9 The psychological reality of grammar: a student's-eye view of cognitive science

Thomas G. Bever

Never mind about the why:
figure out the how, and the why will take care of itself.
Lord Peter Wimsey

Introduction

This is a cautionary historical tale about the scientific study of what we know, what we do, and how we acquire what we know and do. There are two opposed paradigms in psychological science that deal with these interrelated topics: The behaviorist *prescribes* possible adult structures in terms of a theory of what can be learned; the rationalist *explores* adult structures in order to find out what a developmental theory must explain. The essential moral of the recent history of cognitive science is that it is more productive to explore mental life than prescribe it.

The classical period

Experimental cognitive psychology was born the first time at the end of the nineteenth century. The pervasive bureaucratic work in psychology by Wundt (1874) and the thoughtful organization by James (1890) demonstrated that experiments on mental life both can be done and are interesting. Language was an obviously appropriate object of this kind of psychological study. Wundt (1911), in particular, gave great attention to summarizing a century of research on the natural units and structure of language; (see especially, Humboldt, 1835; Chomsky, 1966). He reported a striking conclusion: The natural unit of linguistic knowledge is the *intuition* that a sequence is a sentence. He reasoned as follows:

All of the cited articles coauthored by me were substantially written when I was an associate in George A. Miller's laboratory, first at Harvard University and then at Rockefeller University. That is no accident.

We cannot define sentences as sequences of words because there are single-word sentences (e.g., "Leave").

We cannot define sentences as word uses that have meaningful relations because there are meaningful relations within certain word sequences that, nevertheless, are not sentences (e.g., "the days of the week").

Hence, the sentence must be defined as a sequence that native speakers of a language intuitively believe to convey a complete proposition in a linguistically acceptable form.

At the outset, this framed the problem of linguistic description as a problem of linguistic knowledge – to describe what speakers of a language know when they know a language. Wundt's formal analyses of this knowledge summarized a tradition of continental research on local and exotic languages. Most important was the assignment of purely abstract syntactic structures to sentences, independent of their meaning. Among these structural features were several levels of representation, which expressed grammatical relations between words and phrases. At a surface level, a set of hierarchically embedded frames symbolized the relative unity of word sequences grouped into phrases. For example, in sentence (1), "the" is clearly more related within a unit to "Rubicon" than to "crossing"; similarly, in sentence (2), "was" is closer to "crossed" than to "Rubicon." The surface level also defines a set of surface grammatical relations between the phrases: In sentence (1), "Caesar" is the grammatical subject, namely, the phrase that determines the morphological agreement with the verb. In sentence (2), the corresponding grammatical subject is "The Rubicon." In sentence (3), it is the entire act, "Crossing the Rubicon."

(1) Caesar was crossing the Rubicon.
(2) The Rubicon was crossed by Caesar.
(3) Crossing the Rubicon was what Caesar did.

It was obvious that the propositional relations between the phrases could not be captured by the surface grammatical relations. As Wundt noted, "Caesar" is the acting one in each of these cases. The propositional relations between phrases must be represented by a separate level, the "inner form" of the sentence. At this level, sentences (1) to (3) share the same relations between "Caesar" (agent), "cross" (action), and "Rubicon" (object). The different actual sequences at the surface grammatical level were related to the propositional level by processes of transformation *(umwandlung)*; these processes reorder surface phrases in surface patterns allowed by the particular language. The propositional relations were not purely semantic, but were the formal expression of relations between semantic units of meaning (Blumenthal, 1970; 1975).

The continental model of language was rich and made many claims about the capacity of humans to manipulate abstract entities. But, alas, the theory never

became an object of experimental importance. The reasons are, no doubt, politically and scientifically complex. One sufficient fact is that Wundt classified the study of language as a branch of social psychology, and hence, for him, not a subject fit for experimental investigation. His vast structural catalog of language is more an anthropological survey rather than a scientific scrutiny of the mind. The linguistic model and its richness became lost to scientific psychology.

But the model was not lost to everbody interested in language. It was popularized in the infant field of linguistics by a young professor of German, Leonard Bloomfield. Bloomfield's (1914) enthusiastic exegesis of Wundt's multileveled model might have installed it as the basis for the newly emerging science of language. However, in all social sciences at the time, there was a growing preoccupation with behaviorist injunctions against unobservable entities and relations. Clearly, the notions "intuition" and "inner grammatical level" were not acceptable in the operationalist framework. Bloomfield capitulated to such restrictions as enthusiastically as he had earlier espoused the Wundtian model. His foundational book, *Language* (1933), presents a totally operationalist framework for linguistic theory. In that book, the sentence is hardly mentioned, while meaning is given an amusing (and brief) treatment in terms of the association of stimuli and responses. The interdict had fallen.

The dark ages

Behaviorism is a seductive doctrine that has dominated psychological theories of learning for the majority of this century. It is seductive because it *simultaneously* purports to answer three questions:

What do we learn?
How do we learn?
Why do we learn?

According to behaviorism, the reason that we learn is that the environment provides pleasure when we do – that is, it reinforces our activity. Selective reinforcement accounts for the way we learn; it associates environmentally successful pairs of behaviors and situations as permanently learned units. Accordingly, what we learn must be expressed in terms of definable pairs of behaviors and situations. These principles provide a chain of inference from the motive to learn back to the structure of what is learned, a chain that has fettered generations of psychologists. The functional importance of reinforcement justifies the scientific investigation of isolated sources of pleasure and displeasure. The focus on behavior/situation pairs licenses investigation of the learning of meaningless associations between situations and behaviors. The requirement that learned associations are between definable entities transfers to the operationalist requirement that theoretical terms must be inductively reducible to observable entities.

By the late 1950s, sophisticated elaborations of these principles had crystallized, most notably in the proposal by Hull (1943) that they could account for hierarchically organized chains of behavior. Even when transferred to descriptions of language behavior (Osgood, Suci, & Tannenbaum, 1957), the basic behaviorist doctrine about the structure of what was learned remained: It must offer recognizable links between isolatable situations and behaviors (Fodor, 1966; Bever, 1968). The implications of this doctrine for theories of language were severe. Consider how these procedures affected theories of the relationship between two levels of description, words and phrases. Following Bloomfield's conversion, linguists had adopted the behaviorist restrictions on how to pursue the analysis of language structure. They imposed on themselves a set of *discovery* procedures that would guarantee the scientific acceptability of linguistics and the learnability of the resultant linguistic descriptions. Language was to be described in a hierarchy of levels of learned units such that the units at each level can be expressed as a grouping of units at an intuitively lower level. The lowest level in any such hierarchy was necessarily composed of physically definable units. For example, sentences (4)–(7) could all be resolved to a basic sequence of the same kinds of phrases, a noun phrase, a verb, and an adjective phrase.

(4) Harry was eager.
(5) The boy was eager.
(6) The tall boy was eager to leave.
(7) He was something.

The behaviorist principles demanded that phrases not be free-floating, abstract objects: Each must be reducible back to a single word that could serve as a lexical substitute for the entire phrase, as in sentence (7). In this way, "phrases" were rendered theoretically as units that could be resolved as "words." At the same time, the description gave an account of the fact that words within phrases seem to be more closely related to each other than across phrases. Finally, it seemed possible to hope for a description of all possible types of phrases, since longer ones seemed to resolve into combinations of shorter ones, which in turn could resolve to single words.

This behaviorist implementation of linguistics may seem harmless enough, but it had a particular result: The theoretical notion of the *sentence* could not be described by linguistic theory. This is true for three empirical reasons: First, the number of sentences seemed unfathomably large; second, in a single sentence, there are often relations between words in different phrases; there are grammatical relations between phrases, which cannot be described as superficial facts about the phrases themselves. In addition, as Wundt had noted, it is methodologically impossible to define sentences operationally. To deal with such phenomena as these, linguistic theory would have required levels of representation and theoretical entities that could not be resolved by reduction to independently ob-

servable units. Most behaviorist linguists were sufficiently aware of these problems to leave the description of the sentence alone; the reducible phrase was the pinnacle of behaviorist linguistics.

One unconventional linguist attempted to apply the operationalist descriptive principles to sentences. From the operationalist standpoint, the sentence is a component of a complete discourse in the same sense that a phrase is a component of a sentence. Harris (1957, 1958) capitalized on this perspective: He developed a descriptive scheme in which sentences (and clauses) that can occur in the same discourse frame are reduced to canonical sentence forms. This scheme depends on the fact that sentences occur in structural families: For example, sentences (1)–(3) are part of a larger set of constructions, as follows:

Caesar crossed the Rubicon.
It is the Rubicon that Caesar crossed.
What Caesar did was cross the Rubicon.
The Rubicon is what Caesar crossed.
. . . that Caesar crossed the Rubicon . . .
. . . Caesar's crossing the Rubicon . . .
. . . the Rubicon's being crossed by Caesar . . .
. . . the crossing of the Rubicon by Caesar . . .

Harris noted that all of these variants can often co-occur in the same discourse environment – that is, all of them can be substituted for the blank in the following discourse frame (ignoring changes needed to accommodate the clausal variants):

Caesar marched north.
Then _____ .
This surprised the local inhabitants.

The notion of co-occurrence was intended to be the same as that describing the substitutability of phrases in discourse-based sentence frames. The difference is that unlike phrases, sentences cannot be reduced to canonical words ("it," "did"). Rather, Harris suggested that they be reduced to a *standard canonical sentence* form, the "kernel" of each structural sentence family. Kernel sentences are the simple declarative construction; for example, co-occurrence "transformations" express the relation between the kernel sentence and its variants. The kernel and the passive sentence (2) are related by the co-occurrence transformation, as follows:

"NP_1, V + ed by NP_2 '$\longleftrightarrow$' NP_2 was V ed by NP_1"

There are several important points to retain about this theory. First, the co-occurrence transformations can only relate specific observable sentences. (Harris actually apologized for using abstract terms of the previous kind, but argued that this was only a descriptively temporary notion, not a true violation of operationalist purity; see Katz & Bever, 1975.) Second, the relative priority of the kernel

sentence form had an intuitive appeal but still did not unambiguously meet the requirement that it be both observable and operationally definable. Finally, it was inordinately difficult to make the program work in detail. In retrospect, one can view co-occurrence transformational grammar as an insightful attempt to describe sentence structure within a taxonomic paradigm. The failures of the attempt were illuminating and set the scene for the later developments in linguistic theory, to which I return after reviewing progress in psychology of the day.

The middle ages

For several decades, psychological research on language proceeded largely without explicit benefit of linguistic analysis. The beginnings were modest: the study of the processing of unrelated word sequences to discover how universal laws of learning apply to human beings. This deceptively simple paradigm, "verbal learning," became the focus of an enormous amount of research (many hundreds of articles, with several journals devoted to little else – see Underwood & Schulz, 1960; Cofer, 1961; Cofer & Musgrave, 1963, for representative reviews). Enormous effort was devoted to exploring the formation of associations between adjacent and remote words in strings, the influence of different kinds of practice, of different kinds of reinforcement, of subjects' age, mental capacity, and so on. These studies might appear to be about language as well as learning because they used words as stimuli. But they were about words only coincidentally; words were merely handy units that humans could learn to string together in unrelated ways (much later, it became clear that even this was an illusion, but that is another story; Anderson & Bower, 1973). The focus was on learning, motivation, and memory, not language. Of course, one could have viewed this as propaedeutic to an understanding of how words are processed when they are organized in whole sentences. Unfortunately, this promise was never realized. As we shall shortly see, words in whole sentences are processed differently from words in random sequences.

Just as wholes are the natural enemies of parts, gestalt psychology is the natural enemy of associationist behaviorism. Crucial demonstrations had long been available that percepts have higher-order structures that cannot be accounted for by merely associating the parts. Linguistic structures would seem to have been prime candidates for gestalt investigations: Sentences are wholes that bind together and transcend words and phrases as their parts (Lashley, 1951; Mandler & Mandler, 1964). Such obviously true observation rarely stops associationists from going about their business, and had absolutely no impact on the prediction that the study of verbal learning would lead to an understanding of language (Skinner, 1957). The failure of the gestalt demonstrations to undercut associationistic accounts of language was due in part to the inability of gestalt psychol-

ogy to develop a generative theory of possible gestalten; in any domain, most "principles" of how gestalten are formed seemed true but inexplicable. Furthermore, gestalt psychologists themselves had little interest in language, since to them it seemed obvious that language was both abstract and learned, and therefore not the proper object of their investigation (Koffka, 1935). Once again, a methodological preconception barred a potentially fruitful approach to language.

Until the 1950s, linguists and psychologists worked apart, though sharing the fundamental theoretical restrictions of behaviorism. An early burst of psycholinguistics occurred when the two groups discovered that they could mate each others' theories: learning theory was allegedly capable of describing the acquisition of "behavioral hierarchies" of just the type that linguists had found to be the ultimate grammatical aspect of language, namely, words-in-phrases (Osgood & Sebeok, 1954). The first mating between psychology and linguistics was intense but sterile. The reason was that the two disciplines were mutually compatible just because they shared the same behaviorist preconceptions: The psychologist was willing to postulate of the language learner only the inductive capacity to learn what the linguist had already restricted himself to describe. Yet the shared descriptive restrictions robbed linguistic theory – and psychology – of the sentence. The project of the first psycholinguistics – to show how linguistic structures followed from psychological laws of learning – was successful, brilliantly – and pyrrhically.

There was a separate stream of research that emerged during the same period – the study of how language behavior is organized in adults, independent of any theoretical preconceptions about learning (Miller, 1951a, 1951b). It was unarguable that at some point in the understanding of the sounds of spoken language, listeners arrive at an abstract conceptual analysis of its meaning and structure, but it was still arguable that the meaning conveys the structure and not the reverse. Miller and Selfridge (1950) demonstrated the behavioral relevance of sentence structure, in particular, by showing that memory for the word sequences improves as they approach the statistical regularities of English. This suggested that language behavior involves a transfer of the physical signal into a linguistic world that can access regularities of structure and meaning. The perceptual question was formulated in terms of a search for *the* units of speech perception in which the acoustic-to-linguistic transfer takes place from an acoustic to a linguistic representation.

A standard experimental method to investigate this question was based on an everyday fact: Spoken language is extremely resistant to acoustic interference. Imagine yourself next to a large waterfall: Even though the noise is tremendous, you are able to understand somebody talking to you, so long as the conversation is in your language. The question is, why? Clearly, you are using your knowledge of the language to aid your perception. But which aspect of your linguistic

knowledge do you access first? A straightforward hypothesis is that you have memorized the *words* of your language. In this view, the unit of transfer from acoustic to linguistic information in speech perception is the word: A listener first maps the acoustic signal onto separate words and then computes the meaning from them.

The proposal that the unit of speech perception is the word may seem daunting, since there are so many of them. But one thing we know: People do learn thousands of words in their language. Since the number of effectively necessary words is finite (though large), it is possible to imagine that they are the learned basis for speech perception. A laboratory-controlled variant of the speech-by-waterfall experience offered a technique to test the word hypothesis (Miller, Heise, & Lichten, 1951; Miller & Isard, 1963). Suppose we adjust the loudness of a noise source relative to recordings of the words in sentence (8) so that each word is recognized correctly 50 percent of the time when the words are heard alone. If it is the word level at which the acoustic information is mapped onto linguistic representations, then a sentence composed by stringing together the same words should be perceived 25 percent of the time. This follows from the hypothesis that the acoustic shape of each word is mapped independently onto a linguistic representation. The actual facts are striking: When strung together into a sentence, the word sequence is often recognized much more than 50 percent of the time. Most important is the intuitive fact that the words seem acoustically clearer *as you hear them* when they are in a sentence.

(8) Horses, eat

The outcome of a series of such studies was the demonstration that it is at least the *sentence* that is the unit of speech perception. Even sentences that do not make semantic sense enhance perception of the words in them.

This finding created two problems that still dominate investigations of language: How do we use our knowledge of sentences in behavior? What do we know when we know the sentences of our language? The first issue is a direct example of the problem of integrating abstract knowledge with concrete behavior. For decades, psychologists had assumed that beliefs cannot affect perception except in very limited cases. This belief was consistent with the behaviorist strictures on what we can learn: If beliefs can influence perception, then what the child learns cannot be limited to what is in the world to be perceived. The apparent influence of abstract sentence structure on acoustic decisions [and other kinds of sensory judgments (Miller, Bruner, & Postman, 1954)] suggested that perceptual activity involves the *simultaneous* integration of abstract and concrete levels of representation.

There are two views how this integration occurs: Abstract levels of knowledge can directly influence concrete levels ("top-down processing") or can interact

only after lower-level representations are formed ("bottom-up processing"). Since it was first discovered, this issue has been vigorously pursued in language perception, because language offers well-defined, distinct levels of linguistic structure. The pendulum of scientific fashion has swung back and forth, with current investigators in favor of a bottom-up theory of perception (Forster, 1979; Fodor, 1983; Pylyshyn, 1984). This issue may never be entirely resolved. The clear experimental demonstration of the behavioral importance of sentence level discredited behaviorist restrictions on theories of language comprehension. More enduringly, the issue has been the engine for decades of empirical and theoretical research.

Another matter was raised by the acoustic clarity of sentences: How can something as varied as the sentence be the unit of transfer from acoustic to linguistic representations? Unlike the number of words, there is no meaningful upper bound to the number of sentences; hence, when we listen, we cannot transfer information from acoustic to linguistic structure sentence by sentence. Clearly, we must know a system of rules for describing the sentences in our language, and we must apply that categorical knowledge actively during speech behaviors such as comprehension.

This conclusion from experiments on adult behavior presented an unanticipated challenge to the behaviorist doctrine about learning. The sentence level of knowledge plays an active role in perception, yet that level cannot be described or acquired according to behaviorist principles. By ignoring the problem of learning altogether and focusing on adult behavior, it was discovered that adults use representations that cannot be learned by induction: People actively use categorical rules in ongoing behavior. This seemed typical of many behaviors once some thought was given to the question (Miller, Galanter, & Pribram, 1960). Behaviorism was surely doomed as a theory of language, but the final fall awaited a viable theory of what kind of knowledge generates the sentences of a language.

What *is* a sentence? Harris had attempted to include the sentence level within an operationalist framework, but his success was limited by the discovery procedures. In the late 1950s Chomsky (1957a) offered a new kind of grammar as a formal answer to this question: A sentence is what the grammar describes as a sentence. The motivations and details of this new theory were similar in many respect to those summarized by Wundt (though without awareness of this similarity at the time) (Chomsky, 1957b). The configuration of the syntactic model that described sentence structure was also similar to that of much earlier times and included a reformulated notion of transformation. But the new transformational grammar had novel formal devices, as well as a completely new goal for grammatical description. The grammar was generative: It described sentence structures of a language as a natural part of human knowledge. The new approach flatly rejected operationalist discovery procedures and allowed abstract

terms and rules that were not reducible to observable entities. This represents in linguistics the same kind of shift away from behaviorist principles of learning that was occurring within psychology. Chomsky was also diligent in pointing out the general failures of behaviorist accounts of language (Chomsky, 1959). The goal of linguistic analysis was to describe an adult's linguistic knowledge, not language behavior: The staple facts that reflect that knowledge are intuitions about the acceptability of sentences in one's language. Hence, masses of data relevant to linguistic analysis are easy to collect; one merely consults a native speaker on how he feels about a language sequence. By assigning sentence status to only those sequences that are intuitively acceptable, the grammar constitutes a theory of the speaker's underlying linguistic knowledge.

The most important feature of the new syntactic model was that several levels of representation are included in the description of every sentence. Obviously, only one of these levels is directly observable. This level, the surface structure, corresponds roughly to the kind of phrase structure Wundt described for surface grammatical relations, as well as that arrived at by behaviorist linguists. Every sentence at this level was paired with an underlying "kernel" structure, which presented the propositional relations for each verb in a canonical form.

A set of transformations specified the possible pairs of deep and surface phrase structures. Unlike co-occurrence transformations, these transformations operate unidirectionally. Each transformation applies a successive deformation of the input tree – for example, changing a structure that will end up as an active into one that will end up as a passive:

tree 1 =) tree 2

Certain transformations combined kernel trees into complex sentences like sentences (9) and (10) (⟨⟩ represents a position into which a complement sentence is to be inserted):

(9) John defied Bill to go.
(10) John desired Bill to go.

The way in which different kernel sentences are combined reveals interesting differences in the deep-structure organization of sentences that are superficially similar. For example, in the complex sentences with

John past defy Bill ⟨⟩ = =) John past defy Bill to go
Bill go

John past desire ⟨⟩ = =) John past desire Bill to go
Bill go

"desire" and "defy" appear to have superficially identical structures. But they have different corresponding structural variations, which attests to their under-

lying distinctness. For example, the complement sentence can be passivized as a whole unit with "desire" but not with "defy." Contrast your intuitions about the acceptability of the following two sequences:

(11) For Bill to go was desired by John.
(12) For Bill to go was defied by John (unacceptable).

The fact that the complement sentence can act as a whole unit in passive constructions demonstrated that at the deep structure level the entire complement sentence is a direct object of verbs like "desire," whereas only the following noun phrase is the direct object of verbs like "defy."

Similar kinds of reasons demonstrated the difference at the underlying structure level of other superficially identical sentence structures:

(13) John was eager to please.
(14) John was easy to please.

John past BE eager ⟨⟩
John please someone ==⟩ John was eager to please

It past BE easy ⟨⟩
Someone please John ==⟩ John was easy to please

Thus, knowing the syntax of a language consisted of knowing a set of phrase structure rules that generate the underlying kernel structures and the transformations that deform and combine those underlying structures into sentences. The kernel structures were not representations of meaning of sentences but of their underlying syntactic configuration. The meaning of a sentence was a function of semantic operations on underlying structure representation (Katz & Postal, 1964; Chomsky, 1965); the sound was produced by operations on the surface representation.

Linguists made a firm point of insisting that, at most, a grammar was a model of "competence" – what the speaker knows. This was distinguished from "performance" – how the speaker implements his knowledge. But, despite this distinction, the syntactic model had great appeal as a model of the processes we carry out when we talk and listen. It offered a precise answer to the question of what we know when they know the sentences in their language: We know the different coherent levels of representation and the linguistic *rules* that interrelate those levels. It was tempting to postulate that the theory of what we know is a theory of what we do, thus answering two questions simultaneously:

What do we know when we know a language?
What do we do when we use what we know?

The first answers were couched in terms of the linguistic model and a direct assumption about how the model was linked to behavioral systems. We know

two systems of syntactic rules and the levels of representation that they describe; phrase/structure rules define underlying kernel structures, and transformations derive surface phrase structures from the kernels. This knowledge is linked to behavior in such a way that every syntactic operation corresponds to a psychological process. The hypothesis linking language behavior and knowledge was that they are identical.

The way to test the direct linking hypothesis was to demonstrate the "psychological reality" in ongoing behavior of linguistic structures. Surface phrase structures were taken to be noncontroversial because they were allegedly definable within the behaviorist framework. (In fact, they probably are not; Bever, 1968b; Bever, Foder, & Weksel, 1968.) But the claim that *deep* structures were psychologically "real" was a direct challenge to behaviorism: Deep structures were obviously abstract in the sense that they are not actual sentences, but rather the "inner form" to which Wundt referred. Note that even the simplest kernel structure ("he (past leave)") is not an actual sentence. A set of morphological transformations always applies to even the simplest sentences (e.g., changing "he past leave" to "he left"). The further claim that transformations were real mental processes was an additional challenge, both because the rules are themselves abstract and because they define intermediate levels of phrase structure representations as they do their transformational work.

The first step was to define a structural paradigm that would generate a family of studies of the psychological validity of the grammar as a mental model. Three optional transformations defined eight sentence types in a three-dimensional cube (Miller, 1962b). Sentences can be either active or passive, declarative or interrogative, positive or negative. In linguistic theory of the day, each of these dimensions corresponded to transformation. Accordingly, in testing the behavioral implications of the formal relations between these sentences as arrayed on the transformational cube, one could test the psychological relevance of the grammatical model.

The initial results were breathtaking. The amount of time it takes to produce a sentence, given another variant of it, is a function of the distance between them on the transformational cube (Miller & McKean, 1964); transformational distance between sentences also predicts confusability between them in memory (Mehler, 1963; Clifton, Kurcz, & Jenkins, 1965; Clifton & Odom, 1966). Furthermore, the ease of memorizing the sentences was predicted by the number of transformations that have applied to them: Simple declaratives are easier than passives, which are easier than passive questions, and so on (Mehler, 1963). Finally, such transformationally complex sentence forms as passives were more resistant to acoustic distortion (Compton, 1967), took longer to comprehend than corresponding actives (McMahon, 1963; Gough, 1965, 1966), and put a greater load on immediate memory (Savin & Perchonock, 1965).

It is hard to convey how exciting these developments were. It appeared that there was to be a continuing direct connection between linguistic and psychological research: Linguistic analysis would support structural analyses, which would directly become hypotheses for investigation in language behavior. Abstract models of linguistic structure and performance could give clear direction and critical importance to empirical research (Chomsky & Miller, 1963; Miller & Chomsky, 1963). The hypothesized link between the structure of knowledge and the processes of behavior was wildly successful. The golden age had arrived.

The enlightenment

It soon became clear that either the linking hypothesis was wrong or the grammar was wrong, or both. The support for the psychological relevance of transformations had been based only on those three that defined the transformational "cube." But the overall program implied a broad empirical hypothesis about the relation between all rules in a grammar and sentence processing: the so-called derivational theory of complexity. (The complexity of processing a sentence corresponds to the number of transformations in its description.)

It had long been known that the derivational hypothesis was wrong for many constructions. As a predictive principle, it was both too strong and too weak (Fodor & Garrett, 1967; Bever, 1970). It was too weak because it failed to predict the obvious complexity of numerous kinds of sentences.

(15) Center embedding: *the oyster the oyster the oyster split split split.* (Contrast to the structurally identical, but easier, *the reporter everyone I met trusts predicted the coup.)*

(16) Object relative clauses without relative pronouns: *the horse raced past the barn fell.* (Contrast to the structurally identical, but easier, *the horse ridden past the barn fell.)*

It was too strong because it incorrectly predicted that various sentence constructions with more transformations are harder to understand than corresponding sentences with fewer transformations:

(17) Heavy noun phrase (NP) shift: *We showed the long-awaited and astoundingly beautiful pictures of the Himalayan trip to Mary.* is transformationally less complex than the easier *We showed Mary the long-awaited and breathtakingly beautiful pictures of the Himalayan trip.*

(18) Extraposition: *That Bill left early with Mary surprised Hank.* is transformationally less complex than the easier *It surprised Hank that Bill left early with Mary.*

These direct demonstrations of the inadequacy of the derivational theory of complexity were backed up by several experiments that examined the implications for perceptual difficulty of transformations other than those that defined the transformational cube. For example, the transformation that optionally moves a

particle/preposition to the position following the verb was well motivated as starting with the *Verb+particle NP* deep structure and then being transformed to the *Verb-NP-prt* sequence, as shown subsequently. This order of structures is indicated by the fact that verbs are lexically associated with some particle/prepositions and not others (e.g., we have "call up" and "call NP over" but not "call NP under"). If such verbal sequences are entered as complex lexical items and *then* transformed, we can capture these lexical regularities.

John called up Mary $==>$ John called Mary up
verb prt NP verb NP prt

Despite this clear motivation for treating the *Verb-prt-NP* variant as less complex transformationally, such sentences turned out to be processed more slowly than the corresponding transformed versions (Bever & Mehler, 1967). Other studies of specific transformations also failed to show that perceptual difficulty corresponds to the number of transformations (Bever, Fodor, Garrett, & Mehler, 1966; Jenkins, Fodor, & Saporta, 1965; Fodor & Garrett, 1967).

Such failures were baffling in light of the initial success of the structures arranged on the transformation cube. The failures motivated reconsideration of the theoretical interpretation of the three-dimensional model. This further consideration revealed that if the grammatical theory were taken literally, it would not have motivated many of the original predictions. First, the linguistically provided route from one construction to another is not along the surface of the three-dimensional cube; rather, it must involve returning to the underlying kernel structure and then reemerging to the target structure. For example, the grammatically defined pathway from the negative-passive to the active construction is not provided by undoing the negative and then the passive; rather, it is provided by undoing the morphologically necessary transformations, the passive transformation, and then the negative transformation, to recover the kernel structure: then that structure must have the morphological transformations reapplied to it to produce the declarative sentence. Each time two sentences are related, it must be by way of a return to the inner kernel structure and a new path out to the target sentence structure. This strict interpretation of the grammar had the consequence that confusions between two complex sentences adjacent on the transformation cube (e.g., between the negative passive and the passive question) would be far less likely than between two simpler structures (the negative and the question). Yet, this was not confirmed. Sentences were confusible as a function of their adjacency on the *surface* of the sentence cube. This fact had to be reinterpreted in light of the lack of motivation due to a strict implementation of the grammar. One possibility was that their language cube was a real but temporary representation that subjects themselves constructed to deal with the temporarily repetitive experimental problem set for them: The repeated presentation

of a small number of sentence types stimulated their arrangement on a cube (Bever, 1970a). This offered an explanation of why small variations in response technique could change the evidence for the transformational cube. For example, the frequency of decisions that a sentence is appearing for the first time is low if a transformational variant of it appeared recently (Clifton & Odom, 1966). The amount of interference, however, appears to be a function of the similarity of the two sentences at the underlying structure level: For example, active and passive constructions mutually interact more than active and negative constructions. This is explained by the fact that the "negative" is marked in the deep structure, as well as occasioning a transformational rule (Katz & Postal, 1964; Chomsky, 1965).

It was also necessary to understand why the relative difficulties of comprehension and perception were predicted by the transformational distance from the kernel structure on the cube. First, two of the transformations actually changed the meaning of the sentence: question and negation. It is not surprising, upon reflection, that such sentences might be relatively hard to process precisely because of their semantic complexity (i.e., the negative has the same propositional structure as the declarative, with an additional semantic operator). The importance of semantic analysis was emphasized by the finding that passive constructions are relatively complex *only* when they are semantically reversible (Sobin, 1966; Walker, Gough, & Wall, 1968; e.g., "Vercingetorix was crossed by Caesar" is more difficult than its corresponding active construction, but "The Rubicon was crossed by Caesar" is not).

The conclusion of such demonstrations, and of further theoretical analysis, was that there is no direct evidence for grammatical transformations in speech behaviors. The behavioral relevance of grammatical levels of representation fared better. Surface phrase structure can be construed as making the behavioral claim that the members of a phrase have a stronger bond to each other than to the members of other phrases. This was directly tested by Levelt (1970). He demonstrated that surface phrase structure can be derived by hierarchical clustering analyses of subjects' reports of "relatedness" among the word pairs of a sentence (see Chapters 4 and 8 of this volume). Accordingly, words are better prompts for other words in the same surface phrase than in different phrases (Johnson, 1965; Suci, Ammon, & Gamlin, 1967; Stewart & Gough, 1967).

Other kinds of studies demonstrated that surface phrase structure is imposed on sentences as we perceive them (Mehler & Carey, 1967; Mehler, Bever, & Carey, 1970). The surface structure pattern was dramatically revealed in the perceived location of nonspeech sounds that interrupt a sentence. Such sounds characteristically are reported as having occurred in a phrase break, especially between clauses, rather than in their actual location (Fodor & Bever, 1965; Garrett, 1966). This effect is not due to guessing biases (Garrett, 1965; Holmes,

1970; Seitz, 1972; Bever & Hurtig, 1974) or to variation in word sequence probability (Bever, Lackner, & Stolz, 1969). Carefully designed materials also showed that this effect is not due to local intonational cues; for example, the italicized sequence in sentences (20) and (21) is common to both syntactic organizations (Garrett, Bever, & Fodor, 1966). Yet, the perceived location of the nonspeech sound objectively at the slash is determined by the syntactically driven clause boundary (marked by the # for each structural version). As the authors of this particular study put it, the use of surface phrase structure to organize speech during perception is "active." This was a convincing demonstration of what was foreshadowed by the studies of the comprehension of speech-in-noise: Grammatical knowledge is actively deployed in a way that clarifies the speech signal.

(20) In her *hope of marrying An/na was impractical*
#

(21) Harry's hope of *marrying An/na was impractical*
#

The most tendentious hypothesis in generative grammar was that every sentence has a deep structure representation, which is abstract. This hypothesis was the subject of greatest debate in linguistic circles because it was the most striking challenge to behaviorist principles. On the one hand, deep structure is not directly observable in the surface sequence; on the other, it is not definable in semantic terms. Rather, it is a formal structure that mediates between surface form and meaning. Several psychologists proposed that deep syntactic structure was, in fact, the "code" in which sentences are organized and retained (Mehler, 1963; Miller, 1962b). The "coding hypothesis" was just formulated in terms of a "schema plus correction" model for memory based on Bartlett (1932). Sentences are retained in the simple declarative form, plus transformationally motivated "tags" that express the actual construction of the sentence (passive, negative, etc.) (Mehler, 1963; Miller, 1962a). This theory was originally based on a literal interpretation of the language cube in which the kernel structure was taken to be the simple declarative sentence. There was further support for the interpretation that the coding hypothesis applied to abstract underlying structures as defined in syntactic theory. *Ceteris paribus*, perceived relatedness between words is greater when they are in the same deep structure phrase [Levelt, 1970; e.g., "John" and "eager" would be perceived to be more closely related than "John" in sentence (13) and "easy" in sentence (14)]. Words are better prompts for other words in the same deep structure phrase than in other phrases (Walker, Gough, & Wall, 1968; Davidson, 1969). In fact, a word is a better prompt for recall of the entire sentence if it is represented several times in the deep structure representation [e.g., "John" would be a better prompt in sentence (13) than in sentence (14), since only in sentence (13) is it the subject of two deep structure predicates; Blumenthal, 1967; Blumenthal & Boakes, 1967; Wanner, 1968].

The nonspeech interruption studies offered evidence that the deep structure is actively computed during speech perception. The perceived location of nonspeech sounds is influenced by the "clause" structure at the *deep* structure level of representation, as shown in sentences (22) and (23) (Bever, Lackner, & Kirk, 1969). See the preceding discussion of sentences (4) and (5). The slash indicates the objective location on the nonspeech sound, and the # indicates the relatively dominant perceived location of the sound.

(22) The general defied the t/roops to fight
#

(23) The general desired the t/roops to fight
#

Thus, the levels of representation proposed by transformational grammar were vindicated as active participants in sentence perception; even the deep structure level of representation is actively computed as part of comprehension. But it seemed clear that the grammatical rules that defined these levels of representation did not systematically correspond to psychological operations. In that case, how are grammatical levels of representation elicited during speech processing?

One possibility is that the representation-forming processes of perception are entirely distinct from grammatical operations (Fodor & Garrett, 1966; Fodor, Bever, & Garrett, 1974). In one version of this view, listeners acquire an extragrammatical set of *perceptual strategies* that map surface structures onto deep structures (Bever, 1968a, 1970a). These strategies are not grammatical rules but state relations between levels of representations based on salient features at the surface level. The most powerful of these strategies (in English) is that the surface sequence "NP-verb-NP" corresponds to "agent verb object." That strategy gives a nontransformational explanation for the one fact that unambiguously had supported the hypothesis that transformations correspond to psychological operations: Passives are harder than actives. This is explained by the fact that passives violate the NP-verb-NP strategy. Similarly, the strategy explains the salience of NP-verb-NP structures as exemplified in sentence (16). It also explains the preference for the transformed version of sentences with particle/prepositions, and the transformed version places the object directly after the verb.

Research on sentence comprehension in children confirmed the behavioral independence of such perceptual strategies. During their third year, children rely heavily on this strategy for comprehension, although there is no concomitant change in their linguistic knowledge (Bever, Mehler, Valian & Epstein, 1969; Bever, 1970; Maratsos, 1972; DeVilliers & DeVilliers, 1972; Slobin & Bever, 1980). The relation between grammatical rules and the processes of speech production seemed similarly remote: Both intuition and speech errors suggest that speaking involves mapping underlying structures onto surface structures (Fodor, Bever, & Garrett, 1974; Garrett, 1975). But the mapping mechanisms them-

selves are not grammatical rules: Speakers appear to talk by inserting words into syntactically defined frames (Garrett, 1975; Bever, Carroll, & Hurtig, 1975).

Distinguishing grammatical rules from psychological strategies stimulated investigation of how the latter interact with other psychological processes, such as attention, perception, and memory. There are capacity limits that require immediate memory to be cleared periodically for new input (Miller, 1957; Miller, Galanter, & Pribram, 1960). The perceptual strategies can clear immediate memory by recoding surface sequences on deep structure propositions; this motivated the hypothesis that the proposition is the unit of recoding during speech perception (Bever, 1970a; Bever, Kirk, & Lackner, 1969; Fodor, Bever, & Garrett, 1974). This proposal gives special status to the end of each proposition, since it is there that definitive recoding can take place. In fact, just at the end of clauses, reaction times to clicks are slow (Abrams & Bever, 1969), detection of clicks is poor (Bever, Hurtig, & Handel, 1974), tones are hard to discriminate (Holmes & Forster, 1972), and evoked potentials are suppressed (Seitz, 1973), whereas the relative magnitude of orienting responses to shock suggest greatest preoccupation at that point (Bever et al. 1969). The loss of attention capacity was ostensibly due to the mental load associated with the final stage of recoding the sentence into a deep representation (Abrams & Bever, 1969; Bever, Garrett, & Hurtig, 1973). At first, it was argued that the surface clause was the relevant unit that defined the scope of the perceptual strategies (Fodor & Bever, 1965); then the deep structure "sentence" (Bever, Lackner, & Kirk, 1969). Finally, it became clear that the relevant unit as a psychological object was in part defined as a "functionally complete" proposition (Carroll, 1964; Carroll & Tanenhaus, 1975).

The decade of research on speech processing between 1965 and 1975 offered a complete account of the relation between linguistic theory and behavioral systems: Grammar defines the levels of representation, but ordinary behavior depends on statistically valid strategies. Grammatical rules may find behavioral instantiation, but only as a backup system slowly brought into play on the rare cases when the behavioral strategies fail (Bever, 1972).

The moral of this experience is clear. We made progress by separating the question of what people understand and say from how they understand and say it. The straightforward attempt to use the grammatical model directly as a processing model failed. The question of what humans know about language is not only distinct from how children learn it, it is distinct from how adults use it. In retrospect, this should not have been a surprising result. It is a philosophical truism that there is a difference between knowing *that* X and knowing *how to* X – knowing that a sound sequence is an arpeggio on a French horn is quite different from playing one. Musical knowledge may inform both performers and listeners about the structure inherent in their shared experience, but the knowledge

does not describe the actual experiential processes. The same distinction is true of linguistic knowledge.

Further consideration also suggests a straightforward functional reason why grammars are not models of specific kinds of language behaviors: There are too many classes of language behavior, each with its own neurological and physical constraints. In particular, humans ordinarily both talk and comprehend; the constraints on the ear are obviously different from those on the mouth. A grammar represents exactly those aspects of language that are true, no matter how the language is used. The representation of such knowledge must be abstracted away from any particular system of use.

The most positive result of this exercise was the demonstration of the importance of abstract levels of linguistic representation during language behavior. It definitively rejected a behaviorist model of language learning, which cannot account for the incorporation of such abstract structures. It also offered the hope of a new kind of gestalt psychology, one in which the relevant gestalten would be given a theoretical foundation; the grammar can be called on to define the "good figures" of language (Neisser, 1967). The golden age was tarnished, but there was a solid prospect for a period of normal science in which abstract mental structures could be taken as the object of serious inquiry.

Modern times

Behaviorism springs eternal. New theories in both linguistics and experimental psychology proposed arguments against the notion of a syntactic deep structure. Generative semanticists in linguistics grappled with the problem of the relationship between deep structure and semantic representations. Their argument was that if a shared deep structure represents the relation between an active and a passive construction, a shared deep structure can underlie both an active and a causative structure (e.g., expressing the structural relation between "John killed Bill" and "John caused Bill to die"); or it can underlie an active and a superficially remote construction ("I'm telling you, that what happened was that the thing that John caused Bill to do was changed from being alive to being not alive"). Generative semanticists also noted technical problems with the transformational syntactic model that focused on its relation to semantic structure. They eventually proposed that the semantic representation of a sentence *is* its deep structure (Lakoff, 1971a; Lakoff & Ross, 1976; McCawley, 1968; Postal, 1967). In this view, there was no intermediate, purely syntactic configuration that represented the inner grammatical structures. Rather, they were viewed as an arbitrary subset of the semantic representation (Lakoff, 1972a). The entire semantic representation itself served as the input to the transformations, which derived surface phrase structures from it.

This theory had great appeal and caused considerable controversy among syntacticians. It appeared to simplify linguistic description, and to do so without recourse to a purely formal level of syntactic representation, but this position was not tenable: Semantic representations themselves either must be stated in a normal form or must comprise all conceptual knowledge. Either way, the generative semantics program collapses. If the semantic representation is in a canonical, propositional, or some other normal form, the model still includes an intermediate formal level of representation that mediates structurally between thoughts and the outer form of language (as in McCawley, 1976). If semantic representations are purely conceptual, they must include all possible human knowledge (Lakoff, 1971a, 1972a, 1973). This conclusion was a *reductio ad totam*, since it is impossible to define grammatical rules for all possible knowledge (Katz & Bever, 1975). For this and other reasons generative semantics was largely abandoned as a grammatical project (see Katz & Bever, 1975; Newmeyer, 1980, for general discussions).

The wars in linguistics highlighted another problem: Linguistic theory changes like quicksilver. Psychologists think they have their hands on it, but it slips through. The rapid development of linguistic theory is one of the reasons it is such an influential discipline in behavioral science. It is also one of the reasons that merging linguistics and experimental psychology is difficult. It takes a month to develop a new syntactic analysis of a phenomenon; it takes a year to develop an experimental investigation of it. All too often, the psychologist is in the position of completing an arduous series of empirical studies, only to discover that the linguistic theory underlying them is no longer operative. During the 1970s, syntactic theory received particularly great attention in linguistics; the transformational model evolved and alternative models emerged. The rejection of the early transformational grammar as a literal model of processing had discouraged psychologists from attending to linguistic theory. The sudden multiplicity of syntactic theories confirmed the psychologists' suspicion that linguists were hopelessly fickle.

For all that, psychologists were not to be left out of attempts to develop a behaviorist revival of language. The salient target was the deep structure coding hypothesis, which gives priority to deep syntactic structure as the code for remembered sentences. The obvious alternative hypothesis is that people actually remember sentences in terms of semantic schemata. Many experimental demonstrations of the syntactic coding hypothesis seemed interpretable as degenerate cases of semantic rather than syntactic coding; for example, "John" in sentence (13) is not only the subject of two deep structure sentences, it is also the agent of two propositional "ideas." There were also positive demonstrations of the importance of nonsyntactic representations. Sentences in a story are misrecalled or gratuitously imputed to the story in a way consistent with the story line (Brans-

ford & Franks, 1971, 1972) – a story originally with sentences about a robbery, shots, and dying may be incorrectly recalled as having included a sentence about killing. Such results, and the apparent disarray in linguistic theory, encouraged many to assume that linguistics-based psychology of language was an adventure of the past. In fact, some took this discouraging thought to its behaviorist conclusion: Syntax is not real; only probabilistic language behavior exists. What appears to be evidence of syntax is the gratuitous by-product of perceptual processes (Clark & Haviland, 1974), speech behavior learning strategies (Bates, 1977; Bates & MacWhinney, 1982), or an overimaginative and flawed linguistic methodology (Schank, 1973, 1975).

Nonetheless, the same crucial empirical problem remained, both for linguistics and for psychology: Sentences have behaviorally relevant inner forms, phrase-grouped representations that are neither conceptual nor superficial. Independently of how they are behaviorally integrated with conceptual world knowledge, these structures play a role in the formal description of linguistic knowledge, aspects of sentence comprehension, and the concrete analysis of language learning. One can see some clear trends in the evolution of syntactic theory that may lead to a renewed collaboration between psychology and linguistics. The overwhelming shift has been away from restrictions expressed in transformations to verb-based structural restrictions. Three syntactic models have emerged in this direction (see Sells, 1985, and Tanenhaus, 1986, for a review of current linguistic and psycholinguistic issues, respectively).

(1) Government and binding theory: The original transformational model, with the following changes from the original:

Increase in the importance of lexical information in determining sentence structure.
Reduction in the variety of transformations to one movement rule.
Introduction of explicit "filters" that constrain the distinct levels of representation and relations between the levels of representation.

In the original transformational model, the transformations themselves specified the relation between the deep and the surface level; a derivation is possible only if there are particular transformations that construct it. In the current variant, there is only one transformational rule, essentially a universal movement rule that randomly changes the deep tree; constraints on possible trees then filter out potentially ungrammatical structures (Chomsky, 1981, 1985).

(2) Lexical functional grammar: A separate model (which may be a notational variant of the preceding model) treats most transformational variations as specified within the lexicon itself (Bresnan & Kaplan, 1983).

(3) Finally, a recently introduced variant of phrase structure proposes to describe sentence structure. Unlike earlier models of phrase structure, this model in-

cludes an enriched repertoire of the kinds of abstract constituent nodes that are intended to overcome the empirical inadequacy of earlier phrase structure systems (Gazdar, Klein, Pullum, & Sag, 1985).

Each of these three syntactic models has its champions as a model of speech behavior: Crain and Fodor (1985) for generalized phase structure grammar; Bresnan and Kaplan (1983) for lexical functional grammar; Frazier (1985) and Freedman and Forster (1985) for government and binding theory. The common argument is that the only thing wrong with the previous attempts to take the syntactic theory as a psychological model was that the syntactic theory itself was wrong. Yet the old difficulties persist; consistent linking assumptions between the formal model and the behaving person are hard to state. A consistent theory of direct implementation of grammatical knowledge in behavior continues to elude us. (See Bever, 1987, for a discussion of current attempts to formulate direct linguistic linking assumptions.)

Back Again?

Several recent developments in cognitive psychology and the study of language suggest a certain recycling of previous approaches. First, concomitant with the increasing emphasis on lexically driven grammatical analyses, a large body of research in the last decade has been devoted to "lexical access," the recognition, immediate processing and recall of single words in various contexts, including sentences (see Simpson, 1981; Seidenberg & Tanenhaus 1986; and Tanenhaus, Carlson, & Seidenberg, 1985, for reviews). Sentence-level syntax offered few structures for a psychologist to study after establishing the behavioral relevance of abstract levels of representation. A language has a small number of constructions that can act as crucial cases, for example, paired constructions like sentences (13) and (14) or (22) and (23), which are identical in their surface representation but different at the underlying level; or paired constructions like sentences (24) and (25), which share a substantial acoustic segment, with distinct surface phrase structures.

But a language has a large number of words. It will take a long time to run out of things to study about them. And it appears that such work will always be relevant; surely, whatever we find out today about processing words will be relevant to a future integrated theory of language behavior. Unfortunately, scientific history tells us that this is not necessarily so; the only people who we can be absolutely sure will profit from our scientific work, brilliant or parochial, are the future historians of science. Otherwise, we would today be profiting from the decades of research on verbal learning. Today's studies seem to offer more hope;

they typically (but not always) involve the relation among words in sentences. They sometimes focus on the relation between lexical levels and other levels of processing. Such studies offer an unequivocal answer to the question of how much of sentence processing is controlled by categorical information in words. But many lexically focused studies are done entirely without the benefit of a more general theory of language behavior; the word has returned, and it is not always good.

This raises a separate but intellectually coordinated development, now in its first blush of high fashion in cognitive psychology: "connectionist" models of complex behavior (Feldman, 1986; McClelland & Rumelhart, 1986). Connectionism is a descriptive term for a class of theories based on associative connections between elements. There are few constraints on how abstract the units themselves can be or how many clusters they are arranged in. In this sense, connectionism is associationism without behaviorism. How the modern technological version will fare remains to be seen. Practitioners of the art take the laudable stance that the theories should be precise enough to be testable. But "testability" is easily conflated with "modeled," and much energy is given to instantiating specific proposals in computer programs. This is intriguing and guarantees a theory's local consistency, but by no means guarantees its correctness. Superficially, successful computational models can be the undoing of the science they purport to aid. The ability of a model to deal with a large but manageable number of isolated facts has little to do with having a correct theory. In fact, factual success can unnaturally prolong the life of an incorrect theory. As the unknown wag says, if Ptolemy had had a computer, we would still believe that we are the center of the universe.

One of the things we know about words is that a language has a manageably small number of them. Hence, studies of lexical processes and connectionist models go well together. Such associative models can even simulate the effect of morphological combination rules (Rumelhart et al., 1986). This success tempts one to speculate that the consequences of such linguistic rules are represented in an associative network without any actual representation of the rules themselves: That is, the system *is* the rule (see Rumelhart & McClelland, 1968, chap. 18; Feldman, 1986). This is a conflation of competence and performance opposite to the conflation of transformational grammar and a performance model, discussed previously. Twenty-five years ago, psycholinguists claimed that the grammar is the performance system; now the corresponding connectionist claim would be that the performance system is the grammar. The grammarian's claim was wrong because categorical grammatical knowledge cannot explain the continua of actual behavior. The connectionist claim may turn out to be wrong for the corresponding reason: Contingent probabilities cannot explain categorical knowledge. Of course, given enough associative connections between explicit and internal

elements, any rule-governed behavior can be approximated. Such models are the equivalent of the sculptor's plaster of paris: Given them an articulate structure and they will shape themselves to it. Associative simulations do not explain morphological structure in lexical morpheme combinations; they only conform to it as an automatic consequence of simulating a description of the combinations. If there is anything we know about sentences, it is that a language has an *un*manageably large number of them. This contrast was one of the strong motivations for embarking on the study of the acquisition and deployment of rules in behavior rather than memorized connections and automatic elicitations (Miller, Galanter, and Pribram, 1960). No doubt connections and elicitations exist and dominate much measurable behavior in an adult. The perennial puzzle is how to prestructure associations so that they systematically *impose* rulelike organization on what they learn. The specific challenges for an associative theory of language learning remain: What is the nature of the categorical information encoded in words, sentences, and rules? How is such categorical information learned and represented? Last (but not least), how does it come to be used and translated into contingent connections? (See Lachter & Bever, 1988.)

There is a recent development in one interpretation of grammar that appears to reunite the study of syntax with a theory of language acquisition. The child acquires the rich internal structure of his or her language from scanty data (Chomsky, 1965), a fact that implies that language structure must be deeply constrained to accord with the child's specific language capacity. This underlies the recent development of so-called learnability theory – an attempt to construct a model of language acquisition that explains specific linguistic rules and constraints on the grounds that they make language learnable (Wexler & Culicover, 1980; Osheron, Stob, & Weinstein, 1982; Pinker, 1984).

Such functionalist proposals are innocuous as a selectional model of learning. But if they are taken to constrain the form of possible linguistic theories, we will have come full circle to the methodological position on linguistic analysis that characterized behaviorism. To be sure, current syntactic models are not constrained by either behaviorist or associative descriptive principles. But we must be warned by our own history: Theories of learning are not good prospects as theories of knowledge.

Conclusion: the mystery of structure

This has been a tale about the merits of being methodologically flexible and answering only one question at a time. Progress was made in understanding how language behavior works by divorcing it from the question of how it is learned. Further progress depended on making a distinction between the models of linguistic knowledge and language behavior. Despite all of this progress, however,

we are left with an essential empirical puzzle, the answer to which will involve solutions to most of the issues raised in this chapter:

How does syntactic structure improve the acuity of the ear?

Note

I have received comments from many linguists and psychologists on an earlier version of this chapter. Generally, psychologists find it an accurate picture of the intellectual period; linguists tend to think that it overlooks much of what was happening within linguistics at the time. The typical comment is: "But linguists were not confused about that matter (e.g., the derivational theory of complexity), so why report the psychologists' confusion over it?" I am not convinced that most linguists were – and are – not just as confused as most psychologists on most matters. However, such responses prompt me to add an explanatory caveat: This chapter is intended to give an account of how the world looked to psychologists and psycholinguists at the forefront of the field during the years of 1955–75. A parallel history of developments within linguistics would be of parallel interest – especially with respect to how linguists viewed language behavior. For example, during the period covered in this chapter, almost all linguists disdained research on the role of syntactic and semantic structures in language processing. This situation has changed somewhat in the last decade, though not always with a clarifying effect, as mentioned at the end of the chapter.

References

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

Anderson, J. R., & Bower, G. H. (1973). *Human Associative Memory.* Washington, D.C.: V. H. Winston.

Bartlett, F. C. (1932). *Remembering.* Cambridge: Cambridge University Press.

Bates, E. (1977). Comments on the paper by Culicover and Wexler. In P. Culicover, T. Wascow, & A. Akamagian (Eds.), *Formal Syntax.* New York: Academic Press.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (Eds.), *Language Acquisition: The State of the Art.* New York: Cambridge University Press.

Berry, R. (1970). A critical review of noise location during simultaneously presented setences. Ph.D. dissertation, University of Illinois.

Bever, T. G. (1968a). A survey of some recent work in psycholinguistics. In W. J. Plath (Ed.), *Specification and utilization of a Transformational Grammar: Scientific Report Number Three.* Yorktown Heights, N.Y.: Thomas J. Watson Research Center, International Business Machines Corporation.

Bever, T. G. (1968b). Associations to stimulus–response theories of language. In T. R. Dixon & D. L. Horton (Eds.), *Verbal Behavior and General Behavior Theory.* Englewood Cliffs, N.J.: Prentice-Hall, pp. 478–94.

Bever, T. G. (1970a). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the Development of Language.* New York: Wiley.

Bever, T. G. (1970b). The integrated study of language. In J. Morton (Ed.), *Biological and Social Perspectives in Language.* Logos Press, pp. 158–206.

Bever, T. G. (1972). The limits of intuition. *Foundations of Language, 8,* 411–12.

Bever, T. G. (1973). Serial position and response biases do not account for the effect of syntactic structure on the location of brief noises during sentences. *Journal of Psycholinguistic Research, 2,* 287–8.

Bever, T. G. (1974). The ascent of the specious of there's a lot we don't know about mirrors. In David Cohen (Ed.), *Explaining Linguistic Phenomena*. Washington, D.C. Hemisphere, pp. 173–200.

Bever, T. G. (1987). Three paradigms for the study of language and cognition. In M. Macken et al. (Eds.), *Proceedings of the SCLI Workshop on Language and Context*.

Bever, T. G., Carroll, J. M., & Hurtig, R. (1976). Analogy of ungrammatical sequences that are utterable and comprehensible are the origins of new grammars in language acquisition and linguistic evolution. In T. G. Bever et al. (Eds.) *An Integrated Theory of Linguistic Ability*. New York: Crowell, pp. 149–82.

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

Bever, T. G., Fodor, J. A., & Weksel, W. (1965). Is linguistics empirical? *Psychological Review, 72*, 493–500.

Bever, T. G., Fodor, J. A., and Weksel, W. (1968). A formal limitation of associationism. In T. R. Dixon & D. L.Horton (Eds.), *Verbal Learning and General Behavior Theory*, Englewood Cliffs, N.J.: Prentice-Hall.

Bever, T. G., Garrett, M. F., & Hurtig, R. (1973). Ambiguity increases complexity of perceptually incomplete clauses. *Memory and Cognition, 1*, 279–81.

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

Bever, T. G., Hurtig, R., Handel, A. (1976). Analytic tasks stimulate right-ear processing in listening to speech. *Neuropsychologia, 14*, 174–81.

Bever, T. G., Kirk, R., & Lackner, J. (1969). An autonomic reflection of syntactic structure. *Neuropsychologia, 7*, 23–8.

Bever, T. G., Lackner, J. R., & Kirk, R. (1969). The underlying structures of sentences are the primary units of immediate speech processing. *Perception and Psychophysics, 5*, 225–31.

Bever, T. G., Lackner, J. R., & Stolz, W. (1969). Transitional probability is not a general mechanism for the segmentation of speech. *Journal of Experimental Psychology, 79*, 387–94.

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

Bever, T. G., Mehler, J., Valian, V., Epstein, J., & Morrissey, H. (1969). Linguistic capacity of young children. (Unpublished.) Referred to in J. Fodor, T. G. Bever, & M. F. Garrett (Eds.), *The Psychology of Language*. New York: McGraw-Hill, p. 502.

Bloomfield, L. (1914). *The Study of Language*. New York: Henry Holt.

Bloomfield, L. (1933). Prompted recall of sentences. *Journal of Verbal Learning and Verbal Behavior, 6*, 203–6.

Blumenthal, A. L., (1970). *Language and Psychology*. New York: Wiley.

Blumenthal, A. L. (1975). A reappraisal of Wilhelm Wundt. *American Psychologist, 30*, 1081–8.

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

Bransford, J. D., Barclay, J., & Franks, J. J. (1972). Sentence memory: Constructive vs. interpretive approach. *Cognitive Psychology, 3*, 193–209.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2*, 331–50.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations for comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11*, 717–26.

Bresnan, J. (1978). A realistic transformational grammar. In M. Halle, J. Bresnan, & G. Miller (Eds.), *Linguistic Theory and Psychological Reality*. Cambridge, Mass.: MIT Press.

Bresnan, J., & Kaplan, R. (1983). Grammars as the mental representation of language. In J. Bresnan (Ed.), *The Mental Representation of Language*. Cambridge, Mass.: MIT Press.

Carey, P. W., Mehler, J., & Bever, T. G. (1970). Judging the veracity of ambiguous sentences. *Journal of Verbal Learning and Verbal Behavior, 9,* 243–354.

Carroll, J. (1967). *Language and Thought.* Englewood Cliffs, N.J.: Prentice-Hall.

Chomsky, N. (1957b). *Syntactic Structures.* The Hague: Mouton.

Chomsky, N. (1959). Review of "Verbal Behavior." *Language, 30,* 26–58.

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

Chomsky, N. (1966). *Cartesian Linguistics.* New York: Harper & Row.

Chomsky, N. (1967). The formal nature of language. In E. Lennenberg (Ed.), *Biological Foundations of Language,* New York: Wiley, 397–442.

Chomsky, N. (1975). *Logical Structure of Linguistic Theory.* New York: Plenum.

Chomsky, N. (1981). *Lectures on Government and Binding.* Dordrecht, Holland: Foris.

Chomsky, N. (1985). *Knowledge of Language.* New York: Praeger.

Chomsky, N. & Miller, G. A. (1958). Finite state languages. *Information and Control, 1,* 91–112.

Chomsky, N. & Miller, G. A. (1963). Introduction to the formal analysis of natural languages. In D. Luce, R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology.* New York: Wiley.

Clark, H. H. (1965). Some structural properties of simple active and passive sentences. *Journal of Verbal Learning and Verbal Behavior, 4,* 365–370.

Clark, H. H., & Haviland, S. E. (1974). Psychological processes in linguistic explanation. In D. Cohen (Ed.), *Explaining Linguistic Phenomena.* Washington, D.C.: Hemisphere.

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

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

Cofer, C. N. (Ed.) (1961). *Verbal Learning and Verbal Behavior.* New York: McGraw-Hill.

Cofer, C. N. & Musgrave, B. S. (Eds.). (1963). *Verbal Behavior and Learning: Problems and Processes.* New York: McGraw-Hill.

Compton, A. J. (1967). Aural perception of different syntactic structures and length. *Language and Speech, 10,* 81–87.

Crain, S., & Fodor, J. D. (1985a). How can grammars help the parser? In D. Dowty, L. Kartunnen, & A. M. Zwicky, (Eds.). *Natural language parsing: Psychology, computational and theoretical perspectives.* London: Cambridge University Press.

Crain, S. & Fodor, J. D. (1985b). Rules and contraints in sentence processing. *Proceedings of the 15th New England Linguistic Society, 15,* Amherst, Mass.: New England Linguistic Society.

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

DeVilliers, P. A., & DeVilliers, J. G. (1972). Early judgments of semantic and syntactic acceptability by children. *Journal of Psycholinguistic Research, 1,* 299–310.

Epstein, W. (1961). The influence of syntactical structure on learning. *American Journal of Psychology, 74,* 80–5.

Feldman, J. A. (1986). *Neural Representation of Conceptual Knowledge.* Cognitive Science Technical Report, Rochester, N.Y.: University of Rochester.

Fodor, J. A. (1965). Could meaning be an rm? *Journal of Verbal Learning and Verbal Behavior, 4,* 73–81.

Fodor, J. A. (1983). *The Modularity of Mind.* Boston: MIT Press.

Fodor, J. A., & Bever T. G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior, 4,* 414–20.

Fodor, J. A., Bever, T. G., & Garrett, M. R. (1974). *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar.* New York: McGraw-Hill.

Fodor, J. A., & Garrett, M. F. (1967). Some syntactic determinants of sentential complexity. *Perception and Psychophysics, 2,* 289–96.

Forster, K. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*. Hillsdale, N.J.: Erlbaum, pp. 27–85.

Frazier, L. (1985). Modularity and the representational hypothesis. *Proceedings of the 15th Northeast Linguistics Society, 15*. Amherst, Mass.: New England Linguistic Society.

Frazier, L. (1986). The mapping between grammar and processor. In I. Gopnick & M. Gopnick (Eds.), *From Models to Modules: Studies in Cognitive Science from the McGill Workshops*. Norwood, N.J.: Ablex Press.

Freedman, S. E., & Forster, K. I. (1985). The psychological status of overgenerated sentences. *Cognition, 19*, 101–32.

Garrett, M. F. (1965). Syntactic structures and judgments of auditory events. Ph.D. dissertation, University of Illinois.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation*. New York: Academic Press, pp. 133–77.

Garrett, M. F., Bever, T. G., & Fodor, J. A. (1966). The active use of grammar in speech perception. *Perception and Psychophysics, 1*, 30–2.

Gazdar, G., Klein, E., Pullum, G., & Sag, I. (1985). *Generalized Phrase Structure Grammar*. Cambridge, Mass.: Harvard University Press.

Gough, P. B. (1965). Grammatical transformations and speed of understanding. *Journal of Verbal Learning and Verbal Behavior, 4*, 107–11.

Gough, P. B. (1966). The verification of sentences. The effects of delay of evidence and sentence length. *Journal of Verbal Learning and Verbal Behavior, 5*, 492–6.

Harris, Z. (1957). Co-occurrence and transformation in linguistic structure. *Language, 33*, 283–340.

Harris, Z. (1958). Linguistic transformations for informational retrieval. Washington, D.C.: *Proceedings of the International Conference on Scientific Information*, 2, NAS-WRC.

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

Hull, C. L. (1943). *Principles of Behavior*. New York: Appleton-Century-Crofts.

Humboldt, W. V. (1835/1903). *Ueber die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*. In *Gessamelte Schriften*, Vol. 7. Berlin: Koeniglich Preussische Akademie der Wissenschaften.

James, W. (1890). *Principles of Psychology*. New York: Henry Holt.

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

Johnson, N. F. (1965). The psychological reality of phrase structure rules. *Journal of Verbal Learning and Verbal Behavior, 4*, 469–75.

Katz, J. J., & Bever, T. G. (1975). The fall and rise of empiricism. In T. G. Bever, J. Katz, & D. T. Langendoen (Eds.), *An Integrated Theory of Linguistic Ability*. New York: T. Y. Crowell.

Katz, J. J. & Postal, P. M. (1964). *An Integrated Theory of Linguistic Descriptions*. Cambridge, Mass.: MIT Press.

Koffka, K. (1935). *Principle of Gestalt Psychology*. New York: Harcourt, Brace & World.

Koplin, J. H., & Davis, J. (1966). Grammatical transformataions and recognition memory of sentences. *Psychonomic Science, 6*, 257–8.

Lachter, J., & Bever, T. G. (1988). The relation between linguistic structure and theories of learning – A constructive critique of some connectionist learning models. *Cognition*.

Lakoff, G. (1971a). On generative semantics. In D. Steinberg & L. Jakobovits (Eds.), *Semantics*. Cambridge: Cambridge University Press, pp. 232–96.

Lakoff, G. (1971b). Presupposition and relative well-formedness. In D. Steinberg & L. Jakobovits (Eds.), *Semantics*. Cambridge: Cambridge University Press, pp. 329–40.

Lakoff, G. (1972a). The arbitrary basis of transformational grammar. *Language, 48*, 76–85.

Lakoff, G. (1972b). Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society*, 182–228.

Lakoff, G. (1973). Fuzzy grammar and the performance/competence terminology game. *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*, 271–91.

Lakoff, G., & Ross, J. (1976). Is deep structure necessary? In J. McCawley (Ed.), *Syntax and Semantics*, Vol. 7. New York: Academic Press, pp. 159–64.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral Mechanisms in Behavior: The Hixon Symposium*. New York: Wiley.

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

Licklider, J. C. R., & Miller, G. A. (1951). The perception of speech. In S. S. Stevens (Ed.), *Handbook of Experimental Psychology*. New York: Wiley, pp. 1040–74.

Mackay, D. G., & Bever, T. G. (1967). In search of ambiguity. *Perception and Psychophysics, 2*, 193–200.

Mandler, G., & Mandler, J. (1964). Serial position effects in sentences. *Journal of Verbal Learning and Verbal Behavior, 3*, 195–202.

Maratsos, M. (1974). Children who get worse at understanding the passive: A replication of Bever. *Journal of Psycholinguistic Research, 3*, 65–74.

McCawley, J. (1968a). The role of semantics in grammar. In E. Bach & R. Harms (Eds.), *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston, pp. 127–70.

McCawley, J. (1968b). Lexical insertion in a transformational grammar without deep structure. *Papers from the Fourth Regional Meeting of the Chicago Linguistic Society*, 71–80. Reprinted in McCawley (1976), pp. 155–66.

McCawley, J. (1976). *Grammar and Meaning*. New York: Academic Press.

McClelland, J. L., & Rumelhart, D. E. (1986). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114*, 159–88.

McMahon, L. (1963). Grammatical analysis as part of understanding a sentence. Ph.D. dissertation, Harvard University.

Mehler, J. (1963). Some effects of grammatical transformations on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior, 2*, 346–51.

Mehler, J., Bever, T. G. & Carey, P. (1967). What we look at when we read. *Perception and Psychophysics, 2*, 213–18.

Mehler, J., & Carey, P. (1967). Role of surface and base structure in the perception of sentences. *Journal of Verbal Learning and Verbal Behavior, 6*, 335–8.

Miller, G. A. (1951a). Speech and language. In S. S. Stevens (Ed.), *Handbook of Experimental Psychology*. New York: Wiley, pp. 769–810.

Miller, G. A. (1951b). *Language and Communication*. New York: McGraw-Hill.

Miller, G. A. (1954). Psycholinguistics. In G. Lindzey (Ed.), *Handbook of Social Psychology*. Cambridge, Mass.: Addison-Wesley, pp. 693–708.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Miller, G. A. (1962a). Decision units in the perception of speech. *IRE Transactions on Information Theory, 2*, IT-8.

Miller. G. A. (1962b). Some psychological studies of grammar. *American Psychologist, 7*, 748–62. (Presidential address to the Eastern Psychological Association, April 1962.)

Miller, G. A. (1963). Comments on Professor Postman's paper. In C. N. Cofer & B. Musgrave (Eds.), *Verbal Behavior and Learning*. New York: McGraw-Hill, pp. 321–29.

Miller, G. A., Bruner, J. S., & Postman, L. (1954). Familiarity of letter sequences and tachistoscopic identification. *Journal of General Psychology, 50*, 129–39.

Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In D. Luce, R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology*. New York: Wiley.

Miller, G. A., Galanter, E., & Pribram, K. (1960). *Plans and the Structure of Behavior*. New York: Holt.

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology, 41*, 329–35.

Miller, G. A., & Isard, S. (1963). Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior, 2*, 217–28.

Miller, G. A., & McKean, K. O. (1964). Chronometric study of some relations between sentences. *The Quarterly Journal of Experimental Psychology, 16*, 297–308.

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of Acoustical Society of America, 27*, 338–52.

Miller, G. A., & Selfridge, J. A. (1950). Verbal context and the recall of meaningful material. *American Journal of Psychology, 63*, 176–85.

Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

Newmeyer, F. (1980). *Linguistic Theory in America*. New York: Academic Press.

Osgood, C. E., & Sebeok, T. A., (Eds.). (1954). *Psycholinguistics: A Survey of Theory and Research Problems*. Supplement to the *International Journal of American Linguistics, 20* (4). Reprinted by Indiana University Press, 1969.

Osgood, C., Suci, G., & Tannenbaum, P. (1957). *The Measurement of Meaning*. Urbana: University of Illinois Press.

Osheron, D. N., Stob, M., & Weinstein, S. (1982). Learning strategies. *Information and Control, 53*, 32–51.

Pinker, S. (1984). *Language Learnability and Language Development*. Cambridge, Mass.: Harvard University Press.

Postal, P. (1967). Linguistic anarchy notes. In. J. McCawley (Ed.), *Syntax and Semantics*, Vol. 7. New York: Academic Press, pp. 201–26.

Pylyshyn, A. (1984). *Computation and Cognition*. Boston: MIT Press.

Rumelhart, D. E., & McClelland J. L. (1986). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, and the PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations of the Microstructure of Cognition*. Cambridge, Mass.: MIT Press.

Savin, H. & Perchonock, E. (1965). Grammatical structure and the immediate recall of English sentences. *Journal of Verbal Learning and Verbal Behavior, 4*, 348–59.

Schank, R. C. (1973). Identification of conceptualizations underlying natural language. In R. C. Schank & K. M. Colby (Eds.), *Computer Models of Thought and Language*. San Francisco: Freeman, pp. 187–247.

Schank, R. C. (1976). The role of memory in language processing. In C. N. Cofer (Ed.), *The Structure of Human Memory*. San Francisco: Freeman, pp. 162–89.

Seidenberg, M. S., & Tanenhaus, M. K. (1986). Modularity and lexical access. In I. Gopnick & M. Gopnick (Eds.), *From Models to Modules: Studies in Cognitive Science from the McGill Workshops*. Norwood, N.J.: Ablex, pp. 117–34.

Seitz, M. (1972). AER and the perception of speech. Ph.D. dissertation, University of Washington.

Sells, P. (1985). *Lectures on Contemporary Syntactic Theories*. Palo Alto, Calif.: CSLI Publications.

Simpson, G. B. (1981). Meaning, dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior, 20*, 120–36.

Skinner, B. F. (1957). *Verbal Behavior*. New York: Appleton-Century-Crofts.

Slobin, D. I. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior, 5*, 219–27.

Slobin, D., & Bever, T. G. (1980). Children use canonical sentence schemas in sentence perception. *Cognition, 12*, 229–65.

Stewart, C., & Gough, P. (1967). Constituent search in immediate memory for sentences. *Proceedings of the Midwestern Psychology Association.*

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

Tanenhaus, M. K. (in press). Psycholinguistics: an overview. In F. Newmeyer (Eds.), *Linguistics: The Cambridge Survey.* Volume 3, Biological and Psychological Perspectives. Cambridge, Mass.: MIT Press.

Tanenhaus, M. K. & Carroll, J. M. (1975). Functional clause hierarchy . . . and Nouniness. In R. Grossman, J. San, & T. Vance (Eds.), *Papers from the Parasessian on Functionalism.* Chicago: Chicago Linguistics Society.

Tanenhaus, M. K., Carlson, G. N., & Seidenberg, M. S. (1985). Do listeners compute linguistic representations? In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural Language Parsing*, New York: Cambridge University Press, pp. 359–408.

Tanenhaus, M. K., Carroll, J. M., & Bever, T. G. (1976). Sentence-picture verification models as theories of sentence comprehension: A critique of Carpenter and Just. *Psychological Review, 83*, 310–17.

Underwood, B., & Schultz, R. (1960). *Meaningfulness and Verbal Learning.* Philadelphia: Lippincott.

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

Wanner, E. (1968). On remembering, forgetting, and understanding sentences: A study of the deep structure hypothesis. Ph.D. dissertation, Harvard University.

Wexler, K., & Culicover, P. (1980). *Formal Principles in Language Acquisition.* Cambridge, Mass.: MIT Press.

Wingfield, A., & Klein, J. F. (1971). Syntactic structure and acoustic pattern in speech perception. *Perception and Psychophysics, 9*, 23–5.

Wundt, W. (1874). *Grundzuge der psysiologischen Psychologie.* Leipzig: Engelmann.

Wundt, W. (1911). *Volkerpsychologie & Erster Band: Die Sprache.* Leipzig: Englemann.