This excerpt from

Sentence Comprehension.
David J. Townsend and Thomas G. Bever.
© 2001 The MIT Press.

is provided in screen-viewable form for personal use only by members of MIT CogNet.

Unauthorized use or dissemination of this information is expressly forbidden.

If you have any questions about this material, please contact cognetadmin@cognet.mit.edu.

Chapter 2
Classical Evidence for the Sentence

This chapter and the next present the historical background for contemporary models of sentence comprehension. The long history of scientific approaches to the sentence also serves as a cautionary historical tale about the scientific study of what we know, what we do, and how we acquire what we know and do. As we discussed in the first chapter, two opposing paradigms in psychological science deal with these interrelated topics. The behaviorist *prescribes* possible adult structures in terms of a theory of what can be learned from explicit data. The rationalist *explores* adult structures, including those that are implicit, to find out what a developmental theory must explain. In this chapter, we outline a century of alternations between approaching the sentence as defined by linguistic knowledge, and treating it as an associatively processed behavioral concept. This history leaves a residue of consistently reappearing associative and structural facts about language that any comprehension theory must account for. In addition, it lays out some options on how to integrate structural theories of linguistics knowledge with associative and statistical properties of language behavior.

2.1 Early Ideas about the Sentence

Experimental cognitive psychology was born the first time at the end of the nineteenth century. The pervasive paradigmatic 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, summarized a century of research on the natural units and structure of language (see especially Humboldt 1835/1903; Chomsky 1966). Wundt came to a striking conclusion: The natural unit of linguistic knowledge is the *intuition* that a sequence is a sentence. He reasoned as follows:

• We cannot define sentences as sequences of words because there are single-word sentences (e.g., “Stay”).

• 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., "Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday").
• 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 the description of linguistic knowledge: The goal of linguistic description is 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. The structural features included 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*, despite being adjacent to both. Similarly, in sentence (2), *was* is intuitively closer to *crossed* than to *Rubicon*.

(1) Caesar was crossing the Rubicon.

(2) The Rubicon was crossed by Caesar.

(3) Cross the Rubicon was what Caesar did.

The surface level also defines a set of surface grammatical relations between the phrases. In sentence (1), *Caesar* is the grammatical subject—that is, 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*.

It was obvious that surface grammatical relations could not capture the propositional relations between the phrases. Wundt noted that *Caesar* is the acting one in each of sentences (1) to (3) despite its different positions. The propositional relations between phrases are represented by a separate level that Wundt called the *inner form* of the sentence. At this level, sentences (1) to (3) share the same relations between *Caesar* (agent), *cross* (action), and *Rubicon* (patient). The different actual sequences at the surface grammatical level are related to the propositional level by mapping processes called *Umwandlungen* (literally, "transformations"). These processes reorder surface phrases into the surface patterns allowed by the particular language. The relations between elements of a proposition are not purely semantic, but are the formal expression of relations between semantic units of meaning (Blumenthal 1970, 1975). That is, even the propositional form is arranged according to a system.

2.2 Banishment of the Sentence

The continental model of language was rich and made many claims about the capacity of humans to manipulate abstract entities. But the theory never became an object of experimental importance. The reasons are, no doubt, scientifically, even politically 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 laboratory experimental investigation. His vast structural catalog of language is more an anthropological survey than a scientific scrutiny of mental processes. The continental linguistic model and its richness became lost to scientific psychology.

But Wundt's model was not lost to everybody 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. The notions "intuition" and "inner grammatical level" were not acceptable in a framework that required operational definitions.

Even Bloomfield capitulated to such restrictions as enthusiastically as he had earlier espoused the Wundtian model. His foundational book, *Language* (1933), presents a behaviorist framework for linguistic theory. In that book, the sentence is hardly mentioned, while meaning is given cursory treatment in terms of the reinforced association of stimuli and responses.

2.2.1 Behaviorism, Stimulus-Response Theory, and the Sentence

Behaviorism is a seductive doctrine that dominated psychological theories of learning for most of the twentieth 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 we learn is that the environment provides pleasure when we do. That is, the environment reinforces only certain activities in certain circumstances, and those activities become habits. This 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 an appealing chain of inference from the motive to learn back to the structure of what is learned.

The classic behaviorist model in behavioral science is the S-R schema, laid out by Watson (1919) and given more formal definitions and philosophical justification by Skinner (1957). This model of behavior describes every behavior as a response to a particular stimulus, ranging from relatively automatic behaviors such as ducking at a loud noise to obviously learned behaviors, such as stopping at a red light. In each case, a period of training inculcates the habit of producing a particular response to a particular stimulus. The training consists of presentation of the stimulus *S* and then some form of positive reinforcement if the correct response *R* is produced: every time a S-R sequence is followed by a positive reinforcement, the S-R bond is strengthened, (4a), just as it is weakened when followed by punishment, (4b). That is, behaviors are "associated" with stimulus configurations, by virtue of independent "reinforcement." This simple architecture has tremendous power implicit in it, and the principles dominated behavioral science for roughly four decades.

(4) a. Red light → stop, Positive Reinforcement (driving instructor says "good")
    b. Red light → go, Punishment ("crash!")

Part of the philosophical surround of associationism is behaviorism, the principle that only observable stimuli, responses, reinforcements, and punishments can count as part of a theory. This restriction seemed viable so long as fairly simple behaviors were at issue, but a complex behavior such as language seemed to resist such treatment —where are the stimuli, the responses, the reinforcements? Ultimately, various S-R theorists proposed that language behavior could be accounted for with a model in which each word served as a stimulus for the next, building up an overall structure out of local associative relations (e.g., Staats 1961; Kendler and Kendler 1962). For example, in producing the sentence in (5a), *the* can be taken as the stimulus that elicits *horse*, which in turn elicits *races*, as in (5b).

(5) a. The horse races.
    b. The → horse → races.

The S-R treatments of language also formulated a distinction between function words (e.g., *the*) and content words, based on the different kinds of references they have (content words ostensibly have externally definable reference; function words do not). This allowed differential reinforcement of two kinds of sequence information, function-to-function and content-to-content. For example, sentence (5a) is composed of two overlapping sequences:

(5) c. The → X_ → Y es
    d. horse → races

From a structural standpoint, (5c) captures part of the structural regularities relating to noun-verb agreement, and contrasts with a different sequence (5f) when the noun is plural, as in (5e). The sequential probability in (5d) captures the meaning relation, in which it is frequently true of *horses* that they are the agent of *race*.

(5) e. The horses race.
f. The → X es → Y_

While initially couched as a model of how language behavior could be learned and maintained, this scheme can also be interpreted as a model of language comprehension. In that model, the learned sequences of adjacent elements are internally represented as automatically characterizing a sentence as it is encountered.

Chomsky and others noted fifty years ago that such a classic S-R sequencing of elements is not adequate to describe linguistic facts and that similar limitations apply to unadorned S-R theory for the interpretation of sentences. The most salient reason is that sentences manifestly have elaborate hierarchical structure and long distance dependencies in which the two parts of associated components are separated by an arbitrary distance. Any model limited to expressing the associative relation between elements no more than a limited distance apart will not be able necessarily to represent hierarchical relations, and will surely be inadequate to represent long distance dependencies.

(6) a. The hors*es* that were raced past the barn *are* falling.
b. The horse_ that was raced past the barn *is* falling.

Chomsky generalized his critique of Skinner's proposals on language to the entire S-R program. At a general level, he noted that it is impossible to define independently what *counts* as a stimulus, a response, or a reinforcement in any normal complex situation. There is no independent definition of stimulus, or of response or reinforcement for that matter—it is all determined after the fact. If a motorist stops at an intersection, if we did not know already about stoplights, how would we know what had actually controlled the behavior, how would we know what the effective behavior is, and how would we know where the reinforcement is? Rather, we think we know each of the three components because after the fact we can analyze them. Similarly, without already knowing about the relation between noun-number and verb inflection, how would we (or the listener) know what aspect of the speaker's behavior controlled the relation; indeed, how would we know there is a relation? Furthermore, in language, how do we know exactly which pieces of the preceding string should be taken as the stimulus for the current word—that is, what exactly should be reinforced? We will see that this problem corresponds in modeling language comprehension with associative models to the "grain" question, namely, how does a system select the "relevant" amount and level of information to form complex associations subject to reinforcement?

2.2.2 Mediation Theory and Linguistic Knowledge

The first proposals within the S-R framework that attempted to meet Chomksy's challenge also antedate the classic psycholinguistic period and are primarily due to Charles Osgood (1963; Osgood, Suci, and Tannenbaum 1957). Osgood was a S-R psychologist trained by Clark Hull, who elaborated the role of internal entities that "mediate" external stimuli with external responses. Hull and his students had addressed the problem of how to adapt simple S-R theory to the learning of complex chains—for example, when a rat learns a maze with two turns before reaching the reinforcing goal (as in (7a)). The last correct turn (R_2) is reinforced by the presence of the reward in the goal location, but how does the final reinforcement affect the first correct turn (R_1)? Hull's proposal (1943) was that both "responses" and "reinforcement" could be analyzed as "fractionating" into parts that can be related to each other at points prior to the final response and reinforcement. For instance, the first turn has two components, the actual immediate response (e.g., R_1, "turn right") and a fraction of the goal response (e.g., "Go toward the goal")—that is, the initial correct turn is both a local behavior and part of the final goal-directed behavior. Hull represented this by postulating a fractional response (r_{g_1}) and fractional reinforcement (s_g) at early points in a chain, as in (7b):

(7) a. $S_1 \rightarrow R_1$, $S_2 \rightarrow R_2$, reinforcement by Goal
b. $S_1 \rightarrow R_1$
r_{g_1}, reinforced by s_g
$S_2 \rightarrow R_2$, reinforced by Goal

Of course, this leads to the postulation of a kind of variable that can intervene between the stimulus and response, a hypothesis anathema to behaviorists. There were many proposals and much worry about how such intervening variables could come to exist, based on the principle that only observable stimuli, responses, and reinforcements can play a causal role in behavior.

Even if it is not possible for behaviorist principles to account for the theoretical notions of fractionated responses and reinforcements, those concepts can be applied to the description of a complex behavior like language, and at least potentially maintain the usefulness of associationism. Osgood first explored the application of this kind of schema to the representation of word meaning. According to his proposal, a word stimulates a set of particular intervening response components with varying strength (r_{g_1}, r_{g_2}, ...), which in turn are connected to a set of intervening stimulus components (s_{g_1}, s_{g_2}, ...), which connect to overt responses.

For subsequent developments it is instructive to analyze some critical features of this model. First, the input is an explicit word. Each word is connected to an intervening set of meaning "r-s" *components* or *features* in which the *r* is a fractional, or internal, part of the goal response and serves as an internal stimulus for other behaviors, which may also be internal. Each word has an activation strength level associated with each feature, ranging from 0 to a ceiling level. The features themselves are characteristically learned by way of massed exposure to a wide range of word-word pairs. The claim was that out of many such pairs, similarities in meanings would coalesce around activations of particular intervening r-s components, thereby isolating them as relevant to differentiating meanings from each other. Random intervening components would tend to fall out of general use because they would not be relevant to extracting unique responses from any input words. (At times, Osgood also entertained the view that at least some of the intervening components could be innately constrained, while the connection strength to them would be based on experience.) One might label the separate intervening r-s to the word *elephant* as in (8).

(8) a. r_g-s_1 = "animate"
 b. r_g-s_2 = "large"

Finally, any given word could be the critical factor in a variety of actual output responses, because the intervening components could be stimulated by other mental factors, adjacent words, background level, local context, and so on. Thus, while *elephant* as an input always resulted in raising the activation level of each intervening component by a fixed amount, the resulting output could vary because of other factors that govern which response receives the overall highest activation: sometimes *elephant* can elicit *cat* or *house*, but rarely *force*.

Thus, unlike Skinner, Osgood recognized explicitly that in a S-R explanation of a complex phenomenon like language, a fundamental problem was how to reinforce and therefore represent "abstract" intervening r-s connections that mediate between explicit stimuli and responses. Osgood built on the Hullian concept of fractional responses and reinforcement to explain language behavior. In Osgood's scheme for representing and learning word meaning, an overt stimulus (e.g., an appearance of a word) automatically stimulates the universal set of intervening "semantic" r-s modules (what we might think of as intervening "features"). The overt response results from a particular set of those r-s modules. Reinforcement of that response spreads "backward" to reinforce the particular intervening r-s modules that played a positive role in eliciting the overt response. Osgood (1963) extended these principles to explain the role of syntactic structure in language behaviors.

In this sense, Osgood was an enlightened associationist who postulated intervening entities that seem quite abstract. But he was also a "behaviorist," hewing to the position that all causally relevant intervening variables in behavior must ultimately be resolvable to components of actual stimuli or responses. He shared with most psychologists of his day the view that mental life is made up of the componentially reinforced shards of overt stimuli and overt responses. In this way, he wedded himself to the fundamental weakness of the Skinnerian program, despite his enlightened views on linguistic representations. Osgood's behaviorism made him continually vulnerable to the same kind of attacks as did Skinner's.

For example, J. A. Fodor (1965) articulated a number of difficulties with this model of meaning. Many of these were technical matters involving the mechanics of how any associative model could work. Most general was the observation, like Chomsky's on Skinner, that the model does not offer independent behaviorist definitions of its independent terms. In particular, there is no principled way to determine which aspects of an input word are just the ones that should be reinforced with a strengthening of connections to intervening components. The same question arises with respect to the connections from those components to the output responses.

Despite its shortcomings, Osgood's attempt to deal with meaning in an associative framework offers a background and general description of features of such models, including some that are contemporary (see chapter 4). A specific input (a word) can activate a set of internal components, which along with activations contributed by other sources, can result in a range of outputs in which the specific input is a critical, but not the sole, factor. In light of current connectionist models that we discuss below and in chapter 4, Osgood was a pioneer—he saw how a distributed representation model could be integrated with fundamental principles of associationism to explain complex abstract behavior.

Reinforcement was an important concept because it justified the scientific investigation of isolated sources of pleasure and displeasure. The focus on behavior-situation pairs licensed investigation of the learning of meaningless associations between situations and behaviors. The requirement that learned associations occur between definable entities transferred to the operationalist requirement that we reduce theoretical terms to observable entities.

By the late 1950s, sophisticated elaborations of these principles had crystallized, most notably in the proposal by Hull (1943) that these principles could account for chains of behavior. Even when transferred to descriptions of language behavior, these principles retained the basic behaviorist doctrine about the structure of what was learned (Osgood, Suci, and Tannenbaum 1957): There must be recognizable links between isolatable situations and behaviors (Fodor 1966; Bever 1968a). The implications of this doctrine for theories of language were severe. Consider how these procedures affected theories of the relationship between 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 (9a) to (9d) could all be resolved to a basic sequence of the same kinds of phrases—a noun phrase, a verb, and an adjective phrase.

(9) a. Harry was eager.
 b. The boy was eager.
 c. The tall boy was eager to leave.
 d. 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 (9d). In this way, they rendered "phrases" 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 was 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 into single words.

In the 1930s and 1940s, psychological research on language proceeded largely without explicit benefit of linguistic analysis. Psychologists studied the processing of unrelated word sequences with the goal of discovering how universal laws of learning apply to human beings. This deceptively simple paradigm, called *verbal learning*, became the focus of an enormous amount of research. There were many hundreds of such studies, with several journals devoted to little else (for representative reviews see Underwood and Schulz 1960; Cofer 1961; Cofer and Musgrave 1963). Great effort was devoted to exploring the formation of associations between adjacent and remote words in strings and 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. The focus was on learning, motivation, and memory, not language. Of course, one could have viewed this as leading to an understanding of how words are processed when they are organized in whole sentences. Unfortunately, this promise was never realized. As we will see shortly, words in whole sentences are processed differently from words in random sequences.

Just as wholes are the natural enemies of parts, Gestalt psychology was 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. It would appear that linguistic structures are prime candidates for gestalt investigations. After all, sentences are wholes that bind together and transcend words and phrases as their parts (Lashley 1951; Mandler and Mandler

1964). Such an obviously true observation, however, rarely stops associationists from going about their business, and, in this case, it 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 that "good figures" undercut associationistic accounts of language was due in part to the inability of Gestalt psychology to develop a general theory of possible good figures. 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 even though they shared the fundamental theoretical restrictions of behaviorism and some beliefs about the basic role of S-R associationism. An early burst of psycholinguistics occurred when the two groups discovered that they could mate their theories. Learning theory with mediating responses of the sort developed by Osgood 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 and Sebeok 1954). Although the first mating between psychology and linguistics was briefly intense, it was sterile. The reason was that the two disciplines were mutually compatible just because they shared the same behaviorist preconceptions. Psychologists were willing to postulate of the language learner only the inductive capacity to learn what linguists had already restricted themselves to describing. 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.

2.3 Revival of the Sentence

We now turn to the beginning of modern times, starting about fifty years ago. Syntactic investigations increasingly focused on sentence-level analyses. Correspondingly, experimental investigations of real-language behavior lead psychologists to the empirically demonstrated explanatory power of the sentence level.

2.3.1 The Sentence in Behaviorist Linguistics

The behaviorist implementation of linguistics may seem harmless enough, but it had a particular result: linguistic theory could not describe the sentence. This is true for three empirical reasons. First, the number of sentences in unfathomably large. Second, in a single sentence, there are often relations between words in different phrases.

Third, there are grammatical relations between phrases that cannot be described as superficial facts about the phrases themselves. In addition, as Wundt had noted, it is impossible to define sentences without appealing to speakers' intuitions. 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 observable 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. Harris (1957, 1958) invoked the operationalist view that the sentence is a component of a complete discourse in the same sense that a phrase is a component of a sentence. 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) to (3) are part of a larger set of constructions, as follows:

(10) a. Caesar crossed the Rubicon.
b. It is the Rubicon that Caesar crossed.
c. What Caesar did was cross the Rubicon.
d. The Rubicon is what Caesar crossed.
e. that Caesar crossed the Rubicon ...
f. Caesar's crossing the Rubicon ...
g. the Rubicon's being crossed by Caesar ...
h. the crossing of the Rubicon by Caesar ...

Harris noted that each of these variants can occur in the same discourse environment. That is, we can substitute each of them for the blank in the following discourse frame (ignoring changes needed to accommodate the clausal variants):

(11) a. Caesar marched north.
b. Then ______.
c. This surprised the local inhabitants.

Harris intended the notion of co-occurrence to be the same as that describing the substitutability of phrases in discourse-based sentence frames. The difference is that we cannot reduce sentences—unlike phrases—to canonical words (*it*, *did*). Rather, Harris suggested that we can reduce the sentences of a structural-sentence family to a standard canonical sentence form that he called the *kernel*. A kernel sentence is the simple declarative construction. Co-occurrence "transformations" express the relation between the kernel sentence and its variants. For example, the kernel and the passive sentence (2) are related by the following co-occurrence transformation:

(12) "NP_1 $V + ed$ $NP_2 \leftrightarrow NP_2$ 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. 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 we return after reviewing progress in the psychology of the day.

2.3.2 The Unit of Perception

During the same period (the 1950s) there was a separate stream of research on how adults organize language behavior 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. In particular, Miller and Selfridge (1950) demonstrated the behavioral relevance of sentence structure by showing that memory for word sequences improves as they approach the statistical regularities of English. This result suggested that language behavior involves a transfer of the physical signal into a linguistic model 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.

A standard experimental method to investigate the nature of the unit of perception was based on the ordinary fact that 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 your language to aid your perception. But which aspect of your linguistic knowledge do you access first and how does it help you hear better? 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 their sequence.

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, and Lichten 1951; Miller and Isard 1963). Suppose we adjust the loudness of a noise source relative to recordings of the words in a two-word sentence (13) so that each word is recognized correctly 50 percent of the time when it is heard alone.

(13) a. Horses eat.
b. Horses cry.

If it is the word level at which the acoustic information is mapped onto linguistic representations, 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, however, are striking. When strung together into a sentence, the word sequence is often recognized much more than 50 percent of the time (Miller, Heise, and Lichten 1951). Most important is the intuitive fact that when the words are in a sentence they seem acoustically clearer *as you hear them*. The outcome of a series of such studies was the conclusion that it is at least the sentence that is the unit of speech perception. Even well-formed sentences that do not make semantic sense (e.g., 13b) enhance perception of the words in them.

Miller's conclusion on the role of sentences in speech perception created two questions 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 question 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 ideal was consistent with the behaviorist strictures on what we can learn: if beliefs can influence perception, 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; see Miller, Bruner, and Postman 1954) suggested that perceptual activity involves the *simultaneous* integration of abstract and concrete levels of representation.

There are two views about 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"). Research on language perception has vigorously pursued the nature of the integration of different levels of knowledge because there are well-defined, distinct levels of structure in language. The pendulum of scientific fashion has swung back and forth, with some investigators in favor of a bottom-up theory of perception (Forster 1979; Fodor

1983; Pylyshyn 1984, to name a few) and others in favor of a top-down theory (e.g., Marslen-Wilson and Welsh 1978). While the issue may never be entirely resolved, Miller's clear experimental demonstration of the behavioral importance of the 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.

The acoustic clarity of sentences raised another question. 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 using prememorized sentences. 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.

The 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. Once some thought was given to the question, it seemed clear that people actively use categorical rules in many sorts of behavior (Miller, Galanter, and 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. Such a theory emerged at the time that Miller's research on speech perception raised the question of the nature of knowledge about language.

2.3.3 Transformational Grammar

What *is* a sentence? Harris had attempted to include the sentence level within an operationalist framework, but discovery procedures limited his success. In the late 1950s Chomsky (1957) 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 respects to those of Wundt (Chomsky 1957).

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 the sentence structures of a language as a natural and creative 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 approach represented 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). For Chomsky the goal of linguistic anlaysis was to describe an adult's linguistic knowledge, not language behaivor. 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 or she 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 relatively 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 is paired with an underlying kernel or "deep" structure, which presents the propositional relations for each verb in a canonical form. A set of transformations specifies the possible pairs of deep and surface phrase structures. Unlike Harris's co-occurrence transformations, these transforamtions operate in one direction only. Each transformation applies a successive deformation of the input tree, for example, changing a structure that would end up as an active into one that will end up as a passive:

(14) tree 1 $\xrightarrow{\hspace{6em}}$ tree 2

Certain transformations combined kernel trees into complex sentences. The way different kernel sentences are combined reveals interesting differences in the deep structure organization of sentences that are superficially similar. For example, sentences with *desire* and *defy* have superficially identical structures:

(15) a. John desired Bill to go.
b. John defied Bill to go.

But (15a) and (15b) represent the combination of very different kernel sentences. This is reflected in the fact that the entire phrase *Bill to go* is the intuitive object of *desired*, but only *Bill* is the intuitive object of *defied*. We can illustrate these kernel sentences with the following, where $\diamond$ represents a position into which a complement sentence is to be inserted:

(16) a. John past desire $\diamond$ Bill to go
$\rightarrow$ John desired Bill to go
b. John past defy Bill $\diamond$ Bill to go
$\rightarrow$ John defied Bill to go

Also attesting to the underlying distinctness of (15a) and (15b) are their different structural variations. For example, with *desire* but not with *defy*, we can use the whole complement sentence as a unit in a passive sentence. Contrast your intuitions about the acceptability of the following two sequences:

(17) a. For Bill to go was desired by John.
b. *For Bill to go was defied by John.

(We follow standard practice of indicating unacceptable sentences with an asterisk.) 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:

(18) a. John was eager to please.
b. John was easy to please.

(19) a. John past BE eager ◇ John please someone
→ John was eager to please
b. It past BE easy ◇ Someone please John
→ John was easy to please

In (18a) *John* is the underlying subject of both *eager* and *please*, but in (18b) *John* has only the underlying object position of *please*.

In Chomsky's (1957) syntactic theory, 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 the 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 and Postal 1964; Chomsky 1965). The sound was produced by operations on the surface representation.

2.3.4 A Direct Link Between Knowledge and Behavior

Linguists made a firm point of insisting that, at most, a grammar was a model of "competence"—that is, what the speaker knows. This was contrasted with effects of *performance*, actual systems of language behaviors such as speaking and understanding. Part of the motive for this distinction was the observation that sentences can be intuitively "grammatical" while being difficult to understand, and conversely. For example, Chomsky and Miller (1963) noted that center-embedded sentences become very difficult once there is an embedding within an embedding. This is not a conceptual problem, as shown by the transparent comprehensibility of the same propositions with no embedded structures; see (20c).

(20) a. The house the boy built fell.
b. The house the boy the cat scratched built fell.
c. The cat scratched the boy who built the house that fell.

Furthermore, some center-embedded sentence are sometimes comprehensible (Bever 1970a).

(21) The reporter everyone I met trusts had predicted the coup.

Thus, it seemed clear that there is a distinction between what we "know" about sentences and what we can "do" with them. The grammar offered a precise answer to the question of what we know when we know the sentences in our language. We know the different coherent levels of representation and the linguistic *rules* that interrelate those levels. Linguistics distinguished competence from "performance"—how the speaker implements this knowledge. Despite this distinction the syntactic model had great appeal as a model of the processes we carry out when we talk and listen. 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?

Chomsky's (1957) syntactic model provided an answer to the first question. 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. When the model first appeared, the initial answer to the second question was couched in terms of the linguistic model and a direct assumption about how to link the model to behavioral systems. It was assumed that 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. This hypothesis was explicit in the so-called derivational theory of complexity, that the behavioral complexity of processing a sentence corresponds to the number of transformations in its found description. Of course, that hypothesis was only one of many that might be postulated about the relation between grammar and processing mechanisms, as Chomsky and Miller noted.

The distinction between competence and performance created an aura in which it could appear that "competence' was some kind of abstraction that had no particular implication for mental structures: it was performance that carried out that mapping and created behavioral expressions of the grammar. Technically (and importantly), this common view was a mistake: ‘‘competence’’ itself was a theory of actual knowledge, and therefore it pertained directly to mental structures. Furthermore, the most essential behavior in determining competence was the linguistic intuition of native speakers about well-formed sentences (just as it was for Wundt). Grammaticality intuitions are real behaviors in their own right, and a grammar that accurately distinguishes between grammatical and ungrammatical sentences already explains a vast range of behavior. Nonetheless, many believed that it was important to demonstrate a direct link between knowledge and behavior and thereby to demonstrate the so-called psychological reality of linguistic structures in ongoing behavior.

It was not controversial to claim that surface phrase structures were ‘‘psychologically real’’ because they were allegedly definable within the behaviorist framework. (In fact, they probably are not. See Bever 1968a; Bever, Fodor, and Garrett 1968.) But the claim that *deep* structures were psychologically real was a direct challenge to behaviorism. Deep structures were abstract in the sense that they are not actual sentences, but rather the ‘‘inner form’’ of sentences like that 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, as in changing ‘‘he past leave’’ to ‘‘he left.’’ The further claim that transformations are 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 in testing the hypothesis of a direct link between linguistic knowledge and behavior 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; see figure 2.1). Sentences can be either active (i.e., kernel, K) or passive (P), declarative or interrogative (Q), positive or negative (N). In the linguistic theory of the day, each of these dimensions corresponded to transformation:

(22) a. Mary hit Mark. K
b. Mary did not hit Mark. N
c. Mark was hit by Mary. P
d. Did Mary hit Mark? Q
e. Mark was not hit by Mary. NP
f. Didn’t Mary hit Mark? NQ
g. Was Mark hit by Mary? PQ
h. Wasn’t Mark hit by Mary? PNQ

Accordingly, in testing the behavioral implications of the formal relations between these sentences as arrayed on the sentence cube, as shown in figure 2.1, one could test the psychological relevance of the grammatical model.

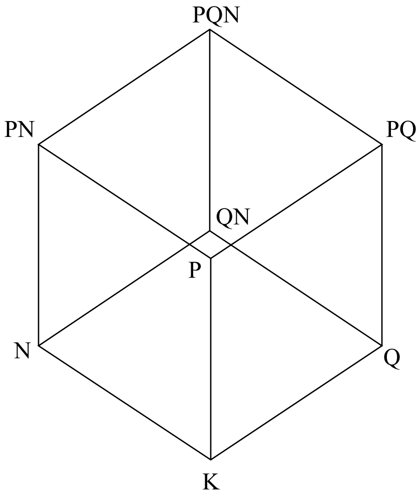

Figure 2.1
The sentence transformation cube (Miller, 1962b, adapted from fig. 12).

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 sentence cube (Miller and McKean 1964). For example, given "Mary hit Mark" it is easier to produce the negative variant than to produce the negative passive version. Transformational distance between sentences also predicts confusability between them in memory (Mehler 1963; Clifton, Kurcz, and Jenkins 1965; Clifton and Odom 1966). For example, a passive-question sentence is more often recalled as a passive than as an active. 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 and 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 and Miller 1963; Miller and Chomsky 1963). The linking hypothesis of a direct mapping from the structure of linguistic knowledge and the processes of language behavior was wildly successful. The golden age had arrived.

2.4 Unlinking Knowledge and Behavior

Alas, 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 sentence cube. But the overall program implied a broad empirical hypothesis about the relation between all rules in a grammar and sentence processing.

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 and Garrett 1967). It was too weak because it failed to predict the obvious complexity of numerous kinds of sentences.

(23) *Center embedding*
   a. The oyster the oyster the oyster split split split.
   b. The reporter everyone I met trusts predicted the coup.

(24) *Reduced-object relative clauses*
   a. The horse raced past the barn fell.
   b. The horse ridden past the barn fell.

In both (23) and (24) the two versions are structurally identical, but in each case the (b) version is easier to understand. (We return to reduced relative structures in chapter 7.)

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

(25) *Heavy noun phrase (NP) shift*
   a. We showed the long-awaited and astoundingly beautiful pictures of the Himalayan trip to Mary.
   b. We showed Mary the long-awaited and astoundingly beautiful pictures of the Himalayan trip.

(26) *Extraposition*
   a. That Bill left early with Mary surprised Hank.
   b. It surprised Hank that Bill left early with Mary.

In both (25) and (26) the (a) version is transformationally less complex than the (b) version but is harder to understand.

Further research backed up these direct demonstrations of the inadequacy of the derivational theory of complexity. These experiments examined the implications for perceptual difficulty of transformations other than those that defined the sentence cube. Several studies of specific transformations failed to show that perceptual difficulty corresponds to the number of transformations (Bever et al. 1966; Bever and Mehler 1967; Fodor and Garrett 1967; Jenkins, Fodor, and Saporta 1965).

Consider, for example, the transformation that optionally moves a particle/preposition to the position following the verb. This transformation was well motivated as starting with the (*Verb* + *particle*) + *NP* deep structure, in which (*Verb* + *particle*) is a complex lexical item, and then being transformed to the *Verb* + *NP* + *particle* sequence:

(27) John called up Mary → John called Mary up
verb prt NP → verb NP prt

The fact that verbs are lexically associated with some particles/prepositions and not others supports this order of structures. For example, we have "call over NP" and "call NP over":

(28) a. John called over the waiter.
b. John called the waiter over.

And, although we have "call under NP," we do not have "call NP under":

(29) a. John called under the bridge.
b. *John called the bridge under.

If *Verb* + *particle* sequences are entered as complex lexical items and *then* transformed, we can capture these lexical facts. Despite this clear motivation for treating the *Verb* + *particle* + *NP* variant as less complex transformationally, such sentences turned out to be processed more slowly than the corresponding transformed versions (Bever and Mehler 1967).

Such failures were baffling in light of the initial success of the structure arranged on the sentence cube. The failures motivated reconsideration of the theoretical interpretation of the three-dimensional model. This further consideration revealed that if we take the grammatical theory literally, it would not have motivated many of the original predictions. That is, 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 does not involve undoing the negative and then the passive. Rather, it involves 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 sentences adjacent on the sentence cube (figure 2.1)—for example, between the negative passive and the passive question—would be far less likely than between two simpler structures, such as 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.

In light of the lack of motivation due to a strict implementation of the grammar, the question remained of why the experimental results on the sentence cube were obtained. One possibility was that the sentence 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 with constant superficial relations between them 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 sentence 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 and 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 and Postal 1964; Chomsky 1965).

Another explanation for why the transformational distance from the kernel structure on the sentence cube predicted the relative difficulties of comprehension and perception appealed to meaning. For example, the question and negative transformations actually changed the meaning of the sentence. It is therefore not surprising that such sentences might be relatively hard to process precisely because of their semantic complexity. That is, the negative has the same propositional structure as the declarative, with an additional semantic operator. The finding that passive constructions are relatively complex *only* when they are semantically reversible supported the importance of semantic analysis (Slobin 1966; Walker, Gough, and Wall 1968). For instance,

(30) The chief of the Gauls was overcome by Caesar.

is more difficult than its corresponding active construction, but

(31) The river Rubicon was overcome by Caesar.

is not. The conclusion of such demonstrations and further theoretical analysis was that there is no direct evidence for the step-by-step application of grammatical transformations in speech behaviors.

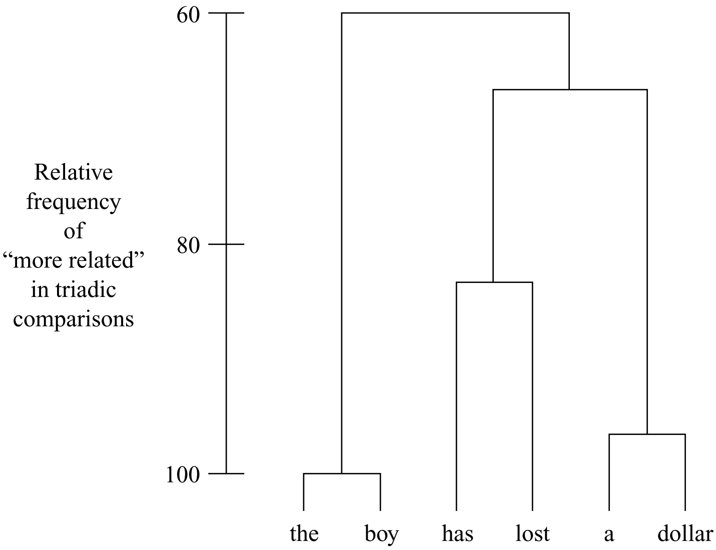

Figure 2.2
A hierarchical clustering analysis of a sentence (adapted from Levelt 1970).

The behavioral relevance of grammatical levels of representation fared better. We can construe surface phrase structure 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. Levelt (1970) directly tested this idea. He demonstrated that a hierarchical clustering analysis can produce a surface phrase structure. Levelt asked subjects to judge the relatedness among the word triads of a sentence, and then used a hierarchical clustering procedure to analyze the judgments. A sample of Levelt’s results, shown in figure 2.2, corresponds closely to linguists’ analysis of phrase structure. Accordingly, words are better prompts for other words in the same surface phrase than in different phrases (Johnson 1965; Stewart and Gough 1967; Suci, Ammon, and Gamlin 1967; Townsend and Saltz 1972). For example, once subjects have read

(32) The tall boy saved the dying woman.

tall is more successful in producing recall for *boy* than for *saved*.

Surface structure is behaviorally relevant for perception as well as for memory. Several studies demonstrated that we impose surface phrase structure on sentences as we perceive them (Mehler, Bever, and Carey 1970; Mehler and Carey 1967). Studies on the perceived location of brief nonspeech sounds that interrupt a sentence dramatically revealed the importance of the surface structure pattern. Subjects characteristically report such sounds as having occurred in a phrase break, especially between clauses, rather than in their actual location (Fodor and Bever 1965; Garrett 1965). For example, listeners judge such nonspeech sounds objectively located at the O in (33a) and (33b) to have occurred at the position of the R:

(33) a. In her *hope of marrying Anna was impractical.*
O R
b. Harry's *hope of marrying Anna was impractical.*
RO

That is, the syntactically driven clause boundary determines judgments of the location of the nonspeech sound. Carefully designed materials showed that this effect is not dependent on variations in intonation. For example, the italicized sequence in sentences (33a) and (33b) is common to both syntactic organizations (Garrett, Bever, and Fodor 1966). Nor is the effect due to guessing biases (Garrett 1965; Holmes and Forster 1970; Bever and Hurtig 1975) or to variation in word sequence probability (Bever, Lackner, and Stolz 1969). As Garrett, Bever, and Fodor (1966) 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 Miller's studies of the comprehension of speech-in-noise: listeners actively deploy grammatical knowledge in a way that clarifies the speech signal. In the next section, we review evidence on the role of deep structure in language behavior.

2.4.1 The Coding Hypothesis and Deep Structure

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, we cannot directly observe deep structure in the surface sequence. On the other hand, we cannot define deep structure in semantic terms. Rather, deep structure is a formal structure that mediates between surface form and meaning. Several psychologists proposed that deep structure was, in fact, the basic schema in which we organize and retain sentences, with transformations specifying modifications of the basic schema (Mehler 1963; Miller 1962b). The *coding hypothesis* was formulated in terms of a "schema plus correction" model for memory, proposed as a general principle by Bartlett (1932). According to the coding hypothesis, we retain sentences in the simple declarative form, plus transformationally motivated "tags" such as passive, negative, and so on that express the actual construction of the sentence (Mehler 1963; Miller 1962a). The coding hypothesis also is a literal interpretation of the sentence cube in which the simple declarative sentence is a kernel structure.

There was further support for the interpretation that the coding hypothesis applied to Chomsky's (1957) abstract underlying structures. All other things being equal, the perceived relatedness between words is greater when they are in the same deep structure phrase (Levelt 1970). For example, in (18a) and (18b), repeated here as (34)

(34) a. John was eager to please.
b. John was easy to please.

John and *eager* are perceived to be more closely related than *John* and *easy*, since, as we noted earlier, in (34a) *John* appears in the same deep structure clause as *eager*, but in (34b) *John* does not appear in the same deep structure clause as *easy*. Words also are better prompts for other words in the same deep structure phrase than in other phrases (Walker, Gough, and Wall 1968; Davidson 1969). In fact, a word is a better prompt for recall of the entire sentence if it appears several times in the deep structure representation. For example, *John* is a better prompt in sentence (34a) than in sentence (34b), since only in sentence (34a) is it the subject of two deep structure predicates (Blumenthal 1967; Blumenthal and Boakes 1967; Wanner 1968).

The studies on locating nonspeech sounds in sentences offered evidence that listeners actively compute deep structure during speech perception. The perceived location of a brief nonspeech sounds is influenced by the "clause" structure at the *deep* structure level of representation, as shown in sentences (35a) and (35b) (Bever, Lackner, and Kirk 1969). The O indicates the objective location on the nonspeech sound, and the R indicates the relatively dominant perceived location of the sound.

(35) a. The general defied the troops to fight.
O R
b. The general desired the troops to fight.
R O

The fact that the nonspeech sound is heard after *troops* in (35a) makes sense if we recall that the noun phrase after *defy* appears in the same deep structure clause as *defy*. The tendency to hear the nonspeech sound before *troops* in (35b) follows from the fact that *troops* is not in the same deep structure clause as *desire*.

These studies vindicated the levels of representation proposed by transformational grammar as active participants in sentence perception. Listeners actively compute the deep structure level of representation as an active 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, the question emerged: How does speech processing elicit grammatical levels of representation?

2.4.2 Perceptual Strategies

One possibility is that the processes that form perceptual representations are entirely distinct from grammatical operations (Fodor and Garrett 1967; Fodor, Bever, and Garrett 1974). In a version of this view, listeners acquire an extragrammatical set of *perceptual strategies* that map surface structures onto deep structures (Bever 1968b, 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 action patient." 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. The fact that passives violate the NP-verb-NP strategy explains the difficulty of the passive. Similarly, the strategy explains the salience of NP-verb-NP structures in sentences with reduced relative clauses, as in *The horse raced past the barn fell*. It also explains the preference for the transformed version of sentences with particle/prepositions, since the transformed version places the patient directly after the action.

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 even though there is no concomitant change in their linguistic knowledge. During this period, for example, older children actually perform more poorly than younger children on acting out passive sentences (Bever et al. 1969; Bever 1970a; Maratsos 1974; DeVilliers and DeVilliers 1972; Slobin and Bever 1982).

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, and Pribram 1960). The perceptual strategies can clear immediate memory by recoding surface sequences on deep structure propositions. This reasoning motivated the hypothesis that the proposition is the unit of recoding during speech perception (Bever 1970a; Bever, Kirk, and Lackner 1969; Fodor, Bever, and Garrett 1974). The recoding hypothesis 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 and Bever 1969), detection of clicks is poor (Bever, Hurtig, and Handel 1975), tones are hard to discriminate (Holmes and Forster 1972), and evoked potentials are suppressed (Seitz 1972), 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 and Bever 1969; Bever, Garrett, and Hurtig 1973). At first, it appeared that the surface clause was the unit that defined the scope of the perceptual strategies (Fodor and Bever 1965). Then it appeared that the deep structure "sentence" was the relevant unit (Bever, Lackner, and Kirk 1969). Finally, it became clear that the relevant unit as a psychological object was as a "functionally complete" proposition consisting of fully specified grammatical roles (Carroll 1978; Carroll and Tanenhaus 1975).

The decade of research on speech processing between 1965 and 1975 offered an account of the relation between certain linguistically defined representations 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 in the rare cases when the behavioral strategies fail (Bever 1972).

The moral of this experience is clear. Cognitive science 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*. For example, 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 available for 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 of these language behaviors has its own neurological and physical constraints. In particular, humans ordinarily both talk and comprehend. Yet 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 phase of research 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 in which the relevant "good figures" would be given a theoretical foundation. The grammar could 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.

2.5 The Search for a Behaviorally Relevant Grammar

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 common deep structure represents the relation between an active and a passive construction, a common deep structure can underlie both an active and a causative structure, expressing the structural relation between the following two sentences:

(36) a. John killed Bill.
b. John caused Bill to die.

A common deep structure can also underlie an active and a superficially remote construction, as in:

(37) 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 1971; Lakoff and Ross 1976; McCawley 1968a, 1968b; 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 1972). 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. However, this position was not tenable. Semantic representations themselves either must be stated in a normal form, or they must comprise all conceptual knowledge. Either way, the generative semantics program collapsed. If the semantic representation is in a canonical, propositional, or some other normal form, the model still included 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 1971, 1972). This conclusion was a reductio ad totam, since it is impossible to define grammatical rules for all possible knowledge (Katz and Bever 1975). For this and other reasons, generative semantics was largely abandoned as a grammatical project (see Katz and Bever 1975; Newmeyer 1980, for general discussions).

The wars in linguistics highlighted the problem that linguistic theory changes like quicksilver. Just when psychologists think they have their hands on it, 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, but 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 *John is eager to please* is not only the subject of two deep structure clauses, 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 (Bransford and Franks 1971, 1972). For example, 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, namely, that syntax is not real, and that only probabilistic language behavior exists. What appears to be evidence of syntax is the gratuitous by-product of perceptual processes (Clark and Haviland 1974), speech behavior learning strategies (Bates 1976; Bates and MacWhinney 1982), or an overimaginative and flawed linguistic methodology (Schank 1973, 1976).

Nonetheless, the same crucial empirical problem remained for both linguistics and psychology. That is, sentences have behaviorally relevant inner forms, and 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 language learning.

During the 1980s there were some clear trends in the evolution of syntactic theory that led to a renewed collaboration between psychology and linguistics. The overwhelming shift was away from restrictions expressed in transformations to verb-based structural restrictions. Three syntactic models emerged in this direction (see Sells 1985 and Tanenhaus 1988 for a review of linguistic and psycholinguistic issues, respectively, and chapter 3).

Government and Binding This theory was the original transformational model with the following changes:

- 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 specify the relation between the deep and the surface level. A derivation is possible only if there are particular transformations that construct it. In the government-and-binding 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).

Lexical-Functional Grammar A separate model that may be a notational variant of the government-and-binding model treated most transformational variations as specified within the lexicon itself (Bresnan and Kaplan 1983). (Note that this kind of model still treated recursion and unbounded dependencies in the same manner as standard transformational models.)

Generalized Phrase Structure Grammar A variant of phrase structure proposed to describe sentence structure. Unlike earlier models of phrase structure, this model included an enriched repertoire of the kinds of abstract constituent nodes that are intended to overcome the empirical inadequacy of earlier phrase structure systems (Gazdar et al. 1985).

Each of these syntactic models had its champions as a model of speech behavior. For example, Frazier (1985) and Freedman and Forster (1985) worked within government-and-binding theory, Bresnan and Kaplan (1983) in Lexical-Functional Grammar, and Crain and Fodor (1985) in generalized phase structure grammar. The common argument was 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 persisted: it is hard to state consistent linking assumptions between the formal model and the behaving person. A consistent theory of direct implementation of grammatical knowledge in behavior continued to be elusive. (See Bever 1987 for a discussion of attempts to formulate direct linguistic linking assumptions.)

2.6 The Reemergence of Associationism

In the 1980s several developments in cognitive psychology and the study of language exhibited a certain recycling of previous approaches. Concomitant with the increasing emphasis on lexically driven grammatical analyses, a large body of research was devoted to "lexical access." This research focused on the recognition, immediate processing, and recall of single words in various contexts including sentences (for reviews, see Simpson 1981; Seidenberg and Tanenhaus 1986; Tanenhaus, Carlson, and Seidenberg 1985).

At the same time, sentence-level syntax offered only a 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 such as paired constructions like sentences (38a) and (38b) or (39a) and (39c), which are identical in their surface representation but different at the underlying level, repeated here:

(38) a. John was eager to please.
b. John was easy to please.

(39) a. The general defied the troops to fight.
b. The general desired the troops to fight.

or paired constructions like sentences (40a) and (40b):

(40) a. In her *hope of marrying Anna was impractical.*
b. Harry's *hope of marrying Anna was impractical.*

which can 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, it is reasonable to expect that whatever we find out today about processing words will be relevant to an integrated theory of language behavior in the future. Unfortunately, our 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.

Associationism returned reinvigorated at the end of the 1980s, in the form of "connectionist" models of complex behavior (Feldman 1986; McClelland and Rumelhart 1986), which are now in high fashion in cognitive science. *Connectionism* is a descriptive term for a class of theories based on associative connections between elements: in this sense, it is a framework for associative models, which makes no particular claims until utilized in a particular way. 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 a necessary link to behaviorism.

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 large 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, Hinton, and Williams 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 as a whole allegedly recapitulates behavior only approximated by the rule (see Rumelhard and McClelland 1968, chap. 18; Feldman 1986).

The connectionist approach to language is often an intentional merging of competence and performance in a manner opposite the conflation of transformational grammar and a performance model, discussed above (see, e.g., Elman 1990). Thirty-five years ago, some psycholinguists claimed that the grammar is the performance system; the corresponding connectionist claim would be that the performance system is the reality that the grammar captures only approximately and vaguely.

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 complementary reason, namely, that 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: give them an articulate structure and enough internal associative connectivity 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 unmanageably 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 relevant contingent connections?

2.7 Conclusion: The Mystery of Structure

The listing of classical approaches to the sentence is 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. Elaboration of structural models of linguistic knowledge has actually laid out a number of consistent features about the nature of language for the psychologist to ponder. The reemergence of a powerful framework for associative modeling offers new ways to embrace frequency-based information within processing models, without embracing behaviorist learning principles as well.

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 the knowledge of syntactic structure improve the acuity of the ear?

Note

This chapter is based on a chapter by Bever that appeared in Hirst 1988. We have received comments from many linguists and psychologists on an earlier version of the 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?" We are not convinced that most linguists were—and are—not just as confused as most psychologists on most matters. However, such responses prompt us 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 from 1955 to 1980, and a bit beyond. 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.

This excerpt from

Sentence Comprehension.
David J. Townsend and Thomas G. Bever.
© 2001 The MIT Press.

is provided in screen-viewable form for personal use only by members of MIT CogNet.

Unauthorized use or dissemination of this information is expressly forbidden.

If you have any questions about this material, please contact cognetadmin@cognet.mit.edu.