

## Chomsky without language

ALBERT S. BREGMAN\*

*McGill University*

After reading Chomsky's *Aspects of the Theory of Syntax* (1965), I became intrigued by the idea of composition. Consider the sentence "The stone was thrown by the angelic little boy". In the *Aspects* approach, the passive voice was viewed as being a simple part of the sentence at the deep structural level, but as coming out on the surface structure as a word-order arrangement. Many features of sentences were treated by Chomsky in this way. Furthermore, the exact form that a particular deep pattern would take when it appeared in the surface structure of the sentence would differ, depending on the other deep patterns which appeared in the same sentence; yet at the deep structural level it was always the same.

This made me think of a problem posed by William James which can be stated as follows. If a piece of behavior or thought is constructed out of a set of distinct units, as the associationists had claimed, how can two observations be explained: (a) the "same" act or thought never reoccurs in exactly the same form on two occasions; (b) the so-called units flow into one another and overlap; we cannot determine the exact end of one and the beginning of another. It seemed that Chomsky had addressed these two problems as they appeared in syntactic structures, and had solved them by postulating a double level of description. His deep structural description labelled the separate identity of each syntactic form and made this identity available to rules of composition. His surface description concerned the patterning that resulted from the composition, a patterning in which the deep forms were modified, merged and modulated. Thus it seemed that Chomsky could eat his cake and have it: he could have a restricted number of invariant forms at the deep level and an infinity of unique patterns at the surface. This situation was made possible by the fact that the deep forms were put together not by the simple concatenation proposed by the associationists but by a rather complicated process of composition.

Chomsky believed that this special aspect of language, deep structure, made it impossible to learn language by association, a process which simply concatenated elements. Therefore the obvious capacity of humans to both perceive

---

\*Reprint requests should be sent to Dr. A. S. Bregman, Department of Psychology, McGill University, 1205 Docteur Penfield Avenue, Montreal, Quebec H3A 1B1, Canada.

and produce language pointed to the existence in humans of a powerful processing mechanism, a language acquisition device, specifically designed to extract the deep regularities of language from observed sentences. This capacity, in the opinion of Chomsky, marked language as unique and a special sign of our humanity. The idea that human psychological events are created by the composition of underlying abstract elements was wonderfully attractive and the more I thought of it, the more it seemed to apply to many aspects of human psychology outside the field of language.

Perception, it seemed, could also be seen as a process of composition. In this view, our perceptual process composes some basic notions to generate a structure that matches the current input. An example, drawn from vision, goes as follows. When we perceive a shiny table top, the distribution of light that hits our retinas is interpreted as arising from at least two distinct (deep) aspects of the table top, its surface color and its glossiness. Furthermore, these two ideals must be taken in *composition*, not one at a time, to account for the pattern of color that hits our eyes, how it changes as our head moves, and how it is slightly different in the two eyes. The resulting percept can be seen as having two levels, a deep one in which the color and glossiness are registered as separate features of the table top and a surface level in which the actual distribution of light is accounted for. In Chomsky's system, transformational rules were given the job of describing the modifications of the deep patterns as they became realized in a sentence. By analogy, the perceptual system can be said to have rules which know (in some sense) how the color and glossiness of surfaces interact in shaping the final distribution of light. We can call these the rules of *composition* for deep visual regularities, and can assign the name *ideals* to the underlying visual concepts themselves.

If simple aspects of the world, say surface color and glossiness, compose their effects upon the evidence that reaches our senses, our perceptual systems face the same horrendous detective problem faced by our language system, that of decomposing the evidence back down into its constitutive regularities. The syntactician refers to this as the *parsing* problem. The problem of finding the deep regularities that have shaped a pattern of data is also the problem faced by the scientists; so our senses must practise scientific explanation in their own modest ways. When they attain two adequate explanations of the data by performing two different compositions of underlying ideals, we have an ambiguous figure. When the perceptual system's choice of underlying factors is different from the one used by the external world in shaping the perceptual input, then we have an illusion. When internal factors dominate the perceptual composition so that it fails to be driven by sensory data, we call it imagery, hallucination or dreaming, depending on other aspects of our state of consciousness.

For about twelve years now, I have been trying to apply this concern with composition and its converse, the recovery of regularities, to problems in auditory perception. One of the most fundamental ways in which simple sources of sound can compose their effects upon the wave pattern that reaches our ears is the process of mixing. Outside the psychoacoustic laboratory, what we hear is always a mixture of the effects of different vibrating sources. An adequate perceptual representation of this input will contain, at the deep level, a separate description of each source in terms of such auditory primitive concepts as loudness, pitch, and location, and some not-so-primitive concepts such as the words spoken by one of the vibrating sources. The actual acoustic consequences of the process of mixing will have to be represented in a surface level of description so that it can be tested for adequacy against the sensory input. One consequence of the mixing of two sounds is the masking of one sound by another. The auditory system has a way of discounting the effects of masking. Whenever an input can be matched by a perceptual construction in which a loud sound is represented as interrupting a soft sound, the masking of the evidence for the soft sound at that point in time is not taken as implying the discontinuation of the soft sound. During masking, the soft sound is represented at the deep level (and indeed actually heard) as continuing through the interruption. Another consequence of the mixture of two patterns of sound is that the elements of the two are interwoven in time. Yet our ears can recover a separate description for each of them under an astonishing variety of conditions. I have reviewed elsewhere some of the techniques that the ears use to recover the underlying streams of sound, the so-called heuristics of auditory streaming (Bregman, 1978, 1981). Here, it is sufficient to point out that the auditory system can be seen to be representing the input as a composition of sources, a "source" being one type of auditory ideal or deep concept.

I think, then, that Chomsky's approach has pointed to a general process of mind, that of building a mental structure for a newly encountered event out of a limited number of deeper concepts that have existed prior to the event (Bregman, 1977). There is good reason to believe that this process occurs in the generation of behavior as well as in perception. Chomsky pointed out that the structures that the speaker and hearer of a language have to deal with are the same ones. This view led him to say that his theories described the competence of an ideal speaker-hearer of a language. This approach went contrary to an established tradition in psychology, one in which the concepts used to describe perception and behavior were entirely different. Skinner gave voice to this idea quite clearly in describing the difficulty of explaining how a child could imitate adult speech. He argued that "the formal similarity of stimulus and response need not make the response more likely to occur or

supply any help in its execution. In fact there is *no* similarity between a pattern of sounds and the muscular responses that produce a similar pattern" (Skinner, 1957, p. 59).

While it is possible that an account of the behavior of very young infants may have to be written in a vocabulary that is quite independent from that of perception, the example of human language seems to show that, at least in this field, the behavior of an adult has to be expressed in the same concepts as that used to describe percepts. It is important to see, however, that this equivalence is not a special feature of the domain of language. We also perceive meaning in the non-linguistic behavior of others. For example, I may see a man's struggles with a sack of potatoes as being organized around the goal of getting the sack onto his back. If I am correctly interpreting his actions, both his behavior and my perception are being organized by this same goal. In his behavior, the goal organizes the transitions from one movement to the next. For me, my perception of his goal accounts for the transitions between one of his movements and the next and allows me to predict his future movements.

It is possible to argue that any type of behavior is a composition formed from underlying abstract "ideals" and that these very same ideals form the basis for our understanding of the behavior of other people.

Just as the ideals in language (the syntactic forms for example) are very abstract, so are those that generate non-linguistic behavior. Consider the act of *grasping*. I can grasp with my finger and my thumb, with two fingers, with my knees, with my elbows, with chopsticks, with one knee and a chopstick and so on. The way grasping is *instantiated* in my behavior depends on what other elements are appearing at the same time. I must obviously possess a formula that defines grasping, one that is no less abstract or far from the surface of an action than is the passive element in Chomsky's account of the syntactic form of a sentence.

Why are we built this way? To answer this question we need to look more closely at what function the abstract elements or ideals serve. My view is that in any individual situation, a composition of basic ideals is created. In the field of perception, the composition serves as a model of the situation. The model is built from a set of more fundamental patterns, each one capturing some important regularity of our world. Because the model of the specific situation is built from a vocabulary that captures fundamental regularities, such as *surface, slant, object, above, person, goal*, and so on, the model serves as a powerful calculating system. Each regularity carries with it rules for calculating what may happen next or what actions may be carried out successfully.

In the field of action, the composition has a parallel function. It organizes behavior around abstract action-ideals, such as *grasp, hurry, go to, be polite*,

*sell*, and so on. Each such action-ideal captures some regularity in our interaction with the physical or social world and represents a type or aspect of coherent and effective action.

One can see, furthermore, that thought sequences are compositions whose structure also derives from ideals. In human thought, these ideals have enormous variety, ranging from such specific entities as *Felix, the cat* to such all-purpose organizational patterns as *means-end*, or *classification*, and such computational patterns as *negative* or *hypothesizing*. I mean these italicized words to refer to patterns that structure our thought and not to the words that we use to label such patterns. However, every word of our language links to one of these ideals.

The human child's development of these underlying ideals constitutes what Piaget called "the development of objectivity", a phrase that he employed to refer to the process by which inner patterns of the mind come to have forms that match the forms of external reality. Any adequate mental system for modelling reality will have to have elements whose own forms and patterns reflect external ones.

While the brains of other species and the computational systems which they embody must also have this property, the human brain has developed an astounding capacity to handle a lot of these modelling patterns at once. Furthermore it can extend its intellectual system by adding more and more of them. This is the process that Piaget studied so extensively. As the human child develops, it builds up a stock of mental patterns of control ("schemes") that, in combination, control thought and action. The key word in the previous sentence is *combination*, and in that word lies the greatest challenge for the psychology of the future.

For in considering an extendible modelling system that uses composable patterns, several critical questions emerge. How can patterns of different levels of abstraction combine to model a specific event? What calls them forth and composes them appropriately? Remember that they are not to be strung together like beads on a chain but must be composed to form a structure that describes the various aspects of the current situation, the molar and the molecular, the abstract and the concrete, at the same time. The understanding of a single sentence of language must surely involve the participation of hundreds, if not thousands, of computational schemes, some innate and some learned. How can they all be pulled in at the right time and cooperate smoothly? This is a problem in system architecture, the solution of which may some day tell us how it is possible for a being to be truly intelligent. Closely linked to this issue of architecture is the problem of cognitive growth, or learning as it used to be called. How can a learning process create new patterns

for modelling previously uncaptured aspects of reality and add them to the current set without playing havoc with the already-existing system?

I believe that the critical questions that I have mentioned, those concerning the combinability and extendibility of the set of ideals, will have to serve as the focus for any serious future theory of learning. Indeed, one could speculate that these two problems are so closely linked that it is the same design feature in the human brain that both permits it to combine ideals as freely as it does and to add new ones to its repertoire. I believe that a theory of learning that addressed these issues could serve as a point of unification of psychological theory. Perceptual theory, for example, would be seen to fit in smoothly with learning theory because both would be concerned with ideals—perceptual theory with how we use them to model situations and learning theory with how we acquire and modify them. An extended version of Piaget's theory would probably be closest to the mark.

Paradoxically, it seems that despite the fact that Chomsky has identified himself with the nativist tradition in psychology, his most important influence may ultimately lie in having provoked the development of an adequate theory of learning.

## References

- Bregman, A. S. (1977) Perception and behavior as compositions of ideals. *Cog. Psychol.* 9, 250–292.  
 Bregman, A. S. (1978) The formation of auditory streams. In Requin, J. (ed.), *Attention and Performance VII*. Hillsdale, NJ, Lawrence Erlbaum Associates.  
 Bregman, A. S. (1981) Asking the “what for” question in auditory perception. In Kubovy, M. and Pomerantz, J. R. (eds.), *Perceptual Organization*. Hillsdale, NJ, Lawrence Erlbaum Associates.  
 Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, Mass., MIT Press.  
 Skinner, B. F. (1957) *Verbal Behavior*. New York, Appleton-Century-Crofts.