘it was hard to please the old general’ while

- (20) a. The old general was [hard to please].  
 b. [The old general was hard] and [wouldn't give up].  
 c. [ [The old general was eager] to please] ].  
 d. [The old general was eager], and [wouldn't give up].

A second example of the use of knowledge of internal/external potentialities in applying Strategy A appears in Bever, Lackner, and Kirk's experiment on click location (see above). Clicks were also placed in the noun phrase following the verbs (for other groups of subjects). The numbers in sentences (21a) and (21b) below indicate the different objective click positions that were used).

- (21) a.                   1   2  
 that general desired soldiers to fight  
                                  1   2  
 b. that general defied soldiers to fight

Bever et al. found that the difference in subjective click location for the constructions in (21a) and (21b) was primarily due to responses to clicks objectively in the verb. The previous studies of click location had shown that clicks following and preceding a clause break are reported as occurring in the clause break equally often—that is, it is not the case that pre-clause break clicks are more sensitive to structural effects. Bever et al. suggest that their results show that the listener segments the sequences by using information inherent to the potential internal structure/external structure pairs which each complement verb can have. Verb phrase complement verbs characteristically must have a direct object (for example, 'soldiers') that is simultaneously the subject of a complement sentence. In contrast to this, noun phrase complement verbs can have at least the following kinds of objects:

- (22) a. direct object that is also the subject of a complement sentence  
 (... desired soldiers to fight)  
 b. a complement sentence as direct object  
 (... desired that they fight)  
 c. direct object that is also the direct object of a complement sentence  
 (if the complement sentence is in the passive)  
 (... desired them to be fought)

(20b) is not related to \*'It was eager to please the general.' That is, the abstract sentences internal to (20a) can be reflected in the actual sentences 'It was hard' and 'Somebody pleased the general,' while the sentences internal to (20c) are reflected in 'The general was eager' and 'The general pleased someone.'



- d. direct object without any complement sentence  
 (...desired them)
- e. direct object which is a nominalized complement  
 (...desired them fighting)
- f. complement sentence that has the same subject as the main verb  
 (...desired to fight)

Thus, many of the possible constructions following a noun phrase complement verb begin a new sentence at the level of internal structure, while this is never the case with a verb phrase complement verb. The fact that in structurally homonymous sentences clicks located in the verbs are located differently shows that listeners use the information contained in the verb to predict the internal structure segmentation. When a listener hears ‘defy’ he knows that a direct object must follow; when he hears ‘desire’ he knows that many constructions that follow begin a new internal structure unit. Accordingly, listeners establish segmentation following noun phrase complement verbs (by the time the listener reaches the following noun phrase there is no further difference in the potential internal structure organization of the sentence, so there is no subjective difference in click location for clicks in that location).

### 3.2.2 Relations between clauses

The need for Strategy A follows from the fact that most sentences have more than one internal clause. Not only must the different clauses be segregated from each other, but also the internal relation between the two clauses must be marked. There are two basic types of relations, coordinate ((23a)) and subordinate ((23b–d)).

- (23) a. Wars are distasteful and politicians are always in favor of peace.
- b. Wars are distasteful although *politicians are always in favor of peace*.
- c. Wars *that are distasteful* are a source of political power.
- d. Everybody wants *wars to be distasteful*.

In coordinate constructions both clauses are structurally and conceptually on the same level, while in subordinate constructions the subordinate clause is embedded within a higher, “main” clause; the main clause of such sentences expresses the primary content of the sentence, while subordinate clauses either modify that main content (as in (23b), supplement it (as in (23c), or express a presupposition underlying it (as in (23d), ‘wars can be distasteful’).

The clearest principle is that, *ceteris paribus*, the first N...V... (N) is taken to be the main clause (Strategy B).<sup>3</sup>

**Strategy B**

The first N...V... (N) ... clause (isolated by Strategy A) is the main clause, unless the verb is marked as subordinate.

In English, there are many specific morphemes that mark an initial verb as subordinate, and in such cases Strategy B does not apply. However, various facts demonstrate the relative complexity of sentences in which the first verb is a subordinate verb. Consider first the sentences with a clause as subject ((8e), (8f)). The less complex version is clearly (8e), in which the subordinate verb ('left') is not the first verb in the sequence. In a general study of subjective preference, Bever and Weksel found that subjects indicate a stylistic preference for sentences in which the subordinate clause (marked by a conjunction) follows the main clause (for example, (24a) as opposed to (24b)).

- (24) a. The dog bit the cat because the food was gone.  
b. Because the food was gone, the dog bit the cat.

Clark and Clark (1968) found that sentences in which the subordinate clause occurs first are relatively hard to memorize. These observations do not bear directly on perceptual complexity. In an independent perceptual experiment, Savin (personal communication) found that sentences in which the first verb is in a relative clause (such as (25a)) are more complex than sentences in which the first verb is the main-clause verb (as in (25b)).

In Savin's experiment, more random words are recalled when preceding sentences like (25a) than (25b).

- (25) a. The boy who likes the girl hit the man.  
b. The boy hit the man who likes the girl.

Strategy B accounts for actual mistakes made in other comprehension experiments as well as accounting for the relative complexity of sentences in which the first verb is not the main verb. For example, Blumenthal and Boakes (1961) examined the nature of errors which subjects make in attempting immediate recall of center-embedded sentences ((47a)).

His conclusion was that the main strategy that subjects use is to assume that the first three nouns are a compound subject and that the three verbs are

<sup>3</sup> Note that such strategies capture generalizations which are not necessarily always true. That is, there are exceptions to every strategy—the validity of each strategy is that it holds for most of the cases.

a compound action (as in (16b)). That is, they impose a general “subject–verb” schema onto what they hear.

- (26) a. The man the girl the boy met believed laughed.  
 b. \*The man the girl and the boy met believed and laughed.

In immediate comprehension I found that subjects cannot avoid assuming that an apparent Noun–Verb–Noun (“NVN”) sequence corresponds to a clause even when they are given explicit experience and training that this interpretation is incorrect. Subjects reported immediate paraphrases of center-embedded sentences with apparent NVN sequences (for example, italicized in (27a)).

- (27) a. *The editor authors the newspaper* hired liked laughed.  
 b. The editor the authors newspapers hired liked laughed.

Even after eight trials (with different sentences), the subjects understood the sentences with this property less well than the sentences without it (for example (27b)). That is, the “NVN” sequence in (27a) is so compelling that it may be described as a “linguistic illusion” which training cannot readily overcome.

### 3.2.3 Functional labeling strategies

Not only must listeners isolate internal structure clauses and assign their relations to each other, listeners must also assign the internal structural relations that bind the constituent phrases in each internal sentence. To do this, listeners use a set of labeling strategies that draw on semantic information, probabilistic structural features and knowledge of the potential structure underlying specific lexical items.

3.2.3.1 *Semantic strategies.* A basic strategy for functional assignment is to combine the lexical items in the most plausible way. That is, we use Strategy C whenever possible to assign the correct internal relations within a potential unit independent of syntactic structure. For example, the three lexical items ‘man,’ ‘eats,’ and ‘cookie’ are internally related, as in ‘The man eats the cookie.’

#### **Strategy C**

Constituents are functionally related internally according to semantic constraints.

If Strategy B applies independently of the actual syntactic structure, we might expect that sentences in which the semantic relations are unique are

relatively easy. Schlesinger (1966) supported this prediction by showing that center-embedded sentences are easier to comprehend when the semantic subject–verb objects are semantically constrained. That is, (28a) is easier than (28b). Clark and Clark (1968) demonstrated that if the superficial order of a complex sentence reflects the actual order of described events, then the sentence is relatively easy to retain. That is, (29a) and (29b) are easier than (29c) and (29d).

- (28) a. the question the girl the lion bit answered was complex  
       b. the lion the dog the monkey chased bit died
- (29) a. he spoke before he left  
       b. after he spoke he left  
       c. he left after he spoke  
       d. before he left he spoke

There is some evidence that the presence of unique semantic constraints allows syntactic factors to be bypassed entirely. For example, Slobin (1966) found that the passive construction is no more difficult to verify than the active sentence when the semantic relations are unique. That is, (30a) is no harder to verify than (30b), while (30c) is harder than (30d).

- (30) a. the cookie was eaten by the dog  
       b. the dog ate the cookie  
       c. the horse was followed by the cow  
       d. the cow followed the horse

This finding was extended by Turner and Rommetveit (1967). They showed that children (even in the first grade) respond correctly to a sentence like (30c) only 50 percent of the time when they have to choose a picture appropriate to the sentence. Even at age four, however, they respond correctly to semantically constrained sentences like (30a).

In an ingenious experiment, Mehler and Carey (1968) collected further evidence that subjects may process meaning simultaneously with the processing of syntactic structure. They presented subjects with appropriate and inappropriate pictures following a single sentence; the task of the subjects was to indicate whether the picture was appropriate for the sentence. Two kinds of superficially similar sentences were used, progressive tense (31a) and the participial construction (31b). They found that the latencies (that is, response times) were relatively high for inappropriate pictures, and relatively

high for the participial construction, which was assumed to have a relative complex syntactic structure.

- (31) a. they are fixing benches  
 b. they are performing monkeys

On this basis one would expect the following order of latencies (in order of increasing time to decide about the picture):

<i>Construction</i>	<i>Picture</i>	<i>Predicted</i>	<i>Observed</i>
PROGRESSIVE	appropriate	fastest	fastest
PROGRESSIVE	inappropriate	intermediate	all...
PARTICIPLE	appropriate	intermediate	the...
PARTICIPLE	inappropriate	slowest	same

However, they found that whether a sentence had the more complex syntax or the picture was inappropriate, or both, the reaction time was delayed the same amount. This suggests that subjects process meaning and structure *simultaneously* rather than in sequence; either a relatively complex structure or a complex meaning can add decision time, but since they are processed in parallel, the presence of both, a complex structure and a complex meaning does not add any extra time.

The preceding experiments demonstrate that the most likely semantic organization among a group of phrases can guide the interpretation of sentences, independently of and in parallel with perceptual processing of the syntactic structure. The semantic constraints utilized in the previous experiments were necessarily removed from any natural context, so the effects of generic probability (that is, men usually eat cookies, as opposed to the reverse; if one event precedes another we tend to talk about the first event first and in the main clause, and so on). In the actual application of language, specific contexts must provide far stronger immediate constraints and basis for prediction of the most likely meaning of a sentence independent of its form. Thus, most normal perceptual processing of sentences is probably carried out with little regard to actual sequence or structure; rather, the basic relational functions (actor–action–object–modifier) are assigned on the basis of temporary (“contingent”) and generic (“constant”) semantic probabilities. Strategy C is clearly another process in which the knowledge of linguistically defined syntactic structure is not utilized actively in actual perception. Rather, as in the case of perceptual segmentation, listeners depend heavily on their knowledge of the properties of individual words and groups of words.

3.2.3.2 *Sequential labeling (lexical ordering) strategies.* However, we are capable of understanding sentences in which there are no differential semantic probabilities. Accordingly, a complete account of the mechanisms of speech perception must also include the capacity to analyze the structural relations within a sentence from pure sequential and syntactic information. There is a primary functional labeling strategy, based on the apparent order of the lexical items in a sentence, which applies in the absence of specific semantic information.



#### *Strategy D*

Any *Noun–Verb–Noun* (NVN) sequence within a potential internal unit in the surface structure corresponds to “*actor–action–object*.”

There is some recent experimental evidence that demonstrates the presence of this labeling strategy. The primary finding is that the passive construction is more complex to comprehend than the active (in the absence of semantic constraints, see above). For example, Slobin (1966) found that children verify pictures corresponding to active sentences more quickly than pictures corresponding to passive sentences. Also McMahon (1963) (replicated by Gough 1966) found that generically true (32a, b) or false (32c, d) passives are harder to verify than the corresponding actives. Finally, Savin and Perchonock (1965) showed that the number of unrelated words that can be recalled immediately following a passive sentence is smaller than the number recalled if the test words follow an active sentence.

- (32) a. 5 precedes 13  
 b. 13 is preceded by 5  
 c. 13 precedes 5  
 d. 5 is preceded by 13

The passive construction specifically reverses the assumptions in Strategy D, which is the presumed explanation for the perceptual difficulty of the passive. Of course, the fact that the passive construction is relatively complex perceptually might also be due to its increased length and to its increased transformational complexity. However, the facts pointed out above show that transformational complexity is itself not a *general* explanation of perceptual complexity.

Only the explanation in terms of the violation of Strategy D is consistent with the following experiments. In the picture verification experiment by Mehler and Carey discussed above, it was found that the progressive form is significantly easier to understand than the superficially identical participial construction; the participial construction fails to preserve the NVN = *actor–action–object* property in its surface structure.

Recall that in both the experiment by Blumenthal and by me (examples (26) and (27) above), subjects' errors involved the assumption that the first noun or series of nouns in an apparent N...V...N... sequence is not only the grammatical subject in the external structure but also the actor in the internal structure. That is, listeners impose the "actor–action–(object)" organization on what they hear as part of the basis for segmentation of clauses.

Bever and Mehler (1968) found another example of this constructive tendency in an immediate recall experiment referred to above—a sentence they just heard to maximally conform to an "NVN" sequence. For example, in (33a) the NVN sequence is maintained, while in (33b) it is interrupted:

- (33) a. Quickly the waiter sent the order back.
- b. The waiter quickly sent back the order.

In immediate recall, 87 percent of the syntactic order errors were from stimulus sentences like (33b) to sentences like (33a) rather than the reverse.

These different experimental results converge on one common explanation: any NVN sequence in the surface structure is assumed to correspond directly to *actor–action–object* in the underlying structure. Like the semantic strategies, this process may reflect a statistical preponderance in actual utterances—although little is known about the actual frequencies of construction types at the moment.

3.2.3.5 *Particular lexical strategies.* However, there must be other strategies that supplement sensitivity to surface order and semantic constraints, since we can understand sentences which are not uniquely constrained semantically and which do not maintain the particular "canonical" order of the internal structure in the surface structure implied by Strategy D. Recent experimental work has brought out the fact that there is a heterogeneous set of strategies attached to specific lexical items, primarily function words and verbs.

It is a linguistic truism that inflectional endings and function words can represent the internal structure relations directly in the external structure of sentences. For example, the difference in functional labeling of the nouns in (34a–c) and (34d, e) depends entirely on the change in the preposition.

- (34) a. the laughing at the hunters was impolite
- b. the laughing of the hunters was impolite
- c. the laughing near the hunters was impolite
- d. John rode with Mary

- e. John rode to Mary
- f. John rode by Mary

A recent series of experiments indicates that listeners utilize the intersection of the potential internal structures that adjacent lexical items can have to guide sentence perception. Fodor and Garrett (1967) showed that center-embedded sentences with relative pronouns including ((35a)) are simpler to paraphrase than the same sentences without the relative pronouns ((35b)).

- (35) a. the boy who the man who the girl likes saw laughed
- b. the boy the man the girl likes saw laughed

They interpreted this as due to this perceptual strategy based on the use of the relative pronoun 'who':  $N_1$  *wh*  $N_2$  corresponds to  $N_2$  verb  $N_1$  in the underlying structure. However, again the following theory of perceptual complexity would make the same factual predictions: "more transformations = more psychological complexity." In sentence (35b) a pronoun deletion transformation has applied to transform it from (35a). Consequently, several additional studies have been used to increase the evidence for the argument that listeners project deep structure organization directly from the possible internal constituent structure/external structure combinations associated with the particular lexical items.

A series of experiments has shown that sentences with verbs that take complements (such as 'see') have more psychological complexity than simple transitive verbs (for example, 'hit') even when the complement verbs are used transitively. This finding supports the contention that perceptual processing is guided by the potential internal role that each lexical item *could* play. Thus, complement verbs involve more complexity even when they are employed as simple transitive verbs. Fodor, Garrett, and Bever (1968) showed that center-embedded sentences are harder to paraphrase when they contain a complement verb ((36a)) than when they have a transitive verb in the same position ((36b)) (both visual and auditory presentations were used).

- (36) a. the box the man the child saw carried was empty
- b. the box the man the child hit carried was empty

The preceding experiments involve sentences with two center embeddings that are inordinately difficult constructions in any case. Fodor, Garrett, and Bever also found that the anagram solution for sentences presented in a scrambled order is harder and less accurate if the sentence has a complement verb ((37a)) than a simple transitive verb ((37b)) (subjects were presented with a randomized array of words on cards and asked to make a sentence out of them).



- (37) a. The man whom the child saw carried the box.  
b. The man whom the child hit carried the box.

The results of these last two experiments might be due to nonstructural differences between pure transitive verbs (for example, ‘hit’) and complement verbs (for example, ‘see’) rather than the fact that complement verbs have more potential internal structures; for example, the complement verbs we used characteristically require animate subjects, while the pure transitive verbs do not require animate subjects. To test directly the hypothesis that the relevant independent variable was the number of potential internal structures a verb can have, I compared the results for complement verbs that can take several kinds of complement with those that can take only one complement.

- (38)
- a. John liked it that we slept a lot.
  - b. John liked to kick the bottle.
  - c. John liked Bill to win the race.
  - d. John saw that we slept a lot.
  - e. \*John saw to kick the bottle.
  - f. John saw Bill to be a fool.
  - g. John decided that we should sleep a lot.
  - h. John decided to kick the bottle.
  - i. \*John decided Bill to be a fool.
  - j. John remarked that we slept a lot.
  - k. \*John remarked to kick the bottle.
  - l. \*John remarked Bill to be a fool.

Each complement verb used in the two experiments was classified according to whether it takes three kinds of complements, two or one. For example, (38a–l) show that ‘like’ can have three kinds of complements, while ‘see’ and

TABLE 1A. Mean relative number of subject–verb–object triples correctly recovered per sentence for auditory presentation relative to corresponding transitive verbs (analyzed from data in Fodor, Garrett, and Bever1968: table 1)

Number of potential complements/verb	1	2
Number of cases	4	8
Relative number of SVO triples recovered	.89	.25

TABLE 1B. Relative number of errors in visual presentation (from Fodor, Garrett, and Bever 1968: table 2)

Number of potential complements/verb	1.0	2.0	3.0
Number of cases	5.0	15.0	9.0
Relative number of errors	0.2	1.2	2.0

‘decide’ have two and ‘remark’ has only one. An analysis of the data in both experiments shows that complement verbs with three possible complements are more complex than verbs with only one or two complements. (See Tables 1a, b.)

I have presented the strategies of segmentation, semantic labeling, and sequential labeling separately for purposes of exposition. It is obvious that the operation of one of the strategies can simultaneously aid the operation of another strategy. In actual perception the strategies combine simultaneously to isolate potential internal *actor–action–object... modifier* sentence units and to assign correctly the functional relations within those units. It seems reasonable to suggest that semantic cues are dominant, since structural factors (such as the reversal of the canonical actor–action–object order in passive sentences) do not affect psychological complexity when the semantic relations are unique (according to Slobin). The structural strategies, in turn, project segmentation and internal structural labeling on the basis of general sequential properties of actual sequences (for example, “NVN” in the surface structure corresponds to underlying subject–verb–object or “SVO”), or on the basis of the particular internal/external structural potential of individual lexical items.

The strategies used in speech perception to discover internal structures from external sequences are distinct both from basic linguistic capacities and from the system of intuitions that are described by a grammar. As in the case of the perceptual strategies of numerical judgments, it is not clear whether these linguistic strategies are derived by passive induction over actual experience or whether they are due to autonomous internal developments. In either case it is clear that the perceptual strategies accord closely with experience, particularly Strategies B, C, and D. Furthermore, the deployment of the knowledge of specific lexical internal/external potentialities in perception could not come about without an accumulation of experiences of the lexical potentialities. In this sense the strategies constitute a form of behavioral inductions over actual speech behavior. Thus, just as in the judgments of numerosity, speech behavior can also be described in terms of three aspects: basic capacities, behavioral inductions, and epistemological systems. Of

course, in adult speech behavior these three systems are ordinarily merged together; the presence of the behavioral strategies is brought out in experimental conditions such as those reviewed above; our primitive linguistic capacity (for example, for reference) appears directly in our speech production, while the structural intuitions relevant for linguistic analysis appear only in our conscious epistemological considerations of sentences.

### 3.3 *The development of perceptual strategies of speech in the child*

As in the case of numerical judgments, the child appears to pass through different phases in which his linguistic behavior is successively dominated by each one of the three kinds of cognitive functioning. Consider first the expression of the basic linguistic capacities at age two years. It has been traditionally recognized (cf. McCarthy 1954) that children of this age have the basic capacity to recognize (and often to say) the names of some objects and actions—that is, the capacity for *reference* is already developed, although not widely extended (of course, it is not clear whether they understand the arbitrary nature of the acoustic-referential relations or whether they believe that the names of objects and actions are intrinsic and indivisible from the objects and actions themselves).

A more abstract linguistic notion is the capacity to recognize explicitly the concept of predication as exemplified in the appreciation of the difference between subject–action and action–object relations. Recently we have tested this capacity in young children with a task in which they are requested to act out simple active sentences with toy animals ((39a, b)).

- (39) a. The cow kisses the horse.  
 b. The alligator chases the tiger.

The results of our first experiment in which each child received a total of six sentences of different kinds are presented in Figure 2.<sup>4</sup> Even the children in our youngest age group did extremely well on the simple active sentences. The significance of this simple result (obvious to any parent) is that even the very young child distinguishes the basic functional relations internal to sentences. Furthermore, he appears to be able to distinguish different syntactic construction types, as opposed simply to interpreting the first noun of *any* type of sentence as the actor and the last noun as the object. The evidence for this is

<sup>4</sup> The responses include cases in which the child refused to act out the sentence, but chose the correct alternative acted out for him by the experimenter. They do not include cases, in which the child refused to act out any sentences at all himself. The number of such cases was less than 15% at all ages, and did not materially change the results.

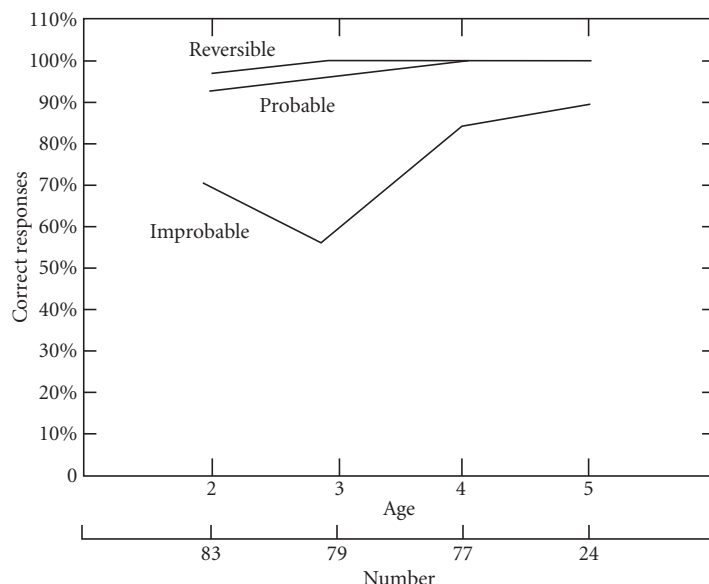


FIGURE 2. The proportion by age of correct responses to reversible active sentences like (39a), probable active sentences like (42a), and improbable active sentences like (42b); each child received only one active sentence of each kind

that while children from 2.0 to 3.0 act out simple active sentences 95 percent correctly, they also do far better than 5 percent on simple passives like (40a, b). (See Figure 3 for the performance on passives.)

- (40) a. The horse is kissed by the cow.  
 b. The tiger is chased by the alligator.

If children at this age always took the first noun as the actor, then they would systematically do as poorly on passives as they do well on actives. Since they perform almost randomly on passives, we must conclude that they can at least distinguish sentences they can understand from sentences they cannot understand. Thus, the basic linguistic capacity evidenced by the two-year-old child includes the notion of reference for objects and actions, the notion of basic functional internal relations, and at least a primitive notion of different sentence structures.

Beilin (1971) has used the emergence of the capacity to recognize the relationship between the active and the passive construction as a measure of the development of the child's capacity to produce linguistic intuitions. (I should emphasize that while the facts are due to Beilin, the interpretation is

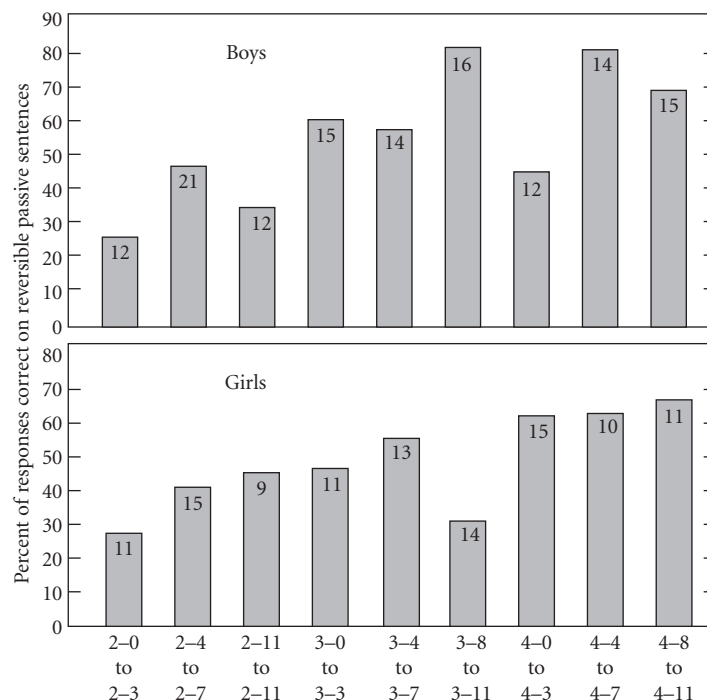


FIGURE 3. The proportion by age and sex of correct responses to reversible passive sentences (e.g., (40a)). Each child received only one reversible passive sentence. (The numbers in each bar indicate the number of children interviewed at that age. The data is broken down by four-month age groups for purposes of the discussion below.)

not necessarily his.) Beilin shows that the child does not appear capable of appreciating the regularity of the relationship between active and passive sentences until about age seven to eight, which is also the age at which the child is alleged to have developed the integer concept.

Indeed, it is support for my claim that the adult form of linguistic and numerical epistemology constitutes the same type of cognitive phenomenon that Beilin finds a correlation between the child's ability to deal correctly with numerical transformations (e.g., recognizing that changing the array in (2a) to that of (2b) doesn't change the number of circles in any of the rows) with the ability to deal correctly with sentential transformations (as exemplified by the relation of the active and the passive).

As in the case of the development of the child's capacity to make judgments of relative numerosity, the linguistic behavior between the ages of two and six

displays a period of relative dependence on perceptual generalizations. Consider first the early appearance of the basis for Strategy B, that the first N...V...(N) sequence is the main clause. In a recent study we have asked young children to act out sentences like (41a).

- (41) a. The cow that jumped walked away.  
 b. The cow jumped and walked away.

Presumably because of memory limitations, children often act out only one of the two clauses of such sentences. Which clause they act out gives us a measure of which clause they consider the most important when they hear it. Our results show that children ~~between 1 percent and 2 percent~~ who perform poorly on acting out both actions in (41b) act out only the first action (the subordinate verb) in (41a); children who do well on sentences in (41b) act out the second action (the main verb in (41a)). That is, children at the beginning of language comprehension pick the first “N...V” sequence on the most important part of a sentence—in other words, they follow Strategy B ~~completely~~; more advanced children learn to discriminate the main verb from the subordinate verb and consider the main verb to be the most important action.

Consider now the development of the basis for a semantic strategy like Strategy C, involving probabilistic constraints. We examined the development of this in the course of the same experiment outlined above by including simple active sentences that either followed (42a) or did not follow (42b) probabilistic constraints. Figure 4 shows the relative sensitivity to the semantic constraints at each age (that is, the percent correct performance on sentences like (42b) subtracted from the percent correct on sentences like (42a)—a large number indicates a high dependence on semantic constraints).

- (42) a. The mother pats the dog.  
 b. The dog pats the mother.

It is obvious why the very young child cannot make use of contextual probabilities: he does not have enough relevant experience to know what the probabilities are. For example, the young child may know the meaning of the word ‘pat’ but may not have heard it enough, or done it enough, to know that usually people pat dogs and not the reverse. Thus, it is not until the third year of life that the kind of contextual probabilities that provide the basis for Strategy C in adult perception develop as guides for sentence comprehension.

Shortly after this development the child goes through a phase in which he depends relatively heavily on something like Strategy D for the comprehension of sentences that do not have semantic constraints. This is brought out by

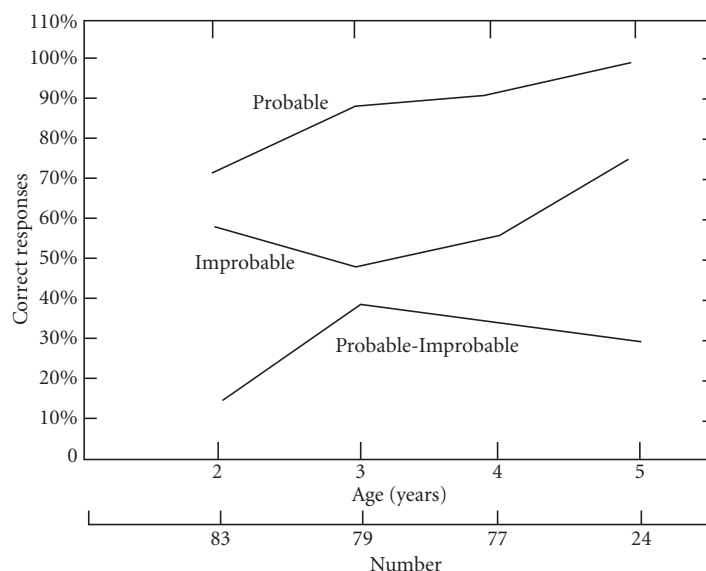


FIGURE 4. The proportion by age of correct responses to probable and improbable sentences, in which each child received one sentence of each kind (roughly half the children at each age received one probable and one improbable active sentence and roughly half received one probable and one improbable passive sentence). The bottom line represents for each age the difference in correct responses between the probable and improbable sentences, and thus is a measure of the children's dependence on probability as an aid to correct performance.

his performance on acting out passive sentences like those in (40a) (see Figure 3). The most important feature of these results is the steady increase in performance until age 3.8 for girls and 4.0 for boys, when there is a sharp (temporary) drop in performance. These results were obtained with only four sentences (of which each child acted out only one), so a larger experiment was run (again by a different experimenter and in a different city) in which twelve reversible passive sentences were used (of which each child acted out three). The results for the passive sentences in this group are presented in Figure 6. Again, the same brief decrease appears at the same ages (although in these materials, the decrease starts at the same time for boys and girls, but lasts to a later age in boys than in girls).

Finally, in a separate experiment, we have studied the performance of the child on cleft-sentence constructions that can reverse the subject-object relation without reversing subjective order, as in (43a, b).

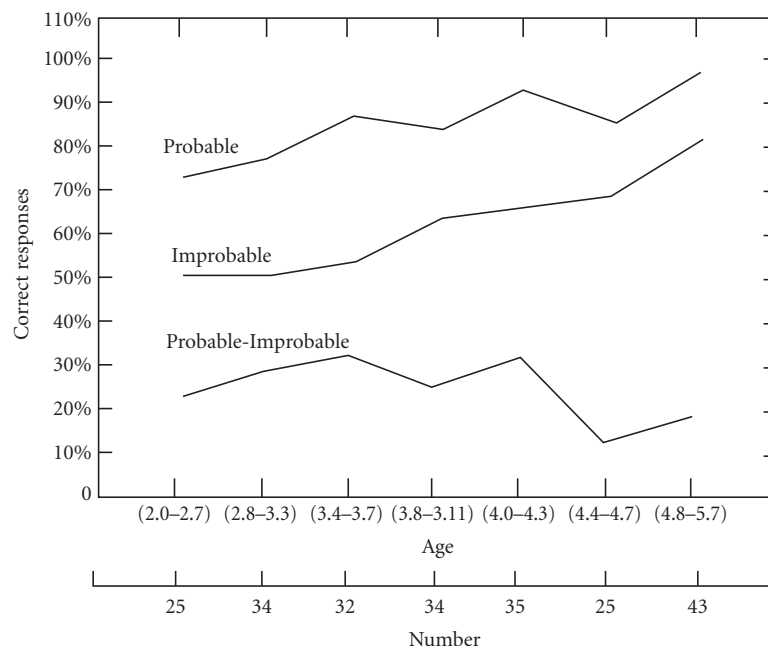


FIGURE 5. The same data display as in Figure 4, except from a different experiment with different children in which each child received three probable sentences and three improbable sentences (balanced within each age group so that the same number of actives and passives of each kind were responded to)

- (43) a. It's the cow that kisses the horse (actor first)  
 b. It's the horse that the cow kisses (object first)

Figure 7 presents the tendency to perform correctly on sentences like (43b). Again, the same decrease in performance appears at about age four.

While any one of these results alone might not be convincing, the constant reappearance of the effect across different experiments with different materials indicates the reliability of the phenomenon. Since each experiment averages across large numbers of children, it is not clear whether the period of the decrease in performance is due to the active development of a perceptual strategy like D or simply to the failure to apply the earlier capacity to understand passives. Of course, in both girls and boys in the experiment in Figure 3, the performance on passives is worse than random at the critical age that indicates an active tendency to use a strategy like D. Similarly, when the reversible passive sentences in the larger experiment are looked at by overall



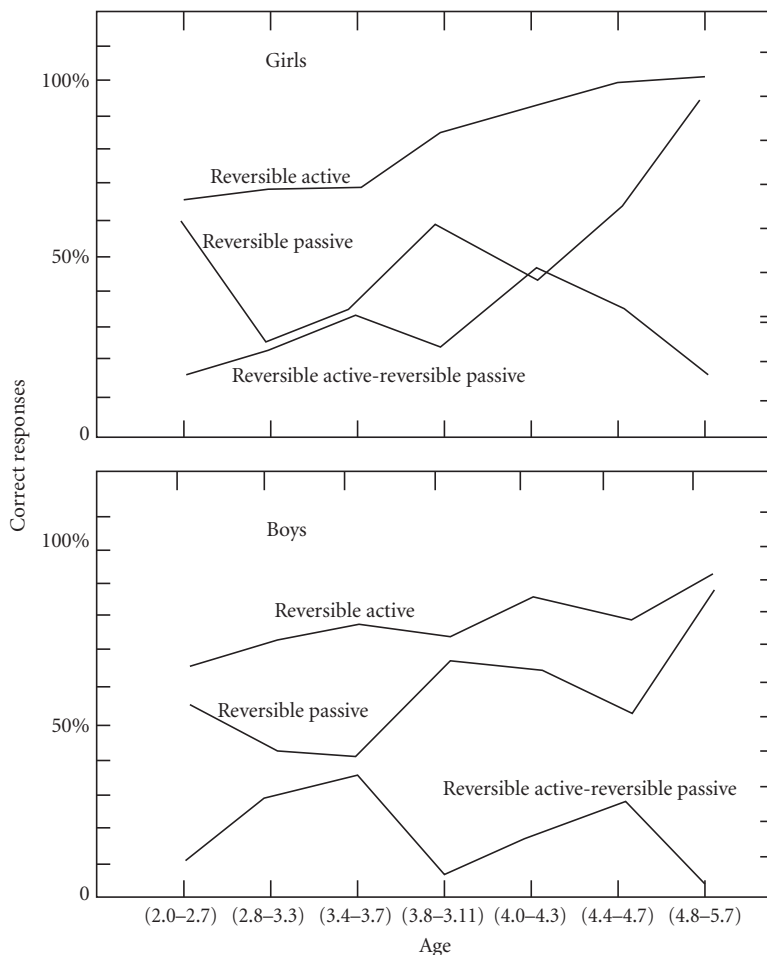


FIGURE 6. The proportion of correct performance by age and sex to reversible active and passive sentences in which each child acted out three sentences of each kind. The bottom line in each graph represents the difference between the performance on actives and passives, and thus is a measure of the children's dependence on the "actor-action-object" order as an aid to correct performance.

difficulty (as measured by overall success during the first three years), the performance on each of the sentences goes below 50 percent (although at slightly different times). Only longitudinal research can resolve the question as to whether all children pass through a phase of overgeneralizing Strategy D or of simply losing their competence and lapsing into random behavior on those semantically reversible sentences that do not conform to SVO pattern.

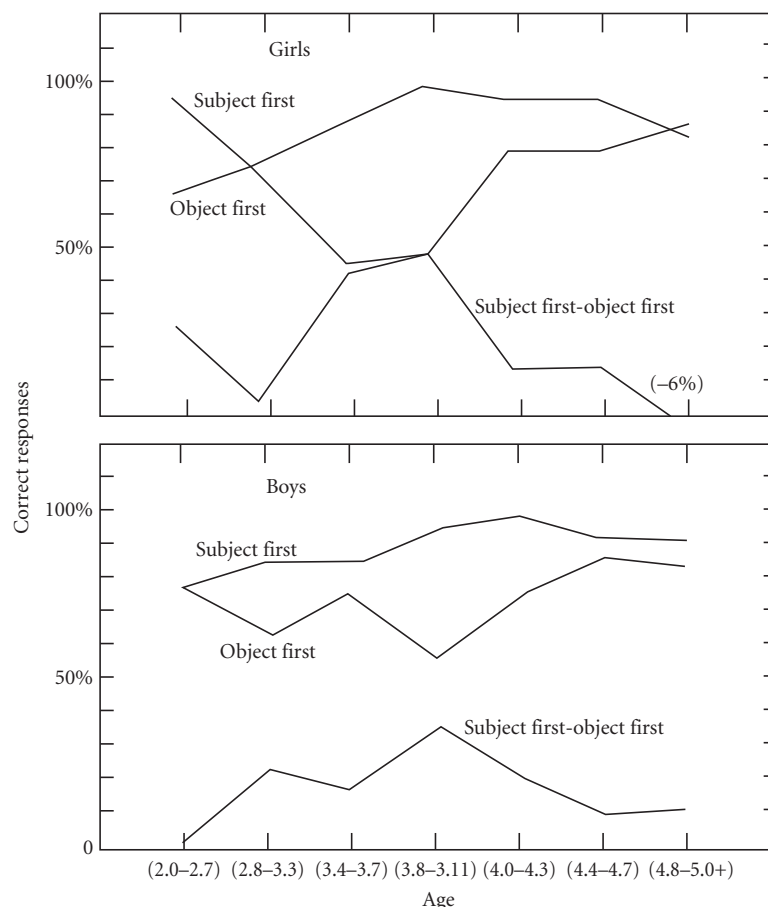


FIGURE 7. The same kind of data display as in Figure 6, except for responses by different children to reversible cleft sentences with subject first (analogous to actives in Figure 6) (e.g., (43a)), and object first (analogous to passives) (e.g., (43b)), in which each subject received three sentences of each kind

However, the fact that we obtain a decrease to below 50 percent even when averaging across children, suggests strongly that all individual children pass through periods of actively applying Strategy D, to misinterpret those sentences in which the first noun is the object rather than the actor.

In any event, we have been able to explore in language the way in which the child may display relative dependence on basic mechanisms, perceptual generalizations, and systematic intuitions as he grows up. This is further support for the distinction between these separable aspects of language behavior that are integrated in adult capacity.

### 3.4 *The enumeration of possible perceptual strategies*

It is an open question as to how the child acquires Strategies B, C, and D. It could be argued that the strategies are formed in response to natural probabilities in the actual speech that the child experiences; in actual speech, sentences may tend to place subordinate clauses second (Strategy B); sentences do usually conform to some sort of contextual constraint (Strategy C); and the active actor–action–object order probably predominates in what mothers say to children (Strategy D). Thus, one could argue that these perceptual strategies are formed by the child as inductions over his experience, as opposed to being due to internal cognitive developments independent of specific experience. However, just as in the case of the acquisition of the strategy of assuming that a relatively large array has a relatively large number of components, one must be prepared to explain *why* it is the case that the perceptual strategies B, C, and D are the ones that the child recognizes as fruitful rather than the many other generalizations that are equally justified by his experience. Thus, even an empiricist view of the acquisition of such perceptual generalizations must include a nativist component that selects certain possible generalizations and rejects others.

I have suggested that the child may extract particular perceptual strategies by selective induction over his early linguistic experience; different topics, different speakers, and different situations justify different perceptual strategies (or at least different relative importance of the strategies). Thus, part of what a child learns when he adapts to the “linguistic style” of a situation is a particular configuration of perceptual strategies which the language used in that situation justifies. Since the number of potential strategies (like the number of sublanguages and of sentences) is infinite, the child must have both a characterization of the set of possible perceptual strategies as well as a routine for the extraction of such strategies from his particular linguistic experience. Analogously, a recognition routine must have *a priori* limits on the kinds of recurrent information it treats as relevant for the formation of perceptual strategies, and a system for the competence and lapsing into random behavior on those semantically reversible distillations of that information into particular strategies.

All we can do at the moment is to define the problem of the specification of possible perceptual strategies. It is clear that probabilistic information about the internal structure and internal/external structure pair is the basis for certain psychological strategies that are developed. But it is not clear that all perceptual strategies are based on experience in this way, nor is it clear which additional linguistic structures are manipulated by strategies. That is, just as

the general study of linguistics seeks to define language universals in terms of the basic structures and universal constraints on possible rules, the study of speech perception must be stated in terms of the basic form of the perceptual mechanism and universal constraints on possible perceptual strategies. Just as certain linguistic structures may be “innate” and some learned, certain perceptual strategies may be basic to all perceptual processes, and some derived from linguistic experience.

#### **4 The influence of the child’s perceptual system on linguistic structure and linguistic behavior in the adult**

A grammar provides the basis for the prediction of new possible sentences from the ones that have already been uttered and heard. That is, the system of grammatical rules that relate internal and external structure is the finite basis for the *acquisition* of linguistic creativity. While the grammatical rules make possible the extrapolation of new sentences from old ones, the system of behavioral strategies make more efficient the *perception* and *production* of sentences.

The preceding sections have demonstrated the independence of the perceptual and grammatical systems for relating internal and external structures of sentences. These systems can manifest themselves as independent systems in the adult and are learned at least partially independently in the young child. However, there is one obvious connection between the two systems in the child: the child will learn the grammar for those sentences that he can understand (at least partially). Conversely, the child will have difficulty in learning the putative grammatical structure underlying sentences that he has difficulty in understanding. Thus, the child’s system of speech perception constrains what he can understand and consequently restricts the kinds of grammar he can learn. To put this another way: the child will learn those grammatical structures most easily which are most consistent with his perceptual system—in those cases in which the grammar offers alternative structures the child will tend to learn only those that are perceptually simple.

In brief, then, the child is simultaneously acquiring two kinds of linguistic systems that can modify each other. It is clear that the structures allowed by the grammar of a language restrict the kinds of perceptual strategies that are learned.

#### **5 Some systems of adult language structure and behavior determined by perception**

Obviously, it is rare that strategies are acquired for the perception of sentences that are not learnable structures. However, it is also rare that rules are

acquired which produce sentences that are impossible to understand. Thus, we can expect that certain aspects of sentence structure reflect the perceptual constraints placed on it by the child as he learns the structure and by the adult as he uses the structure. The following sections present examples of syntactic rules that are acquired in response to perceptual strategies.

### 5.1 *Some syntactic rules*

#### 5.1.1 *The integrity of main clauses in external structure*

Consider the perceptual strategies A and B. These combine to form the perception of an initial “N...V...(N)” sequence as comprising the main independent clause of the sentence. I presented evidence above showing that the very young child’s habits of speech perception are dominated by such a strategy; in those sentences with more than one clause, the first “N...V...(N)” sequence is taken by the young child as the most important clause. Certain facts of adult English syntactic structure appear to accommodate this perceptual strategy; for example, R. Kirk (personal communication) has observed that a subordinate-clause verb that precedes its main verb is generally marked as subordinate by the end of its verb phrase. The most obvious device is the subordinate-clause conjunction, as in (44). In each case the first verb is marked by the subordinate conjunction as subordinate.

- (44) a. *Although* the research was secret the liberated files revealed that it concerned themetabolization of sauce Bearnaise.  
 b. *Because* the demands were non-negotiable nobody wanted any.  
 c. *If* the system corrupts itself the thing to do is to take it over.  
 d. *While* this conference was not attended by any Americans of African antecedence, that fact was obviously an accident.

Subordinate conjunctions are specific lexical items that accommodate to Strategy B by acting as specific markers of those cases in which a subordinate clause occurs before its main clause. There are also certain syntactic verb systems that appear to have formed in response to Strategy B: restrictions on the “syntactically” allowed deletion of words that mark functional relations among clauses. The sentences in (45) and (46) exemplify a heterogeneous set of grammatical restrictions on the stylistic deletion of ‘that’ or ‘the fact’ in initial position.

- (45) a. The fact that the door was discovered to be unlocked amazed the tenants.  
 b. That the door was discovered to be unlocked amazed the tenants.

- c. The fact the door was discovered to be unlocked amazed the tenants.
- d. \*The door was discovered to be unlocked amazed the tenants.
- e. The door was discovered to be unlocked and that amazed the tenants.

For example, (45a) can be reduced to (45b) or (45c) by deletion of one of the initial noun phrases, but not to (45d), in which *both* noun phrases are deleted. The cases in (46) show that, so long as some noun that marks the first clause as a nominalization is in initial position before the clause ‘the door...’, both ‘that’ and ‘the fact’ may be deleted.

- (46) a. The discovery of the fact that the door was unlocked amazed the tenants.
- b. The discovery that the door was unlocked amazed the tenants.
- c. The discovery the door was unlocked amazed the tenants.

Stated in this way, the facts in (45) and (46) exemplify a constraint on internal/external structure relations that requires initial subordinate verbs to be uniquely marked. Notice in (45e) that Strategies A and B do not apply if the sentence has more than one clause so long as the first clause is an independent clause (in the traditional grammarian’s sense of “independent”). The facts in (46) show that this constraint does not apply to the deletion of “complementizers” when they do not precede an initial noun. Perceptual strategy B also predicts certain facts about the deletion of relative pronouns on subject nouns. Consider the grammatical facts in (47).

- (47) a. The man who/that came to buy the giraffe forgot his money.
- b. \*The man came to buy the giraffe forgot his money.
- c. The senator who was in the saddle shot from the hip.
- d. \*The senator was in the saddle shot from the hip.
- e. The senator in the saddle shot from the hip.
- f. The monkey who was running after the bus slipped on a banana.
- g. \*The monkey was running after the bus slipped on a banana.
- h. The monkey running after the bus slipped on a banana.
- i. The monkey who was scared by the dog slipped on a banana.
- j. The monkey scared by the dog slipped on a banana.
- k. The acid that was dissolved in the water became colorless.
- l. ?The acid dissolved in the water became colorless.

It would appear from (47a) and (47b) that there is a general syntactic restriction of the deletion of relative pronouns modifying initial nouns. This restriction follows from Strategy B, since the deletion of the relative pronoun would make the relative-clause verb appear incorrectly to be a main verb of an independent clause. However, there are certain cases in which Strategy B predicts that the relative pronoun can be deleted. For example, the deletion of only the pronoun in (47c) to produce (47d) is blocked, but if the verb ‘was’ is optionally deleted as well, then the pronoun must be deleted to produce (47e). Strategy B allows this, since the subordinate-clause verb ‘was’ is already deleted.

In certain cases the relative pronoun can be deleted even in the presence of a following verb. Thus, (47h) can be derived from (47f) even though the verb form ‘running’ directly follows the noun phrase ‘the man.’ However, the suffix ‘-ing’ at the end of the verb marks it independently as a subordinate-clause verb.

In a small number of cases of relative pronoun deletion, the form of the verb suffixes do not uniquely determine the verb as subordinate. Thus, (47j) can be derived from (47i), even though the past participle verb form ‘scared’ is homonymous with the past form of the verb (that is, to produce the apparent clause ‘the dog scared somebody...’). However, the following participle, ‘by,’ marks the verb ‘scared’ as not having a direct object in the external structure and therefore as being subordinate. Similarly, in (47l) the preposition ‘in’ marks the verb ‘dissolved’ as not having a following direct object in the external structure, and therefore either as intransitive or as a subordinate verb. Thus, Kirk’s observation can be elaborated into a principle for the relationship that exists between the external form of a clause and the internal relations among its elements; it is clear that this principle accommodates to the perceptual Strategy B on the one hand and that it constrains the form of the syntactic rules that relate the internal and external structures on the other hand. Kirk’s principle is stated as follows: if the first noun phrase is followed by a verb not marked as a subordinate verb, by a preceding subordinate conjunction, by inflection (such as *-ing*), or by an immediately following word (such as ‘by’ in (48j) and ‘in’ in (48d)), then the verb is a main verb of an independent clause (of which the noun phrase is the external surface structure subject).

This principle predicts that certain sentences are unacceptable in which the subordinate verb suffixes are homonymous with main verb suffixes. For example, (48a) is not acceptable although it is related to (48b), in the same way as (48c) is to (48d).

- (48) a. ?The horse raced past the barn fell.  
 b. The horse that was raced past the barn fell.  
 c. The horse sent past the barn fell.  
 d. The horse that was sent past the barn fell.  
 e. The horse stumbled past the barn and fell.

The difficulty of (48a) is due to the fact that the verb ‘race’ can occur as an intransitive or as a transitive (unlike ‘stumble’ [pure intransitive]), as in (48e) or ‘send’ [pure transitive] as in (48c, d). In (48a) this facilitates the incorrect assumption that ‘horse’ is the subject of ‘raced’ as a main verb. Notice that, if Kirk’s claim were always true, (48a) should be *ungrammatical* as opposed to merely unacceptable, since it also violates Kirk’s claim. However, to block formally constructions like (48a) with verbs that can act both transitively and intransitively, all constructions with pure transitives (like (48c)) and pure intransitives (like (48e)) would be blocked in a grammar. The alternative is to restrict selectively deletion of ‘that (was)’ to subordinate clauses with verbs that are not potentially phonologically homonymous with intransitive forms. Such a restriction is not only difficult to state, but is also inadequate. Consider the relative acceptability of (49a) over (49b) and (49c) over (49d). For each of these pairs in the less acceptable sentence the “NP... VP” sequence created by deleting ‘that was’ between the NP and VP, is relatively likely as an independent “subject–verb” sentence (italicized in the examples). That is, *any* feature of an initial NP–VP sequence that makes it appear relatively plausible to the listener as a sentence interferes with perception if the NP–VP is not actually a main clause.

- (49) a. *The light airplane pushed past the barn* crashed into the post.  
 b. *The pushing bulldozer pushed past the barn* crashed into the post.  
 c. *The door slammed by the storm* splintered.  
 d. *The door slammed during the storm* splintered.

Thus, Strategies A and B are not offered as grammatical rules but as constraints to which otherwise optional rules may respond when the speaker/listener’s knowledge of individual exceptions does not make them irrelevant. There are some principles that are implicit in the strategies that explain certain other syntactic phenomena of English:

In an “N... V” sequence in which “V<sub>2</sub> is appropriately inflected, “N” is the (external structure) subject of “V,” unless some preceding noun is so marked.



An initial “N” is the external subject of the first appearing “V” (unless blocked for a particular verb by Strategy B).

This explains the fact that the restrictions on relative pronoun deletion also apply to relative clauses in object position. Thus (50b) cannot result from (50a) in modern English, although (50d) is an acceptable variant of (50c).

- (50) a. I ate the apple pie that was yummy.  
 b. \*I ate the apple pie was yummy.  
 c. I saluted the apple pie that my mom made.  
 d. I saluted the apple pie my mom made.  
 e. The flag that was waving above *mom’s apple pie* was groovy.  
 f. The flag waving above *mom’s apple pie* was groovy.

Notice that in (50e, f), Strategy B has marked ‘the flag’ as the subject of the first verb after ‘wave’ so the apparent “NVN” sequence (italicized) is allowed (although of course it may cause perceptual difficulties).<sup>5</sup>

#### 5.1.2 Syntactic restrictions on pronominalization

The structure of coreferential pronominalization is another example in which general perceptual principles appear to constrain formal grammatical structures. Indeed, some authors have recently questioned whether pronominalization is a syntactic phenomenon at all since all attempts to provide a complete account in syntactic terms have failed up to now. However, certain general constraints are stable within syntax. First, whenever two nouns with the same reference appear in one clause, one of them must be transformed into a pronoun. Thus we cannot say (51a) unless there are two distinct ‘Georges’ in mind; rather, we must say (51b):

- (51) a. ?George spoke to George.  
 b. George spoke to himself.

In (50b) the first instance of ‘George’ is said to “govern” the pronominalization of the second instance. It is immediately clear that the “government” of pronominalization always proceeds from left to right within clauses.

<sup>5</sup> Principles like these must apply recursively as exemplified by sentences like ‘Max couldn’t believe that the flag waving above mom’s apple pie was groovy’ in which the principle applies to the embedded sequence after ‘Max couldn’t believe that...’ Note that sentences like (50b) could occur in Old English, presumably because the main object (‘... pie’) was often inflected in the accusative, making its confusion as a subject impossible.

Consequently, the second instance of ‘George’ in (51a) cannot govern the first. If it did, the ungrammatical sentence (51c) would result (on the interpretation that ‘he’ and ‘George’ are coreferential. Note that sentence (51c) can be grammatical if ‘he’ and ‘George’ are different people, but not if they are intended to be the same person).

- (51) c. \*He spoke to George.

This *left–right* constraint on pronominalization also obtains in certain sentences with a main (“independent”) clause and a subordinate (“dependent”) clause; for example, (52a) must be transformed to (52b):

- (52) a. ?George was late although Mary spoke to George.  
 b. George was late although Mary spoke to him.  
 c. \*He was late although Mary spoke to George.

As above, pronominalization cannot proceed right to left; thus, (52c) is not a correct version of (52b). However, pronominalization can proceed right to left if the main clause is to the right of the subordinate clause. In (52a) the ‘George’ on the right can govern pronominalization of the ‘George’ on the left (to yield (53b)), as well as the reverse (to yield (53c)).

- (53) a. Although Mary spoke to George, George was late.  
 b. Although Mary spoke to him, George was late.  
 c. Although Mary spoke to George, he was late.

In general, the only cases in which the left–right constraint can be violated are those in which the governed noun is in a clause subordinate to the governing noun (Langacker 1969; Ross 1967b). Thus, there are two independent constraints on the government of pronominalization: either it proceeds from left to right (to yield (51b), (52b), (53c)), or from main clause to subordinate clause (as in (52b), (53b)). At least one of the constraints must be met; if neither is met, an incorrect sentence like (52c) can result from (52a).

Such a complex system appears at first to be an example of a “pure” linguistic law. However, there is an intuitively clear general principle of all experience that could underlie such complex linguistic constraints. First, for one object to “stand for” another, like a pronoun for a noun, a connection must already be established between them. For example, a picture of a leaf cannot be used to represent a tree unless the viewer already knows the connection. Analogously, in (51c) above, ‘he’ cannot refer to ‘George,’ since the listener does not yet know who ‘he’ is. The constraint that allows a superordinate clause noun to govern the pronominalization of a

subordinate-clause noun may also be interpreted as a linguistic reflection of an obvious regularity of experience: presentation of a whole includes a presentation of its subordinate part, but not vice versa. For example, a picture of a tree also presents a leaf to view since it includes a leaf, but a picture of a leaf does not present a tree (without prior knowledge of the connection, as above). Analogously, a pronoun can appear, even preceding its governing noun, if it is explicitly marked as in a subordinate part of the sentence. Since every sentence has at least one main clause, the listener can predict that a pronoun in a subordinate clause will be governed by a main-clause noun. But a pronoun in an initial main clause does not necessarily have a following subordinate-clause governing noun, since there may be no subordinate clause at all (recall Kirk's observation that subordinate clauses in English are always identifiable as such by the end of the verb phrase). To put it another way, the general perceptual principle is:

A symbol "S<sub>1</sub>" can stand for "S<sub>2</sub>" if (a) the prior connection is known or (b) there is an indication that a connection is about to be established.

The constraints on pronominalization would conform to this principle (it should be noted that more recent linguistic investigations (Lakoff 1970) have brought out some counterexamples to the proposals by Langacker and Ross for the syntactic treatment of pronominalization. Further research is necessary to see if further psychological considerations could account for the new examples).

There are many considerations that show that many of the restrictions on coreferentiality are perceptual as opposed to structural, in any case. Consider first a clearly ungrammatical sequence (54a):

- (54) a. \*John the hit Bill.  
 b. \*John the hit ball.

No manipulation of the semantic constraints (as in (54b)), or of the way in which the sentence is pronounced, or at the gestures accompanying its utterance, can affect its unacceptability. This is characteristic of sequences that are unacceptable on syntactic grounds—semantic or behavioral changes do not reduce their unacceptability. However, semantic manipulations do affect the acceptability of certain sentences with coreferential pronouns, which indicates that pronominalization constraints are not purely syntactic. Consider (55a) and (55b), in which the italicized nouns and pronouns are coreferential.

- (55) a. *The shovel* broke the rake *it* fell on.  
 b. The shovel *it* was below broke *the rake*.

While each of these sentences is acceptable, they cannot combine as in (56a) or (56b–d). Notice that the unacceptability of (56a) is not directly due to the difficulty of grasping the content; even after careful consideration of the content as clarified in (55a, b), (56a) is still nearly incomprehensible. Furthermore, (56a) is not unacceptable because of the actual syntactic relations; consider the acceptability of (57a) and (57b), which have the same syntactic structure as that in (56a), without having two coreferential relations that *cross* each other from one clause to the other.

- (56) a. \*The *shovel*<sub>1</sub> *it*<sub>2</sub> was below broke the *rake*<sub>2</sub> *it*<sub>1</sub> fell on.  
 b. \*The shovel that was above it broke the rake it fell on.  
 c. \*The shovel it was below broke the rake that fell on it.  
 d. \*The shovel that was above it broke the rake that fell on it.
- (57) a. The shovel I was below broke the rake it fell on.  
 b. The shovel it was below broke the rake I fell on.

Thus, the unacceptability of (56a–d) appears to be due to a restriction on having two referential relations crossing over each other. However, if the relations between the nouns and verbs are uniquely determined semantically, coreference relations can cross over between clauses, as in (58) (that is, it is semantically predictable that the *box had contained the ball* and the *ball rolled out of the box*).

- (58) The box it rolled out of scratched the ball it had contained.

Coreference also can occur over between clauses without unique semantic constraints so long as the pronouns differ superficially. For example, (59a) and (59b) can combine into (60a), a perfectly acceptable sentence.

- (59) a. The boy deserved the girl he kissed.  
 b. The boy she pleased deserved the girl.
- (60) a. The boy she pleased deserved the girl he kissed.  
 b. The boy that wanted her deserved the girl he kissed.  
 c. The boy she pleased deserved the girl that kissed him.  
 d. The boy that wanted her deserved the girl that kissed him.

The conclusion from these facts is clear: so long as the referential relations are kept distinct from each other, by general semantic restrictions or by unique referential possibilities of the pronouns, coreferential relations may cross each other; that is, the unacceptability of the sentences in (56) is *not* due to a

syntactic restriction, but to a behavioral one. Certain examples bring out even more clearly the effect of nonstructural behavioral variables on the acceptability of sequences with coreferential pronouns. Thus, (61a) and (61b) are unacceptable while (61c) and (61d) are acceptable, since the two pronouns obviously refer to different people because of the difference in sex.

- (61) a. \*He and he liked the cannabis juice.  
 b. \*Him and him liked the cannabis juice.  
 c. Her and him liked the cannabis juice.  
 d. She and he liked the cannabis juice.  
 e. ?He and him liked the cannabis juice.

In fact, (61e) is almost acceptable even with pronouns of the same sex; the difference in the external form of the pronoun is sufficient to make the sentence acceptable. The dependence of such pronominalization restrictions on nonstructural variables is brought out even more markedly by special execution of (61a, b); these sentences are completely acceptable when spoken with accompanying gestures, first pointing to one (male) person and then another. Thus, many of the factors that govern pronominalization are clearly the linguistic reflection of behavioral constraints on symbols that “stand for” other symbols; those syntactic mechanisms that are involved in the description of pronominalization clearly are grammatical responses to such behavioral constraints.

### 5.1.3 *Syntactic restrictions on pronominal adjective ordering*

The previous two examples of the effect of general psychological principles on structure would appear to be extremely general, if not universal; they both bear on the relation between subordinate and superordinate clauses, which is itself a putatively universal structural distinction. Certain perceptual strategies are language-specific in that they depend on particular properties of a language which themselves are not universal. Consider, for example, the strategies involved in the immediate perceptual segregation of major phrases. The implication of perceptual strategies like A and B is that understanding a sentence involves a marking of the internal relations between the phrases in each clause. This itself assumes that the phrases themselves have been (or are being simultaneously) isolated from each other. It is easy to see the importance of such segregation. For example, in (62a), the perceptual segregation of words in the verb must end (and begin) with ‘called’, while in (62b) it must include the participle ‘up.’ Or in (62c) the subject noun phrase must end with

the word ‘snow,’ while in (62d) it must end with the following word ‘catches.’ If such perceptual segmentation into major phrases is not achieved, the internal relations themselves cannot be assigned.

- (62) a. I called up the not very well painted stairs.
- b. I called up the not very well painted Indian.
- c. The powerful snow catches the travelers.
- d. The powerful snow catches saved the travelers.

Of course, in many instances the semantic relations and unique lexical classifications in English can themselves determine the segmentation. Thus, there is no doubt (even temporarily) about the segmentation of the verb in (63a) or the first noun phrase in (63b). Furthermore, in many instances, stress and intonation patterns can provide the necessary clues (consider the relatively high stress on ‘up’ in (62a) or on ‘catches’ in (62c)).

- (63) a. I called from the not very well liked stairs.
- b. The powerful snow barriers saved the travelers.

In English the presence of certain function words at the beginning of a phrase can uniquely determine what kind of phrase it is, and therefore what to look for at its termination. For example, the determiners ‘the, a, some, one, many,’ and the like, all signify the beginning of a noun phrase. A putative perceptual strategy (E) could be based on this fact:

**Strategy E**

“determiner...” begins a noun phrase.

Consider sentence (64a). This principle leads a listener to expect a noun terminating the noun phrase begun by ‘the’; in (64a), the form class possibilities of ‘pencil’ and ‘fell’ uniquely determine the interpretation that the noun phrase is ‘the ball’ and that ‘fell’ is the verb.

- (64) a. The pencil fell.
- b. The pencils fell.
- Sam
- c. The pencil (that) found fell.
- the boy
- d. The nice pencil fell.
- e. The plastic pencil fell.

In addition, there are certain classes of morphemes and words that uniquely identify the boundary of a head noun phrase; in (64b), the plural morpheme

‘s’ (given that a noun doesn’t follow it), and in (64c), the function word ‘that’ or the proper noun ‘Sam’ or the determiner ‘the’ all signal that the immediately preceding noun was the head noun of its noun phrase. That is, there is a perceptual strategy (E’).

### **Strategy E’**

The first noun after “Determiner...” (or the first noun with a following morpheme that marks the beginning of a new noun phrase) is the head noun, which terminates the noun phrase (independent of nouns in relative clause).

Strategy E’ operates correctly in cases like (64d), where an adjective intervenes between the determiner and the noun. However, since the adjective ‘nice’ is not a noun, Strategy E’ does not establish the segmentation of the noun phrase until the noun ‘pencil’ is heard. Strategy E’ would operate incorrectly in cases like (64e), in which one of the prenominal modifiers is itself lexically marked as a noun. Since ‘plastic’ is the first noun after ‘the,’ Strategy E’ would establish the following segmentation to sentence (64e), which is incorrect:

(65) (The plastic)<sub>NP</sub> (pencil)<sub>NP</sub>...

To block this kind of premature NP segmentation, the strategy E’ must be restated so that it does not establish segmentation of a noun phrase until there is a word that is relatively less nounlike.<sup>6</sup>

### **Strategy E’’**

After “determiner...” the boundary of the head noun phrase is marked by (1) a set of morpheme classes that signal the end of a noun phrase (such as ‘s’) or immediately subsequent morphemes that signify the beginning of a new noun phrase (such as ‘the,’ proper nouns) or a relative clause (such as ‘that’) and (2) a subsequent lexical item that is less uniquely a noun.

E’’ yields the correct segmentation for (64e) (and indeed covers most of the cases in (64)). However, it is not clear whether principle E’’ extends the noun phrase as long as possible, or whether it establishes segmentation at the earliest possible point. If the former is true, then (66a) should be more complex than (66b); in (66a) the word ‘marks’ would be incorrectly included within the noun phrase because, while it is a verb, it is homophonous with a noun (as in (66b)).

<sup>6</sup> I am indebted to M. Halle and J. Grinder for suggestions on this problem.

- (66) a. The plastic pencil marks easily.  
b. The plastic pencil marks were ugly.  
c. The plastic rose fell.  
d. The plastic rose and fell.

On the other hand, if principle E<sub>1</sub> applies at the first possible point, then (66c) should be more complex perceptually than (66d); the word ‘rose’ would not be included within the noun phrase, because, while it is a noun, it is homonymous with a verb (as in (66d)). Future experimentation is necessary to decide this question. In any case, the problems raised by the sequences in (66) are usually resolved by normal intonation and stress patterns.

However, nuances of stress do not resolve the segmentation problem exemplified in (64e), so Strategy E<sub>2</sub> is required for the segmentation of noun phrases with prenominal modifiers. This strategy appears to act as a constraint on the external order of prenominal modifiers that might otherwise be freely ordered. Consider the constraints on the order of adjective classes exemplified in (67).

- (67) a. The red plastic box.  
b. \*The plastic red box.  
c. The large red box.  
d. \*The red large box.  
e. The large plastic box.  
f. \*The plastic large box.  
g. The large red plastic box.  
h. \*The plastic red large box.

Notice first that any two prenominal adjective sequence is acceptable if the first adjective is given contrastive stress (for example, in (67b), the phrase would have to be in a contrasting context like ‘not the metal red box, but the plastic red box...’). However, with neutral stress the order of prenominal adjectives is constrained. Several recent theories (Vendler 1968; Martin 1968) state that adjectives are ordered according to the extent to which an adjective is related lexically to a noun (Vendler), or to which it refers to a “substantive, concrete” quality of an object (Martin); the more “nounlike” an adjective is (on either of these two measures), the closer to the noun it must be. Thus, for example, following Vendler, we can argue that a substance adjective like ‘plastic’ is more like a noun than a color adjective like ‘red,’ in the sense that it occurs in more kinds of constructions as a noun than does ‘red’ (see



(68)); similarly, color adjectives like ‘red’ occur in more constructions as nouns than do size adjectives like ‘large’ (see (69)). Martin has recently suggested a more semantic basis for a scale of “nounlikeness” of adjectives; substance words (‘plastic’) refer to the concrete “inner” structure of the noun they modify; color words (‘red’) refer to the exterior of the object they modify; and size words (‘large’) refer to qualities of the objects they refer to which must be assessed by the speaker relative to other objects of that type.<sup>7</sup>

- (68) a. Red is a color; redness is nice.  
 b. Plastic is a substance; plasticity is nice.  
 c. \*That is made out of red.  
 d. That is made out of plastic.  
 e. \*The red broke.  
 f. The plastic broke.  
 g. ?Reds are of variable quality.  
 h. Plastics are of variable quality.
- (69) a. Red is my favorite color.  
 b. \*Large is my favorite size.  
 c. He splattered some red on me.  
 d. \*He splattered some large on me.  
 e. Red and blue and green are colors.  
 f. ?Large and enormous and tiny are sizes.

Whichever metric of “nounlikeness” is used, the syntactic constraints on prenominal adjective ordering principle are expressed the same way: *in a series of pronominal adjectives, the more nounlike adjectives are ordered to be closer to the head noun they all modify*. The perceptual strategy for the segmentation of noun phrases developed in Strategy E’ can explain this otherwise strange grammatical constraint. If more nounlike adjectives preceded less nounlike adjective, then Strategy E’ would produce premature segmentation. For example, principle 2 of E’ would incorrectly segment the phrase in (67b) as shown in (70) as if it were a correct segmentation, like that of (67a), shown in (71). This follows from the fact that ‘red’ is less nounlike than ‘plastic.’

<sup>7</sup> I have summarized the arguments of Vendler and Martin in vastly abbreviated form. The reader should consult their original work on this problem.

(70) (The plastic)<sub>NP</sub> (red pencil)<sub>NP</sub>

(71) (The red plastic pencil)<sub>NP</sub>

Thus, sequences that violate the general constraint on noncontrastive pre-nominal adjective order are incorrectly segmented by Strategy E'.<sup>8</sup>

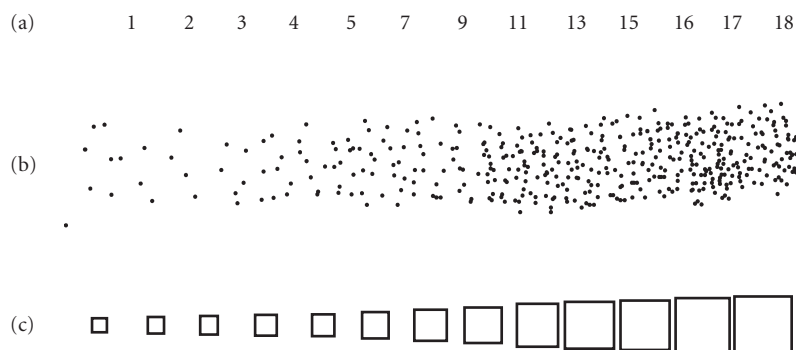


FIGURE 8. Examples of different kinds of perceptual segmentation. In each case there are two main segments with an intermediate transition.

If the above arguments are correct, then the restriction on prenominal adjective ordering is an example of the effect of perceptual strategies on “grammatical” structure. I suggested above that perceptual strategies affect grammatical structures in those cases in which the child acquires the strategies before he acquires certain grammatical structures; grammatical structures acquired after he learns the strategies will be affected by them. Suppose that the child acquired Strategy E' before acquiring the ability to process more than one prenominal adjective at a time; this strategy could be expected to constrain the preferred adjective order that he eventually acquires—given that adjective order is otherwise syntactically free (which is indicated by the fact that, with contrastive stress, any order of two adjectives is possible—see above). Recently we have tested this view of the ontogenesis of prenominal adjective constraints with children between two and five years of age. We present the child with phrases like those in (67), some of which follow the adult constraints (such as (72)) and some of which do not (such as (73)), and

<sup>8</sup> Notice that a general “semantic” account of adjective ordering like Martin’s is incorrect. If adjectives are postposed, then the order is free, as in *I like my pencils red and plastic* or *I like my pencils plastic and red*. That is, the ordering constraint only applies to prenominal ordering. Furthermore, prenominal comparative order is free, as in ‘I never saw a redder larger box.’ This is presumably due to the fact that the comparative suffix ‘-er’ wants every adjective as equally non-nounlike.

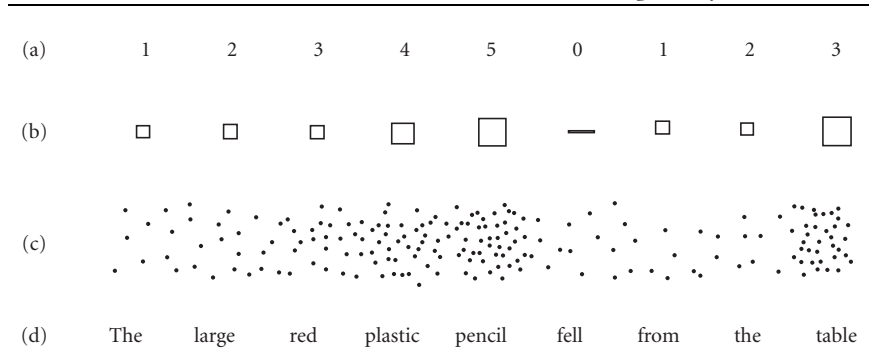


FIGURE 9. Examples of the perceptual segmentation corresponding to correct adjective orderings. In each case, the first major segment terminates after the fifth item.

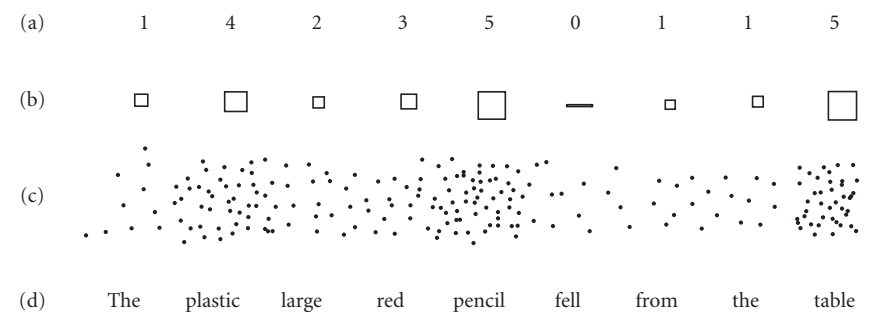


FIGURE 10. Examples of the perceptual segmentation corresponding to incorrect adjective orderings. In each case the first major segment terminates after the second item.

ask him to say back to us what we say (see Bever, Mehler, and Epstein 1968 for details).

(72) The large plastic pencil fell from the table.

(73) The plastic large pencil fell from the table.

The crucial result is that younger children perform better than do older children on the repetition of sequences that do not follow the ordering constraints: the age at which the child's performance deteriorates on this task is just the age at which our other research shows him to be acquiring strategies for speech perception. This is consistent with our proposal that the constraints on prenominal adjective ordering are basically due to perceptual strategies.

The details of a strategy like E<sup>1</sup> are obviously language-specific, since there are many languages without explicit determiners, without pronominal adjectives. However, it is also clear that Strategy E<sup>1</sup> is a special instance of an extremely general principle of perceptual grouping (Principle F). This principle articulates the fact that perceptual segmentation tends to be established only at points in a stimulus where a discontinuity of relations (“R<sub>i</sub>”) is perceived (but not at all such points).

- a. 1 2 3 4 5 7 9 11 13 15 16 17 18
- b. [density gradient]
- c. [size gradient]

### Principle F

In an ordered array of adjacent items ... XYZ ... if XR<sub>j</sub>Y is the same as YR<sub>j</sub>Z, then the array is grouped together. If YR<sub>j</sub>Z is different from XR<sub>j</sub>Y, then the stimulus is segmented (XY) Z.

Consider first the application of Principle F to nonlinguistic stimuli. If asked to segment the sequence of numbers in Figure 8a into groups, there would clearly be two main groups, 1–5 and 13–18, separated by a transition 5–13. The first and third segments are grouped by unit increases, and the middle transition segment by increases of 2.<sup>9</sup> Similarly, Figure 8b and Figure 8c are each made up perceptually of two segments with a transition Figure 8b. The first and third groups are ordered by slightly increasing density, while the middle group is transitional between them. In Figure 8c the first and third segments are ordered by slightly increasing size, while the middle segment is ordered by radically increasing size. Notice that cases like Figure 8c are special instances of contours. A visual edge is defined according to Principle F as a point at which R<sub>i</sub> between two adjacent points differs from R<sub>i-1</sub>.

The constraint on pronominal adjective endings in a sequence of adjectives is that the second must be more “nounlike” than the first. That is, in a sequence of pronominal adjectives, “Adj<sub>1</sub>R<sub>i</sub> Adj<sub>2</sub> R<sub>j</sub> Adj<sub>3</sub> ...,” both R<sub>i</sub> and R<sub>j</sub> are the relation “less nounlike than.”

Consider now the application of Principle F to the segmentation of cases that are directly analogous to the pronominal adjective ordering constraints within noun phrases (Figure 9). The natural segmentation in each of these

<sup>9</sup> Notice that it is ambiguous to which group ‘5’ and ‘13’ belong uniquely. This is again a special case of the question as to whether segmentation is established at the earliest or latest possible point in a sequence. See above.

cases is following the fifth segment, and at no point preceding it (in Figure 9 the nounlikeness of the adjectives corresponds to the largeness of the numbers in Figure 9a, the size of the figures in Figure 9b, and the intensity of the shading in Figure 9c). Suppose that the sequential visual and numerical relations were analogous to a sequence that violates the adjective ordering constraints, as in Figure 10. While a perceptual boundary following the fifth segment remains (just as in the case of the linguistic sequence), there is some uncertainty as to an additional boundary following the second segment in each array. It is exactly this perceptual uncertainty as to perceptual grouping that I have claimed is the basis of the ordering constraints on prenominal adjectives.

Every specific strategy of speech perception is a special case of a general principle of perception, at least in the sense that no general perceptual law may be violated by a language-specific strategy. Thus, the fact that Strategy 1c is a linguistic reflection of Principle F is not an exploitation of Strategy 1c, rather, it is simply a classification of the linguistic strategy in terms of the general principles that it utilizes.

## 5.2 *The reflection in sentential perceptual complexity of universal perceptual restrictions*

During the course of this discussion I have emphasized the ways in which linguists depend on intuitions about sentences as the basic source for data that must be described by a grammar. The demonstration of the tripartite nature of speech behavior articulates the possibility that such intuitions about sentences also are of three basic kinds. For example, it is perfectly clear that the unacceptability of (74a) is due to the violation of the basic linguistic property of reference while (74b) lacks an action. These sequences may be said to be unacceptable as sentences due to violations of basic (universal) linguistic properties. In contrast with this type of violation, the sequences in (75) are unacceptable as sentences due to violations of the rules governing the specific syntactic rules of English.

- (74) a. \*Please pass me the bruck.
- b. \*Tom, Dick and Harry.
- (75) a. \*hoped it for you to win the loot.
- b. \*The group are better than you think.

That is, these sequences violate the surface-level systematic properties of English (notice that the sequences in (74) and (75) could all be uttered and

understood in ordinary speech). I pointed out above that the primitive basic capacities on the one hand and sophisticated epistemological systems are both easier to isolate in adult behavior than the perceptual system. Accordingly, the unacceptability of a sequence that is due to the failure to meet some basic property of all sentences (such as (74j)) is easy to distinguish from unacceptability due to the failure to maintain properties specific to the particular language ((75)). However, sequences that are unacceptable due to the violation of perceptual strategies are relatively hard to identify. Thus, it is not immediately clear whether the ungrammaticality of sentence (76a) is due to linguistic properties of English or due to the mechanisms of speech perception. At first blush it might appear that sentences like (76a) should not be generated by a grammar of English, since they are not immediately acceptable; however, it is possible to argue that there is a near-continuous scale of acceptability between (76a) and (76j) in which the independent variable is the complexity of the sequence that separates the verb ('call') and the particle ('up'). Thus, it is plausible to argue that the apparent unacceptability of (76a) is due to the length of the phrase intervening between the verb and its particle. That is, (76a) is classified as acceptable syntactically, but complex perceptually, because of the load it places on immediate memory of the material between the verb and its particle (notice that the acceptability of (76k) shows that the unacceptability of (76a) is not due to the length of the sentence per se, but to the length of the sequence interrupting the verb and its associated particle).

- (76)
- a. \*John called the not very well liked but quite pretty girl on the next block where Jack has lived for years up.
  - b. \*John called the not very well liked but quite pretty girl who lives on the next block where Jack lived up.
  - c. ?John called the not very well liked but quite pretty girl who lives on the next block up.
  - d. ?John called the not very well liked but quite pretty girl who lives on the block up.
  - e. John called the not very well liked but quite pretty girl up.
  - f. John called the very well liked and quite pretty girl up.
  - g. John called the well liked and quite pretty girl up.
  - h. John called the pretty girl up.
  - i. John called the girl up.
  - j. John called up the girl.

- k. John called up the girl who is not very well liked but quite pretty and who lives on the next block where Jack has lived for years.

5.2.1 *Sequences that interrupt each other—Save the hardest for last*

To generate the acceptable sequences (76h–k), the formal grammar must also generate the less acceptable sequences (76a–g). That is, there is no natural way in which (76a–g) can be blocked by a grammar that also generates (76h–k). However, there is a general perceptual rule (Principle G) that can be used to explain the unacceptability of (76a) on behavioral grounds, and thus explain why it is simultaneously grammatical and unacceptable.

**Principle G**

Sequences with constituents in which each subconstituent contributes information to the internal structure of the constituent are complex in proportion to the complexity of an intervening subsequence.

Notice that immediate memory may be exhausted either by the length of an intervening sequence or by the perceptual complexity of that sequence. Thus, (77a) is more acceptable than the equally long (76a), while (77b) is less acceptable than the equally long (76g).

- (77) a. John called Jane, Mary, Marsha, Sally, Joan, Melissa, Erica, Felicia, Irma, Urania, Galacia and all the other girls in his class up.  
 b. John called girls seen by the sailor he met up.

There are various other linguistic phenomena covered by Principle G. Recently Ross (1967a) has suggested that there is a general constraint on postposition in English which orders “heavier,” or more complex, noun phrases toward the end of a sentence. For example, (78a) is more acceptable than (78b); (78c) is more acceptable than (78d); and so forth.

- (78) a. John called up the girl in the white dress.  
 b. John called the girl in the white dress up.  
 c. John showed the girl the book that I liked a lot.  
 d. John showed the book that I liked a lot to the girl.

These cases are all characterized by the sequence “... Verb XY...” where X and Y both have some unique internal relation to the verb (such as “particle, direct object, indirect object”). Ross’s relative complexity constraint may be viewed as a special extension of Principle G. Consider the sequence “... Verb X Y...” in which Y is less complex than X and both X and Y are related to the verb in the

internal structure (that is, X and Y are dominated by VP). Suppose the complexity of the Verb–X relation taken independently is assigned a value of “x” and Y is assigned a value of “y” where  $y < x$ . Then the complexity of the relations (taken separately) in a sequence “V X Y” is the quantity  $(x + y)$ ; assume that the interaction with short term memory is defined as a factor “m” which is proportional to the complexity of what must be remembered. The overall complexity (including the ordering) of “...Verb XY...” in which X must be held in memory would be  $(x + y + m x)$  and of “Verb Y X” would be  $(x + y + m y)$ . Since by assumption  $y < x$ , the complexity of “V X Y” is greater than that of “V Y X.” That is, in those cases in which the syntax provides free ordering between X and Y, the preferred order is one that places the more complex noun phrase so that it does not interrupt the relation between the verb and the less complex noun phrase; this ordering yields the simplest overall complexity of the sequence. This principle also accounts for the relative acceptability of post-verb ordering. The basic rule is that more complex adverbs are ordered towards the end of the sentence. Thus, (79a) is more acceptable than (79b), while (79c) is more acceptable than (79d) (the intuitive basis for the constraints on adverb order is far weaker than on adjective order).

- (79) a. John walked briskly in a slightly more northerly direction.  
 b. John walked in a slightly more northerly direction briskly.  
 c. John walked north at a slightly brisker pace.  
 d. John walked at a slightly brisker pace north.

Principle F would explain these facts, since in (79a) and (79b) the more complex adverbial phrase comes after the less complex phrase, while both modify the verb.<sup>10</sup> Principle F also accounts for certain stylistic preferences that indicate that

<sup>10</sup> In (79) the adverb category (such as direction or manner) is held constant so the ordering constraints are not due to constraints on the order of adverb categories. The fact that relatively complex adverb phrases are always displaced towards the end of the clause allows us to investigate category restrictions on adverb order, by holding complexity constant and equal between any pair of adverbs. Such comparisons indicate that the canonical postverb adverbial order is DIRECTION MANNER PLACE DURATION FREQUENCY TIME PURPOSE. Thus (a) is correct, while (b) is not. Or to give an example with all categories,

- (a) John walked north fast.  
 (b) John walked fast north.  
 (c) is correct and (d) is appalling. The reader is invited to test my intuitions pair by pair.  
 (c) Georgeala Cough rode north fast in the park briefly often yesterday for fun.  
 (d) Georgeala Cough rode for fun yesterday often briefly in the park fast north.

The source for these constraints is unclear to me at the moment, although it does appear that the direction and manner adverbs modify the verbs while there are sentence modifiers. Furthermore, the order Place ... Purpose appears to be in the direction of increasing abstractness, and (consequently) of increasing psychological complexity. If this observation is true, then Principle D can account for these ordering constraints as well. The relatively complex adverb is ordered relatively late. To test the reality of the constraints themselves, M. Garrett and I played sentences like (a) and (b) to subjects with an



the more complex of two modifiers appears later in a sequence of two. For example, (a) is preferred over (b) in the pairs of examples below:<sup>11</sup>

(80) PRENOMINAL CONJOINED ADJECTIVES OF THE SAME CLASS

- a. The steel and artificially strengthened fiber plastic tube broke.
- b. ? The artificially strengthened fiber plastic and steel tube broke.

CONJOINED PREDICATE NOMINALS

- c. The machine is bulky and incredibly hard to operate without the appearance of at least one malfunction.
- d. ? The machine is incredibly hard to operate without the appearance of at least one malfunction and bulky.

POSTPOSITION OF COMPLEMENT SUBJECT (when the object is not complex)

- e. It amazed Bill that John left early in the morning to catch the train.
- f. ? The fact that John left early in the morning to catch the train amazed Bill.

CONJOINED SENTENCES

- g. It rained while the dog barked at his master's voice mysteriously coming out of a big black funnel.
- h. ? The dog barked at his master's voice mysteriously coming out of a big black funnel while it rained.

5.2.2 Sequences with two simultaneous functions—Three's a crowd



Unfortunately, it is not the case that the perceptual source of the relative unacceptability of a syntactically allowed sequence is always so easily identified and precisely described. Consider (81a), a so-called “double embedding.” It must be generated formally by a contemporary linguistic grammar that also generates (81b). It is not possible to restrict the number of embedded subordinate clauses to one, because of sentences like (81c), which have two embeddings, but are perfectly acceptable. Indeed, since the internal structures of (81a), (81c), and (81d) are identical and only (81a) is unacceptable, no restriction on the form of internal structures themselves can account for the unacceptability of doubly embedded sentences. Rather, it is a function of the way in which the internal relations are presented in the external structure, Fodor and Garrett (1967) suggest that it is the density of the number of internal structure sentence units

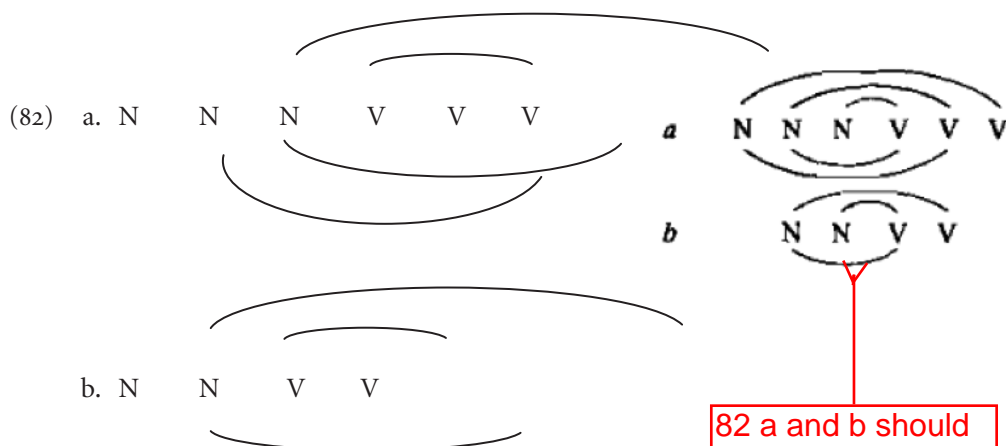
accompanying task (click location) to increase errors. We found that subjects tend strongly to reverse adverb sequences that violate the canonical order stated above.

<sup>11</sup> I am indebted to G. Miller for suggesting some of these examples.

per word in the external structure that exceeds some critical threshold (“density” =  $3/12$  for (81a); “density” =  $3/14$  for (81c)). This proposal is intriguing since it would suggest that at least one dimension of perceptual complexity is quantifiable. However, the proposal is incorrect. The density of internal structures per word is even higher in (81d), but (81d) is entirely comprehensible and acceptable. Thus, the complexity of center-embedded sentences cannot be easily explained away by appeal to any obvious perceptual principle. This creates a dilemma—either we must accept the current form of generative grammar as incorrect, since it cannot avoid generating center-embedded sentences in a natural way, or we must appeal to an unspecified perceptual strategy to account for its difficulty.

- (81) a. The dog the cat the fox was chasing was scratching was yelping.  
 b. The dog the cat was scratching was yelping.  
 c. The fox was chasing the cat that was scratching the dog that was yelping.  
 d. The fox was chasing the cat scratching the yelping dog.

Chomsky and Miller (1963) have attempted to define such a perceptual principle. They argue that any perceptual principle may not interrupt its own operation more than once. In the case of a sentence like (81b) (represented schematically in (82b)) the perceptual assignment of the “actor”–action relation to the first noun and last verb is interrupted by the same assignment to the second noun and first verb. In (81a) (represented in (82a)), the perceptual assignment of actor–action to the first noun and last verb is interrupted by the assignment of the same relation to the second noun and the second verb, which is in turn interrupted by the assignment of the same function to the last noun and the first verb (upper lines in (82) represent subject–verb relations. Lower lines represent verb–object relations).



It is intuitively clear that a self-interrupting operation is more complex than one which does not interrupt itself. However, it is not theoretically motivated that one interruption be acceptable (as in (81b)) and two interruptions be entirely unacceptable (as in (81a)).

It is possible to subsume the relative unacceptability of double-embedded sentences under a general perceptual Principle H, which simultaneously accounts for the perceptual difficulty of a superficially heterogeneous number of types of sentences.

### ***Principle H***

A stimulus may not be perceived as simultaneously having two positions on the same classificatory dimension.

Principle H states the unavoidable fact that a stimulus cannot be perceived in two incompatible ways at the same time. This principle combines with the view of speech perception as a function of direct mapping of external sequences onto internal structures to predict the difficulty of any sequence in which a phrase has a “double function” with respect to such a mapping operation. Before applying Principle H to explain the difficulty of center-embedded sentences, consider first some well-known facts.

Miller and Selfridge (1950) found that sequences with low-order probability approximations to English were difficult to perceive; for example, a sequence like (83) is more difficult than (84).

(83) he went to the newspaper is in deep (2nd-order approximation)

(84) then go ahead and do it if possible (7th-order approximation)

(A “second-order approximation” is generated by giving a subject two words (such as ‘he went’) and asking him to produce the next word of a sentence (‘to’); the next subject is given the last two words of the sequence (‘went to’) and produces the next word (‘the’). A “seventh-order approximation” is generated by giving each subject the last seven words of the sequence each time.) The relative ease of perceiving sentences as they increase in order of approximation was taken by Miller and others as evidence for the organizing role of syntactic structure at levels higher than a single word. For example, in sequence (84) the words form a sentence, while in (83) they do not. However, this does not explain the exact psychological nature of the difficulty of low orders of approximation. In fact, if forming a sentence makes word strings easy, it might be predicted that sequence (83) should be psychologically simpler, since it simultaneously forms two sentences (as in (85) and (86)).

60 Thomas G. Bever

(85) he went to the newspaper

(86) *the newspaper* is in deep

The real basis of the psychological difficulty is clear: the italicized portion of the sequence is vital to each sentence—that is, it has a “double function.” There is a general cognitive restriction that results in psychological complexity whenever such double functions appear. As a visual example, consider the representation of the two squares on the left when they are adjacent. The line labeled “y” is simultaneously shared by the right and left squares. As a result, Figure 11 is generally perceived as a divided rectangle rather than as two adjacent squares. Often such double functions in vision can produce “impossible” figures from the combination of two possible figures, such as Figure 12.

The general psychological principle that governs these visual examples is a special case of Principle H: in a closed system, a component of a stimulus cannot serve two opposite functions at the same time. That is, in Figure 11 line y cannot both end one square *and* begin another; or in Figure 12 the segment labeled “y” cannot both end one kind of figure (the three-dimensional “u” opening right in the segments labeled x–y) *and* begin the other (the three poles in the segments labeled y–z).

There is a related explanation for the psychological difficulty of “center-embedded” sentences. Phillips and Miller (1966) noticed that part of the complexity of center-embedding may be due to the fact that in a sentence like (81a), the second noun is the subject of one clause and the object of another (see (82)). If understanding a sentence involves labeling each word for its logical function in the underlying structure, then the second noun in (81a) could be interpreted as having a “double function” with respect to a strategy that maps external noun sequences onto internal structures, in which the first noun is the object of a verb of which the second noun is the subject. With respect to the preceding noun, it is an object with respect to the following noun. The general double-function hypothesis for perception following from Principle H is this:

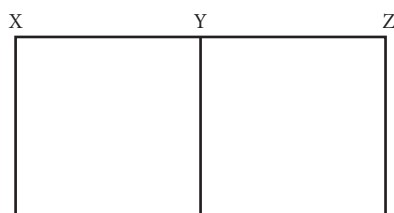


FIGURE 11. Figure most easily seen as a rectangle with one division at “y,” rather than two squares joined at “y”

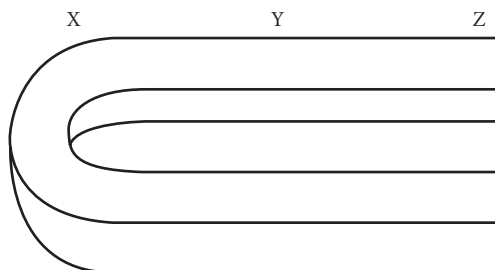


FIGURE 12. Figure that is “impossible” because of the combination of two- and three-dimensional projections at point “y”

### Principle I

In a sequence of constituents  $x, y, z$ , if  $x$  has an internal relation  $R_i$  to  $y$  and  $y$  has the same internal relation to  $z$ , and  $x, y$ , and  $z$  are superficially identical, then the stimulus is relatively complex, due to  $y$ 's double function in the perceptual strategy,  $S_i$

$$S_i : x \ y \rightarrow x \ R_i \ y$$

$S_j$

Notice that the prediction of the perceptual difficulty of center-embedded sentences from Principle I depends on the existence of strategies for the direct perception of the internal structure relations from the external sequence, which define the relations ( $R_i$ ) that adjacent phrases be part of each other. One relevant strategy is presented in Strategy J.

### Strategy J

In...  $NP_1 \ NP_2 \ (VP)$ ...sequence in the external structure,  $NP_1$  is the internal object of an internal structure sentence unit of which  $NP_2$  is the subject.

Of course, Strategy J (like A, B, and C) is not *always* true, as in (87), but it is probably true of external sequences most of the time. The same is true of Strategy K.<sup>12</sup>

(87) *The boy the girl was seen by is here.*

<sup>12</sup> Note that it is not crucial to this explanation that Strategies J and K exist independently, only that the external/internal relations they describe are utilized as listeners hear sentences which justify those strategies (such as (81a)).

### Strategy K

In ...  $V_1$   $V_2$ ... (in which the verbs are finite),  $V_2$  corresponds to the main verb of a sentence with  $V_1$  as the subordinate verb.

The relations assumed by Strategies J and K combine to make single-embedded sentences like (81b) quite simple to perceive. But the same strategies make doubly embedded sentences difficult because of Principle I. With respect to the internal relation set up by Strategy J,  $NP_2$  is simultaneously the left-hand and right-hand member of a strategy in double embeddings, while  $V_2$  is simultaneously the right- and left-hand member of Strategy K. Notice that the superficial identity of the three NPs and Vs in an embedded sentence increases the difficulty since it makes the relation between the first and second and second and third constituent absolutely identical. Thus, if  $N_2$  or  $V_2$  differ superficially from their surrounding phrases, sentences like (81) should become easier. I have not tested this, but it seems to me that (88), in which  $N_2$  and  $V_2$  do differ superficially from their adjacent constituents, is relatively comprehensible (compared with (81a)).

- (88) The dog the destruction the wild fox produced was scaring will run away fast.

There are other kinds of examples in language explained by the double-function Principle I. Consider the complexity of the sentences in (89):

- (89) a. They did not want me not to promise not to help them.  
 b. They did not want me to promise not to help them.

- c. John is not not available for no charge at all.  
 d. John is not available for no charge at all.

(89a) and (89c) are examples of triple negation, which has often been recognized as extremely complex, if acceptable at all. Like the embedded sentences (81a, b), sentences with two negation markers are perfectly comprehensible and acceptable (as in (89b, d)). Principle I applies to predict both the difficulty of sentences with three negations and the relative ease of sentences with two negative markers. Consider the perceptual strategy L, which defines the perceptual operation signaled by a negative marker. Strategy L operates to place the second 'not' in the above sentences as both the scope of the first negation and, simultaneously, the operator on the third negation.

### Strategy L

Negation markers (*not*, *un-*, and the like) apply the operation of semantic negation to their syntactically defined scope.

According to Principle I, any sequence with such a double perceptual function is perceptually complex. This principle also explains the difficulty of many so-called “left-branching” structures. Recently Yngve (1960) has proposed that phrases with a left-branching external hierarchical organization (such as (90a)) are harder to produce and understand than phrases with a right-branching organization (such as (90b)). According to this view (elaborated by Johnson 1965) “left-branching” involves greater load on temporary memory than does “right-branching.” This is allegedly due to the number of hierarchical phrase structure “commitments” for the rest of the sentence that are made by the words in a left-branching structure. For example, the word ‘very’ in (90a) “commits” the talker to an adjectival phrase modifying a noun, while the word ‘the’ in (90b) makes no such commitment. Presumably structures involving more commitments are harder to produce, because they require a greater memory load, to ensure that commitments made earlier in a sentence are fulfilled. This model of speech processing is intended to account for the relative difficulty of sentences like (90c) as compared with (90d).

- (90) a. very big boys
- b. the big boys
- c. Coats collars buckles are strong
- d. Buckles of coats collars are strong

This proposal is coherent as a model of complexity in speech production. But for speech perception it appears that one could argue that left-branching structures should be *simpler* to understand if there is any effect at all, just because the increased number of structural “commitments” that the speaker makes early in a sentence should make it easier (that is, more redundant) for the listener to predict the latter part of the sentence. Thus, even on formal grounds, it is not clear that the amount of left-branching in a sentence should correspond to its perceptual complexity. Furthermore, there are many convincing counterexamples. For example, consider the perceptual simplicity of the highly left-branching sentence in (91a):

- (91) a. After a quite severe appendicitis attack the not very well dressed man  
          fell over.
- b. Buckles of collars of coats are strong.

Thus, left-branching as such cannot be used to predict or explain perceptual complexity. Principle I, however, does account for the perceptual complexity of the cases that appeared to support the left-branching hypothesis (90c). Sentences (90c) and (91b) are predicted to be relatively more difficult than the other sentences in (90) and (91) because of the double function of at least one phrase (note that (90c) becomes immediately comprehensible if the word 'coats' is changed to 'fur' as in (92a), and that (91b) becomes comprehensible if the word 'collars' is changed to 'containers', as in (92b)).

- (92) a. Fur collars' buckles are strong.  
b. Buckles of containers of coats are strong.

That is, while the superficial appearance and phrase structures in (92a, b) are identical with the incomprehensible sentences (90c, d), Principle I does not apply to them because the internal relations between the three phrases now differ—the phrase in the middle is not both the left- and right-hand member of the same external/internal perceptual mapping because of the changes in the internal relations.

Principle I also predicts previously unexplored classes of perceptual complexity that are intuitively of the same sort as the previous examples. Consider the sentences in (93). In each case there is a phrase (indicated by “y”) that is related to a previous phrase in the same way that it relates to a following phrase, and in each case, the sentences are extremely difficult to understand if they are acceptable at all.



- (93) a. They were tired of discussing considering producing toys.
- x                      y                      z
- b. They were tired of the discussion of the consideration of the production of toys.
- x                      y                      z

As in the cases of double embedding, triple negation, and “left-branching,” 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 in which it modifies the third phrase. Consider the relative perceptual ease of these sentences if only two phrases occur:

- (94) a. They were tired of discussing producing toys.  
b. They were tired of the discussion of the production of toys.



The sentences in (93) also become much easier to understand if the internal relations among the three critical phrases are varied:

(95) a. They were tired of discussing  $\overset{x}{\text{discussing}} \overset{y}{\text{ceasing}} \overset{z}{\text{producing}} \text{toys}$ .

b. They were tired of the discussion of the evolution of the production of toys.

(Note that in (95a),  $y$  is the internal structure object of  $z$ , while  $x$  and  $y$  are not directly related. In (95b),  $y$  is the action carried out by  $z$ , but the object of  $x$ .) Finally, the sentences in (93) become perceptually simpler if the superficial form of the critical phrases is varied, even while the internal relations are held constant:

(96) a. They were tired of discussing the consideration of producing toys.  
 b. They were tired of the discussion of considering the production of toys.

In brief, I have tried to show that if speech perception is viewed (at least in part) as a direct mapping of external sequences onto internal structures, then the tautology in Principle H predicts Principle I, which in turn predicts the relative perceptual complexity of double embedding, triple negatives, and left-branching sentences among others. Thus, such sentences may be generated as syntactically (and semantically) acceptable, but be hard to understand nevertheless.<sup>13</sup>

## 6 Three conclusions

These discussions of the role of language behavior in determining language “structure” lead to several modifications of current views of the study of language. First, we must reassess the distinction between “knowledge” of a language and its “actual use,” which places emphasis on the independent primariness of “abstract linguistic knowledge.” Second, we must consider whether the acquisition of language systems is best interpreted in terms of the primary acquisition of a series of grammatically defined rules or in terms

<sup>13</sup> Notice that the notion of “double function” in Principle I does not refer to cases in which a given word may both be a subject and an object in the internal structure. That is, it is not clear that sentence (a) is more complex perceptually than sentence (b), even though ‘boy’ in (a) is both a subject and an object, while in (b) it is only a subject.

of the development of the psychological systems underlying perception and memory. Finally, the demonstration that the structural and behavioral systems of language are often special expression of cognitive universals should expand our conception of the innate components of language acquisition; we now must focus on the problem of how the different innate components are linked together in the course of language learning and how the learned aspects are incorporated in adult language behavior.

### 6.1 *The distinction between linguistic competence and performance in the adult (What is linguistics a science of?)*

The goal of a model of speech perception is to specify how we discover the internal structures of sentences from their external form. The review of experimental work in the first part of this discussion was devoted to explorations of the role of speech perception in the structures isolated in linguistic investigations. The conclusion drawn from these studies was that behavioral processes manipulate linguistically defined structures but do not mirror or directly simulate the grammatical processes that relate those structures within a grammar. Such a conclusion invalidates any model for speech recognition that attempts to directly incorporate grammatical rules as an isolable component of the recognition processes.

The first attempts to integrate transformational grammar with speech behavior were largely concerned with the “psychological reality” of the grammatical rules proposed by linguists (reviewed in Miller 1962; Miller and McNeill 1968). Many psychologists viewed transformational grammar as a novel and radical challenge to their experimental skill. They were particularly unwilling to accept the concept of an “abstract” underlying structure because the current psychological theory could not account for its existence. Thus, the conflict between “behaviorism” and “mentalism” reappeared in discussions of language behavior and motivated many experiments.

One product of this conflict has been the artificial distinction between “linguistic theory” on the one hand and “speech behavior” on the other. Linguists defended themselves against the accumulation of inexplicable psychological facts about speech by invoking the distinction between what we know and what we do. According to this view, “psycholinguistics” was a logical adjunct to “linguistics,” on the following formula:

(97) abstract grammar + “performance principles” = actual speech behavior

WHY IS THERE  
SUCH A BIG GAP  
HERE?



Grammar was taken to be an idealized account of our knowledge. The psychologists' problem appeared to be to find general performance principles that would describe how that ideal grammar is used in behavior.

At first it appeared that many of the processes and structures postulated in transformational grammar would provide direct accounts of behavior. For example, Miller and McKean (1964) found that the time to match pairs of sentences with the same internal structure as a function of the transformations that differentiate their external structures; this result appeared to justify the claim that "one linguistically defined transformation corresponds to one psychological operation." Further research at first appeared to back up this simple competence–performance equation, but more recent research (reviewed above) shows that this is incorrect. In point of fact, grammatically defined structures may be reflected in speech behavior, but not grammatically defined processes. Thus we seem to be in a dilemma: how can we account for the psychological validity of linguistically defined *structures* without taking into account the linguistic *processes* that define those structures and their interrelations?

This dilemma is actually an illusion created by the artificiality of the distinction between "competence" and "performance" in grammatical analysis. A real grammar does not, in fact, describe an abstract linguistic world, but rather a set of intuitions about "grammaticality" held by native speakers. For example, the transformational grammarian appeals to an intuition shared by most of us about our language when he claims that he will consider only facts that pertain to complete sentences. We all agree roughly on what a sentence is, and, no doubt, we could define psychological tests that would identify most sentences most of the time. However, even if the agreement on what is (or is not) a sentence were much weaker than it is, the point would remain the same; the linguist uses an introspective behavioral criterion to choose among his intuitions about a language. He assumes that some of the structural distinctions inherent in a grammar are *consistently* reflected in his intuitions about sentencehood, structural relations, ambiguities, and so on. He uses these consistent reflections in his own behavior to decide what data about the language he must describe.

However, even if our linguistic intuitions are consistent, there is no reason to believe that they are *direct* behavioral reflections of linguistic knowledge. The behavior of having linguistic intuitions may introduce its own properties; that is, there is no guarantee that a linguistic grammar itself is either a *direct* or an *ideal* representation of the linguistic structure. I have emphasized that the discovery of the linguistically pertinent data that the grammar describes is itself a poorly understood psychological process. Therefore, a grammar is *not*

necessarily a unique, basic “nonpsychological” representation of linguistic structure: it is merely the most direct and available of all behavioral reflections of grammatical structure.

In short, for the past ten years we have taken the psychology of linguistic intuitions for granted and have used those intuitions as data relevant to the construction of a universal linguistic grammar. Our apparent problem has been to put grammar and psychology back together again. We are finding that it is impossible to do this *directly* according to the simple equation in (97). Instead we find that we have developed two formulae for the interaction of ideal grammar and speech behavior in grammatical analysis:

- (98) a. ideal grammar + behavioral principles of having “linguistic intuitions”  
       = “linguistic data” (such as the facts in (4))
- b. “linguistic data” + formal grammatical universals = “generative grammar”

Thus, to take linguistic grammar itself as the “basic” structure would be to make the same mistake as does the physicist who takes the parallelogram of forces as the “basic” concept of mechanical systems. The parallelogram of forces is itself derived from a special case of more general physical principles; it has its specific properties due to the specific nature of its application to slow-moving bodies on planes. Similarly, a linguistic grammar may have formal properties that reflect the study of selected subparts of speech behavior (for example, having intuitions about sentences), but which are not reflected in any other kind of speech behavior. Other kinds of speech behavior may bring out additional aspects of the structure of language, and they undoubtedly have laws of their own independent of the structure of language, but all the formalizations of systematic speech behavior including grammar must exemplify at least part of the actual linguistic structure.<sup>14</sup>

This conclusion is in conflict with many recent claims about the relationship of a linguistic grammar and behavior. For example, the common view has been expressed in the following quotations from a conference held in 1966 (Lyons and Wales 1969).

A theory of linguistic knowledge attempts to account for our ‘intuitions’ concerning the language . . . (A theory of performance) is a theory of how, given a certain linguistic competence we actually put it to use, realize it, express it (Wales and Marshall, pp. 29–30).

<sup>14</sup> Note that to take one external capacity as the underlying structure for another is to make the same mistake as do those linguists and psychologists who argue that one actual sentence form (for example, “the active”) is central to other sentence forms (for example “the passive”).

If language were a game, ‘competence’ would be the rules of the game, while the actions of its players would constitute performance (Blumenthal, p. 81).

A search for an analysis of the connection between the way the structural description is specified by the grammar and the way it is ‘specified’ by speakers and hearers . . . is one way of formulating the psycholinguistic problem; the abstract nature of this connection between grammar and recognition is (to be) emphasized . . . the problem (is) which aspects of the structural description are relevant to explanations of particular performance tasks. (Garrett, in the discussion of Fodor and Garrett, p. 175)

These authors agree that there is a linguistic grammar that accounts for our basic linguistic intuitions of sequence acceptability, structural relations, and so forth. It is the psychologists’ problem to explore the “behavioral” relevance of the structures internal to a grammar.

I have argued that a proper understanding of the behavioral and phenomenological nature of “basic linguistic intuitions” forces us to reject the claim that a linguistic grammar is in any sense internal to such linguistic performances as talking and listening. To quote Jonkheere (p. 86 in the same volume):

It does not necessarily follow that the characterization of the rules a person is following in some form of rule-conforming behavior has to go into the explanation of how he follows these rules or performs behavior in conformity with them.

The relationship between linguistic grammar based on intuition and that based on the description of other kinds of explicit language performance may not just be “abstract” (as maintained by Fodor and Garrett) but may be *nonexistent* in some cases. First, apparently “linguistic” intuitions about the relative acceptability of sequences may themselves be functions of one of the systems of speech behavior (for instance, perception), rather than of the system of structurally relevant intuitions. Second, the behavior of producing linguistically relevant intuitions may introduce some properties which are *sui generis* and which appear in *no* other kind of language behavior.

In this paper, I have considered examples of the first kind, in which perceptual mechanisms underlie what initially appear to be idiosyncratic syntactic rules and examples in which the unacceptability of “grammatical” utterances is due to perceptual mechanisms. An example of grammatical structures relevant only to intuitions may be the linguistically defined transformations, since they do not themselves play a direct role in sentence perception. If they also play no role speech production they will remain an example of grammatical mechanisms that are relevant *only* to the behavior of having intuitions about sentences.

Once we accept the possibility that ongoing speech behavior does not use a linguistic grammar, it is no surprise that the mechanisms inherent to ongoing

speech behavior do not manifest transformations or any operations directly on them. An explanation of why producing conscious intuitions about potential sentences elicits transformations that are not utilized in other aspects of speech behavior awaits a full theory of the phenomenology of linguistic intuitions. However, there are some aspects of such intuitions that provide an initially plausible basis for the importance of transformations in linguistic grammars based on those intuitions. The set of intuitions about sentences that are unique to modern transformational grammarians are not intuitions of sequence acceptability, but intuitions of structural relations among sentences. For example, the fact that active and passive constructions are felt to share the same basic grammatical relations and are sensed as somehow corresponding to one another is taken as a motivation for describing both as instances of a common internal structure. If the only goal of grammar were to generate acceptable sequences, the motive for generating active and passive as special instances of the same structure, differentiated by one rule, would be much less strong. Furthermore, even most cases of acceptability judgments involve judgments about *potential* sentences, in which one is asked to extrapolate his linguistic knowledge onto imagined situations, which often stimulates the linguist-informant to aid his “grammaticality” judgment about a particular sequence by thinking of other sequences to which it is closely related. Thus the formal description of a language using transformations depends on intuitions that are irrelevant to most ongoing speech behavior but that emphasize transformational relations between sentences.

The question that arises is: What is the Science of Linguistics a Science of? Linguistic intuitions do not necessarily directly reflect the structure of a language, yet such intuitions are the basic data the linguist uses to verify his grammar. This fact could raise serious doubts as to whether linguistic science is about anything at all, since the nature of the source of its data is so obscure. However, this obscurity is characteristic of every exploration of human behavior. Rather than rejecting linguistic study, we should pursue the course typical of most psychological sciences; give up the belief in an “absolute” intuition about sentences and study the laws of the intuitional process itself.

This course of action has been fruitful in other areas of psychology. Consider the subjectivity of astronomical star-transit judgments, which according to Boring was one of the first problems to arise in the context of what we know today as psychology. For a time, astronomers believed in the “absolute” constancy of their judgment of the instant at which a star crossed a certain reference point. However, it was noticed that different observers produced different judgments, so each pair of astronomers were related by a

“personal equation,” which specified the relative delay in their judgments. Ultimately it was observed that even an individual’s judgment delay was not constant, and would vary from situation to situation.

These observations could have been used to justify rejection of any facts based on personal reaction time, and indeed astronomers turned to other timing techniques as soon as they became available. However, the study of reaction time itself became one of the main areas of experimental psychology. Given that reaction times are not absolute or free of the context in which they occur, psychologists have explored systematically the interaction between reaction time and its context.

The effect of stimulus context on absolute judgment of the stimulus has become a part of almost every branch of psychology. One of the most basic laws governing the interaction between stimuli is the *law of contrast*—for example, the well-known phenomenon of feeling that the ocean is cold on a hot day, while the same ocean at the same temperature feels warm on a cool day. That is, one’s “absolute” judgment of a stimulus can be exaggerated by the difference between the stimulus and its context. This influence by contrast clearly can occur in “intuitions” about grammaticality. For example, (99b) preceded by (99a) may be judged ungrammatical, but contrasted with (99c) it will probably be judged as grammatical.

- (99) a. Who must telephone her?  
       b. Who need telephone her?  
       c. Who want telephone her?

That is, not only are there several reasons for the unacceptability of sequences, but even the notion of structural grammaticality is itself subject to contextual contrast. Often the nature of contextual influences on absolute judgments is less clear than in cases of contrast. For example, it is well known that the perception of an unsaturated spot of color is greatly influenced by its surroundings. Thus, surrounded by a yellow background, a pale green spot may appear blue, while the same green spot appears deep green if it is surrounded by red. These differences in judgment are quite stable, in the sense that even with conscious instruction about the nature of the situation, the perception of the colors is still influenced by the surrounding context in the same way.

Cases like this cannot be described as mere “contrast” effects; in what *a priori* sense does red contrast more directly with green than yellow does? Human observers themselves *contribute* this notion of contrast even in the absence of obvious physical parameters to be contrasted (unlike the case of the influence of hot or cold on the perception of lukewarm, in which the

differences and contrasts have an “objective” contrasting measure). In the case of color perception, it is in the nature of our visual system to contrast red and green in one dimension and blue and yellow in another dimension.

It is quite likely that similar situations obtain between sentences, in which judgments of the grammaticality of one sentence are affected by the other sentences among which it is placed, even when the other sentences do not appear to contrast with the stimulus sentence in as direct a manner as in (99). This proposal is subject to demonstration. Take all the example sentences from several linguistic articles (excluding those sets that contrast directly as in (99)) and present them to subjects either in their original sequence, taken separately from each article, or entirely shuffled from all the articles. Subjects must simply indicate which sentences they think the original articles assumed to be grammatical and which were labeled ungrammatical. It would not be surprising if subjects should replicate the judgments of the original articles much more consistently when presented with the examples in their original order than when presented with all the sentences from the different articles in some random order. If this is true, the experiment will demonstrate that the judgments of “absolute” grammaticality are illusory and that a science of the influence of context on acceptability judgments is as necessary in linguistic research as in every other area of psychology.

Such a criticism does not invalidate linguistics, even without reform. Many intuitions about sentences appear to be strong enough to resist contextual effects, and we can expect that these intuitions will remain constant even when we have developed an understanding of the intuitional process (for example, the relationship between actives and passives, the fact that ‘John hit the ball’ is a sentence of English, and the like). However, recent trends in linguistic research have placed increasing dependence on relatively subtle intuitions (cf. Lakoff 1968; Kiparsky and Kiparsky 1970; Ross 1967a; MacCawley 1969) whose psychological status is extremely unclear. Since there are many sources for intuitional judgments other than grammaticality, and since grammaticality judgments themselves can be influenced by context, subtle intuitions are not to be trusted until we understand the nature of their interaction with factors that are irrelevant to grammaticality. If we depend too much on such intuitions without exploring their nature, linguistic research will perpetuate the defects of introspective mentalism as well as its virtues.

## 6.2 *The acquisition of grammar*

Ideally, a model of language learning should specify how the child discovers the systematic relations between internal and external structures of language



used in talking, listening, and predicting potential sentences in his language. This review of language learning has explored the effects of the system of listening (and presumably talking) in the young child on the system of predicting potential sentences in the adult. The existence of this interaction shows that it is not the case that the predictive grammar is learned independently of the use of language; rather, it is learned in the course of its use.

However, many recent studies have been devoted to exploring the child's acquisition of language in terms of his acquisition of rules allegedly independent of their use. A standard methodology is to observe the child's utterances at a given stage and to then write a "miniature grammar" for his utterances; language development is then described as an ordered series of such "grammars."

There are several methodological difficulties with such a program. First, adult grammars are based on a variety of linguistic *intuitions* about sentences, not actual utterances. The "grammar" for what an adult actually says (and what he understands) would undoubtedly look quite different from the grammar that accounts for his intuitions about sentences *in vitro*. Thus, while a description of the maturation of the child's productive (or perceptual) system for language is interesting, it does not bear directly on his acquisition of a system of linguistic knowledge. Second, any finite set of linguistic data about specific sequences justifies an infinitude of grammars. Which grammar is used to generate a particular corpus of data depends on intuitions *about* the acceptable sequences (like the notion of relations between sentences) as well as *a priori* decisions by the linguist as to what theoretical form a grammar must have, and what kinds of intuitions are relevant to his description. Since young children do not present us with their intuitions *about* sentences and intersentential relations, we cannot narrow the range of possible grammars implied by any finite set of their utterances. Furthermore, we cannot use preconceived notions about the form of grammar underlying a child's utterances (such as the assumption that it is initially nontransformational because this would prejudice the sort of fact that we are trying to ascertain by collecting his utterances in the first place.

Suppose, however, that these difficulties with writing grammars for utterances of preschoolers were somehow overcome by finding a way of eliciting "linguistically relevant" intuitions from young children. One would then be able to study the development of the systems for predicting potential sentences. At each point in the child's development one would still have to examine the structure of his other systems of language behavior to assess their interaction with his alleged "linguistic" intuitions. Thus, like an adult, a child may reject a particular sequence as ungrammatical simply because he

cannot understand it. Of course the situation would be more complex than for an adult, even in the study of a child who could articulate his “linguistic” intuitions, since his perceptual and productive systems for language behavior would themselves be evolving and presumably would interact with each other and with the system of predicting new sentences from old.

Suppose, however, that one solved these problems as well, and were able to distinguish the effects on intuitions due to the different aspects of speech behavior. One then might predict that language structure emerges as some function of transformational rules. The most obvious prediction in this vein would be that the more grammatical rules a subgrammar of English has, the later it is acquired. Furthermore, if one holds the view that the grammar of one’s linguistic knowledge is reflected directly in such behaviors as speech production and perception, then one would predict that sentences involving more transformations are processed relatively poorly by young children.

Brown and Hanlon (1970) have made exactly these assumptions and this prediction for the acquisition of language, rule by rule. They examined the relative frequency of various constructions whose linguistic analysis involves different numbers of transformations. They conclude that almost all their predictions based on relative numbers of transformations are confirmed. However, their results also confirm the hypothesis that sentences in which a relatively large amount of material must be recovered from the deep structure are relatively difficult.

First, three of their predictions involve the relative simplicity of declarative sentences compared to negatives or questions. They argue that this is due to the fact that the affirmatives have one less transformation in their derivation. However, *any* theory of speech processing must take into account the psychological primacy of the positive form of utterance, quite independently of the number of syntactic operations involved.

Even if one granted Brown and Hanlon these three cases, there are nine cases that involve specific deletion of internal structure material (for example, the truncated question form as compared with the question). Finally, the remaining predictions involve the negative question construction, which turns out to be relatively difficult for children. Brown and Hanlon argue that this relative difficulty is due to the fact that an extra transformation (negative) is involved. However, there are various linguistic and behavioral arguments that negative questions also involve the deletion of an entire sentence from their internal structure. That is, ‘didn’t the girl hit the dog’ has an internal structure more nearly corresponding to ‘*the girl hit the dog* didn’t the girl hit the dog,’ from which the (italicized) sentence has been deleted by an optional transformation. As behavioral evidence for this

proposal (which is the only relevant kind of evidence to compare with Brown and Hanlon's explicitly behavioral data), consider the fact that negative questions are ordinarily taken to be assertions (for example, the above sentence asserts that the questioner believes that the girl *did* hit the dog) (cf. Mehler and Bever 1968a).

Thus, Brown and Hanlon's results are equivocal concerning the possibility of predicting the effect of the number of formal transformations on the complexity of a syntactic construction type for children; their data could be explained equally well by a view of sentence complexity according to which the more internal structure material that is implicit in the external structure, the harder the sentence, since the child must contribute more information to the sentence himself.

This chapter has concentrated on the interactions between "linguistic" structures and perceptual mechanisms in the child and adult, although it is clear that mechanisms for *learning* a language affect the eventual structure of a language even more dramatically than do the perceptual systems. I have said little about the effects of general principles of learning on linguistic structure because I do not know anything about how language (or anything else) is learned, while I do have some initial understanding of the mechanisms of perception. There is no doubt however that, as we understand more about the learning of language, we will be able to account for even more of the structures that we find in our adult ability to have intuitions about potential sentences.

The claim that languages are learned via a series of subgrammars of the adult language remains to be demonstrated. However, there are certain nongrammatical behavioral variables that we know to affect the learning of language, even though we do not yet understand the learning process itself. The most obvious behavioral constraint on language acquisition is the development of memory in the young child. The child's immediate and long-term memory must constrain his language ability in vocabulary size, utterance length, and amount of material in the external structure of sentences deleted from their internal structure. The fact that the child starts out with a small vocabulary and short utterances is well attested. The recent research by Brown and Hanlon demonstrates further that the child also has some difficulty with constructions that depend on active reconstruction of deleted internal structure.

### 6.3 *The unity of universals of language and thought in the mind*

Recent discussions of language and linguistic theory have emphasized the extent to which the capacity for language is innate in human beings. The

formal articulation of innate language structures is contained in the universal grammar, which represents all of the formal characteristics and constraints that a grammar for a particular language must reflect. For example, the putative universal grammar states that every language has an internal structure, an external structure, and a set of transformations that map the first onto the second; that there are distinct categories for “noun” and “verb”; and so on. Chomsky (1965) suggested that we must distinguish between *formal* and *substantive* universals. Formal universals describe the *types* of rules that are possible (for example, that there are transformations), while substantive universals describe the universally available stock of terms used in languages (for example, that noun and verb are possible syntactic categories). Many substantive linguistic universals appear to be derivable from more general psychological universals, for example, the universality of the noun–verb distinction in language might be explained as the linguistic reflection of the general cognitive distinction between objects and relations between objects (cf. Chomsky 1965: 28). Thus the concept of “noun” would not have to be taken as a linguistic universal in itself but merely as the linguistic expression of such a cognitive universal. The formal universals are also susceptible to immediate explanation as special instances of general cognitive structures. For example, one could argue that there are transformational systems in other areas of behavior; e.g., the systematic set of transformations involved in interpreting a three-dimensional object from a two-dimensional projection of the object.<sup>15</sup>

This paper has explained the way in which behavioral systems affect all linguistic structures. There are many instances in which the “grammatical” structure of adult linguistic intuitions about potential sentences is influenced by the mechanisms of language perception and learning. The isolation of such cases suggests that there are universal constraints on the form of grammars that are not inherent to the statement of universal grammar itself, but rather to the way in which grammar is learned and the use to which it is put. One general example of this sort is a universal constraint on the amount of ambiguity of internal structural relations in sentences. Many languages represent the internal relations by the order of the words in the external structure and a few selected function words (as does English). Other languages allow

<sup>15</sup> For a clear example of a formal universal that reflects general cognitive structures, consider Chomsky’s proposal that it is a formal linguistic universal that “...proper names must designate objects meeting a condition of spatiotemporal contiguity, and that the same is true of other names designating objects” (1965: 29). Surely one could argue that the same principle applies to the visual apprehension of objects, independent of their name.

relatively free ordering but have a rich system of inflections (such as Russian); some languages have both ways of representing the internal relations to some extent. However, languages that have neither and languages that have both systems to a great extent appear to be very infrequent (if they exist at all). The relevant constraint appears to be that a language may not have *too much ambiguity of the internal relations* in the external forms of sentences. This condition is difficult to state formally as part of the universal grammar because it cannot be phrased equivocally—that is, all languages have some internal-structure ambiguity, so a universal grammar cannot rule out such ambiguous derivations entirely. On the other hand, the frequency of such ambiguities must be restricted. Such a restriction can be interpreted as coming about as a natural function of the fact that a language in which every sentence had an indeterminate internal structure (except from context) would not be learned by children. However, such a restriction is not a part of universal grammar but a statement about the universals of language learning.

One might be tempted to conclude from such investigations as these that our problem is how to “subtract out” general cognitive structures from linguistic structures in order to isolate the “pure” linguistic universals, as depicted in (100).

$$(100) \quad (\text{Apparent Linguistic Universals}) - (\text{Cognitive Universals}) = \text{Real Linguistic Universals}$$

Indeed, the arguments in this chapter might be taken as demonstrations that there is not as much innate structure to language as we had thought, if the “universal grammar” is stripped of those aspects that draw on other psychological systems (cf. McNeill 1970, for considerations of just this possibility).

However, such an enterprise fails to take into consideration the fact that the influences of language and cognition are mutual; one cannot consider one without the other. The isolation of cognitive mechanisms that are utilized in language does not explain them away as linguistic structures any more than the fact that we can name abstract concepts explains how we come to have such concepts. The discovery that certain aspects of language are based on mechanisms of perception, learning, and cognition provides us with a new puzzle about how they become integrated in human communicative behavior.

### Bibliography

- Abrams, K. and Bever, Thomas. G. (1969). Syntactic structure modifies attention during speech perception and recognition. *Quarterly Journal of Experimental Psychology*, 21: 280–90.

- Beilin, Harry (1971). The development of physical concepts. In Theodore Mischel (ed.), *Cognitive Development and Epistemology*. New York: Academic Press, 85–119.
- Bever, Thomas G., Fodor, Jerry A., and Garrett, Merrill (1966). Transformational operations and stimulus complexity. Unpublished ms. MIT.
- Bever, Thomas G., Fodor, Jerry A., and Garrett, Merrill (1966). The psychological segmentation of speech. Paper presented at the 18th International Congress of Psychology, Moscow.
- Bever, Thomas G., Lackner, J., and Kirk, R. (1969). The underlying structure sentence is the primary unit of speech perception. *Perception and Psychophysics*, 5: 225–34.
- Bever, Thomas G., Lackner, J., and Stolz, W. (1969). Transitional probability is not a general mechanism for the segmentation of speech. *Journal of Experimental Psychology*, 79: 387–94.
- Bever, Thomas G. and Mehler, Jacques (1968). The coding hypothesis and short-term memory. Unpublished ms. MIT.
- Bever, Thomas G., Mehler, Jacques, and Epstein, J. (1968). What children do in spite of what they know. *Science*, 162: 921–4.
- Bever, Thomas G., Mehler, Jacques, Valian, Virginia, Epstein, J., and Morrissey, H. (in preparation). Linguistic capacity of young children. Ms.
- Blumenthal, A. L. and Boakes, R. (1961). Prompted recall of sentences. *Journal of Verbal Learning and Verbal Behavior*, 6: 614–16.
- Brown, R. and Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. R. Hayes (ed.), *Cognition and the Development of Language*. New York: Wiley, 11–53.
- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam and Miller, George A. (1963). Introduction to the formal analysis of natural languages. In R. D. Luce, R. R. Bush, and E. Galanter (eds), *Handbook of Mathematical Psychology*, vol. II. New York: John Wiley & Sons, ch. 11.
- Clark, H. H. and Clark, E. V. (1968). Semantic distinctions and memory for complex sentences. *Quarterly Journal of Experimental Psychology*, 20: 129–38.
- Fodor, Jerry A. and Bever, Thomas G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, 4.5: 414–21.
- Fodor, Jerry A. and Garrett, Merrill (1966). Some reflections on competence and performance. In J. Lyons and R. Wales (eds), *Psycholinguistic Papers*. Edinburgh: Edinburgh University Press.
- Fodor, Jerry A. and Garrett, Merrill (1967). Some syntactic determinants of sentential complexity. *Perception and Psychophysics*, 2: 289–96.
- Fodor, Jerry A., Garrett, Merrill, and Bever, Thomas G. (1968). Some syntactic determinants of complexity II: Verb structure. *Perception and Psychophysics*, 3.6: 453–61.
- Garrett, Merrill, Bever, Thomas G., and Fodor, Jerry A. (1966). The active use of grammar in speech perception. *Perception and Psychophysics*, 1: 30–2.
- Gough, Philip B. (1966). The verification of sentences: The effects of delay of evidence and sentence length. *Journal of Verbal Learning and Verbal Behavior*, 5.5: 492–6.

- Johnson, N. I. (1965). The psychological reality of phrase-structure rules. *Journal of Verbal Learning and Verbal Behavior*, 4: 469–75.
- Kaufman, E. L., Lord, M. W., Reese, T. W., and Volkman, J. (1949). The discrimination of visual number. *American Journal of Psychology*, 62: 498–525.
- Kiparsky, P. and Kiparsky, C. (1970). Fact. In M. Bierwisch and K. E. Heidolph (eds), *Progress in Linguistics*. The Hague: Mouton: 143–73. (In press when CBLS was published.)
- Lakoff, George (1968). Instrumental adverbs and the concept of deep structure. *Foundations of Language*, 4: 4–29.
- Lakoff George (1970). An example of a descriptively inadequate interpretive theory. *Linguistic Inquiry*, 1.4: 539–42.
- Langacker, Ronald W. (1969). On pronominalization and the chain of command. In David A. Reibel and Sanford A. Schane (eds), *Modern Studies in English: Readings in Transformational Grammar*. Englewood Cliffs, NJ: Prentice-Hall, 160–86.
- Lyons, J. and Wales, R. J. (eds) (1969). *Psycholinguistics Papers*. Edinburgh: Edinburgh University Press.
- MacCawley, J. (1969). The role of semantics in a grammar. In E. Bach and R. Harms, *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston, 125–70.
- Martin, J. E. (1968). *A Study of the Determinants of Preferred Adjective Order in English*. Unpublished doctoral dissertation. University of Illinois at Urbana-Champaign.
- McCarthy, D. (1954). Language development in children. In L. Carmichael (ed.), *Manual of Child Psychology*. New York: Wiley, 492–630.
- McMahon, Lee (1963). *Grammatical Analysis as Part of Understanding a Sentence*. PhD dissertation. Harvard University.
- McNeill, O. (1970). The development of language. In P. A. Mussen (ed.), *Carmichael's Manual of Child Psychology*, 3rd edition. New York: Wiley, 1061–161.
- Mehler, Jacques (1963). Some effects of grammatical transformations on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, 2: 346–51.
- Mehler, Jacques and Bever, Thomas G. (1968a). Quantification, conservation, and nativism: Reply by J. Mehler and T. G. Bever. *Science*, 162.3857: 979–81.
- Mehler, Jacques and Bever, Thomas G. (1968b). The study of competence in cognitive psychology. *International Journal of Psychology*, 34: 273–80.
- Mehler, Jacques and Carey, P. (1968). The interaction of veracity and syntax in the processing of sentences. *Perception and Psychophysics*, 3: 109–11.
- Miller, George A. (1962). Some psychological studies of grammar. *American Psychologist*, 17: 748–62.
- Miller, George A. and Chomsky, Noam (1963). Finitary models of language users. In R. D. Luce, R. R. Bush, and E. Galanter (eds), *Handbook of Mathematical Psychology*, vol. II. New York: John Wiley & Sons, 419–91.
- Miller, George A. and McKean, Kathryn (1964). A chronometric study of some relations between sentences. *Quarterly Journal of Experimental Psychology*, 16: 297–308.

- Miller, George A. and McNeill, David (1968). Psycholinguistics. In G. Lindze and E. Aronson (eds), *Handbook of Sociology*. Boston: Addison-Wesley, 666–705.
- Miller, George A. and Selfridge, Jennifer A. (1950). Verbal context and the recall of meaningful material. *American Journal of Psychology*, 63: 176–85.
- Phillips, J. and Miller, G. A. (1966). An experimental method to investigate sentence comprehension. Unpublished paper. Harvard Center for Cognitive Studies.
- Ross, John R. (1967a). *Constraints on Variables in Syntax*. PhD dissertation. Cambridge, Mass.: MIT.
- Ross, John R. (1967b). On the cyclic nature of English pronominalization. In John Robert Ross, R. Rusev, Nicolas Ruwet, et al., *To Honor Roman Jakobson: Essays on the Occasion of his Seventieth Birthday, 11 October 1966*. The Hague: Mouton: 1669–82.
- Savin, H. and Perchonock, E. (1965). Grammatical structure and the immediate recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, 4: 348–53.
- Schlesinger, I. (1966). The influence of sentence structure on the reading process. US Office of Naval Research Technical Report 24.
- Slobin, Dan I. (1966). Grammatical transformations and sentence comprehension in child and adulthood. *Journal of Verbal Learning and Verbal Behavior*, 5: 219–27.
- Turner, E. A. and Rommetveit, R. (1967). The acquisition of sentence voice and reversibility. *Child Development*, 38: 649–60.
- Vendler, Zeno (1968). *Adjectives and Nominalizations*. The Hague: Mouton.
- Wales, R. J. and Marshall, J. C. (1966). The organization of linguistic performance. In J. Lyons and R. Wales (eds), *Psycholinguistic Papers*. Edinburgh: Edinburgh University Press, 29–80.
- Yngve, Victor H. (1960). A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society*, 104.5: 444–66.