JOURNAL OF MEMORY AND LANGUAGE 27, 107–123 (1988)

Connectionist Models and Psychological Evidence

JAMES L. MCCLELLAND

Carnegie-Mellon University

In this article I review the connectionist framework for modeling psychological processes, and I examine the role of connectionist models in empirical psychology. I illustrate how modeling can reveal the empirical implications of general principles, and I point out that the connectionist framework is particularly apt for formalizing certain proposed processing principles. The framework has led to the discovery of new classes of explanations for basic findings; it has led to unified accounts of disparate or contradictory phenomena; and it has shed light on the relevance of certain types of evidence for basic questions about the nature of the processing system.

When the study of cognition took hold in the 1960s, it was common to think of the human information-processing system as a device much like a von Neumann computer. Processing was viewed as a sequence of discrete operations (Sternberg, 1969). Memory consisted of a set of separate stores (Atkinson & Shiffrin, 1968; Waugh & Norman, 1965). Complex processes were characterized by flowcharts specifying a sequence of steps to be taken under rigid control of an executive (Clark & Chase, 1972). For many years, this view has prospered. Theories of language processing (Marcus, 1980; Woods, 1970) and language acquisition (Berwick, 1985), of problem solving (Newell & Simon, 1972), of comprehension (Schank, 1981), and of knowledge representation (Minsky, 1975) have all come to psychology from research on the implementation of intelligent processing on von Neumann computers, as have several general theories of the nature of the human information processing system (e.g., J. R. Anderson, 1983; J. R. Anderson & Bower, 1973).

Now a growing group of researchers is beginning to explore a different framework for thinking about cognitive processes. These researchers generally accept the computer metaphor as a useful approximate description of the macrostructure of human thought. But they have come to feel that an alternative framework may be more appropriate for characterizing the microstructure of cognition.

In this article I will describe this alternative, here called the connectionist framework, and then consider the role this framework can play in experimental cognitive psychology. In brief, my argument will be that modeling is often crucial if we are to understand the implications of certain kinds of basic principles of processing that might be proposed, and that the connectionist modeling framework provides a very apt formal language for embodying several principles whose implications deserve to be explored. I will also point out that the connectionist framework has led to the discovery of new principles, that it has aided in the construction of specific models that organize disparate and sometimes contradictory phenomena, and that it has

This article is an elaborated version of a talk given at a symposium of the same title at the Annual Meeting of the Psychonomics Society, New Orleans, LA, November 1986. I would like to thank Dr. P. Carpenter and the other participants of the symposium for helpful suggestions. Requests for reprints should be sent to J. L. McClelland, Department of Psychology, Carnegie-Mellon University, Pittsburgh, PA 15213.

0749-596X/88 $3.00

helped us to clarify what kinds of evidence bear on distinguishing between alternative hypotheses about aspects of human information processing.

The connectionist framework is also beginning to lead to new empirical research. This point is considered briefly in the present article; the articles that follow in this special issue illustrate it in considerable detail (Dell, 1988; Elman & McClelland, 1988; Gluck & Bower, 1988). The articles by Estes (1988) and Massaro (1988) are commentaries on the approach.

THE CONNECTIONIST FRAMEWORK

The term "connectionist models" was introduced by Feldman (1981; Feldman and Ballard, 1982). In these papers, the term is used to refer to a class of models that compute by way of connections among simple processing units. Another phrase often used to describe some connectionist models is *parallel distributed processing* or PDP models (McClelland, Rumelhart, and the PDP Research Group, 1986; Rumelhart, McClelland, and the PDP Research Group, 1986). PDP models are instances of connectionist models that stress the notion that processing activity results from the processing interactions occurring among rather large numbers of processing units.

In this article I intend the phrase "the connectionist framework" to encompass all kinds of connectionist models. The framework may be thought of as providing a set of general assumptions about basic aspects of information processing, and a set of soft constraints on the range of specific assumptions that might be made. In what follows I consider each of several aspects of an information-processing system. I describe the general assumptions connectionist models make about these aspects and I characterize some of the specific assumptions that might be made. The presentation draws heavily on Rumelhart, Hinton, and McClelland (1986), which can be consulted for further details.

Primitives and Their Organization

Like all cognitive models, connectionist models must propose some building blocks and some organization of these building blocks. In connectionist models, the primitives are *units* and *connections*. Units are simple processing devices which take on activation values based on a weighted sum of their inputs from the environment and from other units. Connections provide the medium whereby the units interact with each other; they are weighted, and the weights may be positive or negative, so that a particular input will tend to excite or inhibit the unit that receives it, depending on the sign of the weight (we shall return to these matters when we consider the dynamics of processing below).

Any particular connectionist model will make assumptions about the number of units, their pattern of connectivity to other units, and their interactions with the environment. These assumptions define the architecture of a connectionist model. The set of units and their connections is typically called a *network*.

It should be noted that a very wide variety of architectures is possible. Two are shown in Figs. 1 and 2. One of these, in Fig. 1, from the distributed model of memory examined by McClelland and Rumelhart (1985), shows a set of completely interconnected units, each receiving input from the environment and each projecting back to the environment. In some sense, the network in this figure is the most general possible connectionist architecture, in that all others involve restrictions of this general case. For example, some units may receive no input from the environment; some may send no output outside the net; and some of the interconnections among units in the network may be deleted. There may, furthermore, be restrictions on the values of some of the connections. In the general case, each may be positive or negative, but the architecture may prescribe,

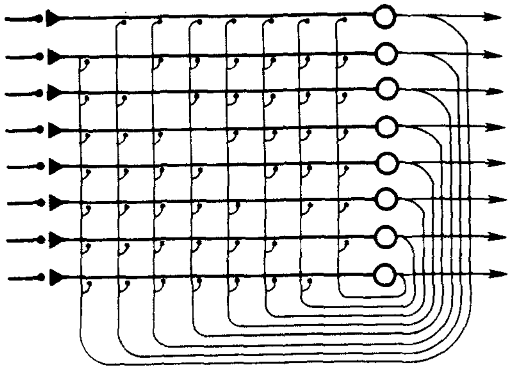

FIG. 1. A fully connected autoassociative network, with connections from each unit to every other unit. Each unit receives input from outside the network and sends output outside the network. All connections may be either positive or negative in this very general formulation. (From "Distributed Memory and the Representation of General and Specific Information," J. L. McClelland and D. E. Rumelhart, 1985 by *Journal of Experimental Psychology: General*, 114, p. 162. Copyright 1985 by the American Psychological Association. Reprinted by permission.)

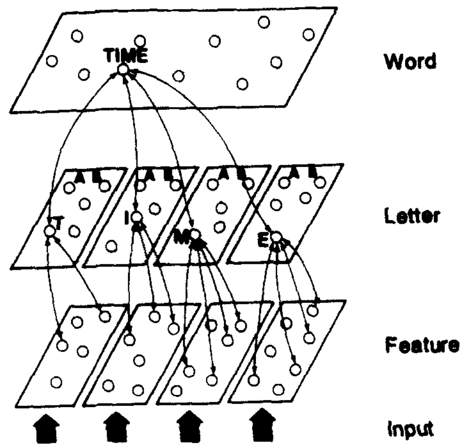

FIG. 2. A sketch of the network used in the interactive-activation model of visual word recognition (McClelland & Rumelhart, 1981). Units within the same rectangle stand for incompatible alternative hypotheses about an input pattern and are mutually inhibitory. Bidirectional excitatory connections between levels are indicated for one word and its constituents. (From "Putting Knowledge in its Place: A Scheme for Programming Parallel Processing Structures on the Fly" by J. L. McClelland, 1985, *Cognitive Science*, 9, p. 115. Copyright 1985 by Ablex Publishing, Reprinted by permission.)

for example, that a certain group of units have mutually inhibitory connections of fixed strength.

Figure 2 gives an example of a more restricted architecture, from the interactive activation model of visual work recognition (McClelland & Rumelhart, 1981). In this model, units stand for hypotheses about displays of letter strings at each of three levels of description: a feature level, a letter level, and a word level. There are excitatory connections (in both directions) between mutually consistent units on adjacent levels, and inhibitory connections between mutually inconsistent units within the same level. Thus the unit for *T* in the first letter position excites and is excited by the units for features of the letter *T*, as well as the units for words that begin with *T*. This unit also inhibits, and is inhibited by, units for other letters in the same letter position.

Active Representation

Representations in connectionist models are patterns of activation over the units in the network. In some ways, these kinds of patterns are similar to representations in other frameworks; after all, representations in a computer are ultimately patterns of 0s and 1s. There are differences, however. For one thing it is quite natural for connectionist representations to be graded, in the sense that each unit's activation need not be one of two binary values. In some models, activations are restricted to binary or some other number of discrete values, but more typically each unit may take on a continuous activation value between some maximum and minimum. A more important difference is this: Connectionist representations are truly active, in the sense that *they give rise to further processing activity directly*, without any need for a central processor or a production-matching-and-application mechanism that examines them and takes action on the basis of the results of this examination.

Models differ in terms of the extent to

which individual processing units can be identified with particular conceptual objects, such as letters, words, and concepts. The models illustrated in Figs. 1 and 2 represent endpoints on a continuum. In the distributed model of memory, each conceptual object is thought of as a pattern of activation over a number of simple processing units. In the interactive activation model of work perception, on the other hand, each unit stands for a primitive conceptual object, such as a letter, a word, or a distinct visual feature. A large number of models lie between these two extremes (see Hinton, McClelland, & Rumelhart, 1986, and Feldman, 1986, for general discussions of the issue of distributed representation).

Processing

Processing in connectionist models occurs through the evolution of patterns of activation over time. This process is governed by assumptions about the exact way in which the activations of units are updated, as a function of their inputs. Updating can be synchronous (all units updated simultaneously) or asynchronous (units updated in random order). Updating generally occurs as follows. First, a *net input* is computed for each unit to be updated. The net input is the sum of the activations of all of the units that project to it, with each contributing activation weighted by the weight on the connection from the contributing unit to the receiving unit.[1] The net input may also include a *bias* term associated with the unit, as well as a term for inputs arising from outside the network. Thus for unit i, its net input is given by

$$\text{net}_i = \sum_j w_{ij}a_j + \text{bias}_i + i_i. \qquad [1]$$

The net input can then be used to set the new activation of the unit according to some monotonic but nonlinear function like the one shown in Fig. 3. Alternatively, the net input can be used to set the activation of the unit probabilistically to one of two discrete values (usually 1 or 0). Another possibility is that the net input may act as a force, tending to drive the activation of the unit up or down a small amount in each time step.[2]

It is typical to use some form of nonlinear activation function, so that the activation of a unit is not simply set equal to the net input or some weighted average of the net input and the previous activation of the unit. Nonlinearities are typically necessary for two reasons: (1) Linear networks are subject to explosive growth of activation due to positive feedback loops unless the weights are severely constrained (see Shrager, Hogg, & Huberman, 1987). (2) Many computations require a layer of nonlinear units between input and output. Without nonlinearities, multiple layers of units add no additional computational power over that offered by a single layer (see Rumelhart, Hinton, & McClelland, 1986, for further explanation).

Knowledge

Crucial to the very idea of cognition is the notion that information processing is guided by knowledge. We recognize the word *the* as a definite article because of knowledge we have about the relation between letter strings and linguistic forms. We infer that a spoon may have been used if we hear "The man stirred the coffee" because of knowledge we have about the kinds of instruments that are used for stir-

[1] In a slightly more general formulation, the net input may be the sum of *products* of the activations of groups of contributing units. In this formulation there is a weight associated with each product, rather than each individual contributing activation. These product terms have no special computational significance, since the effects of multiplicative interactions among inputs can be accomplished by extra layers of units (see Williams, 1986).

[2] Some variants of connectionist models (e.g., Grossberg, 1978) treat the excitatory and inhibitory inputs as separate forces, rather than aggregating them together in a single term.

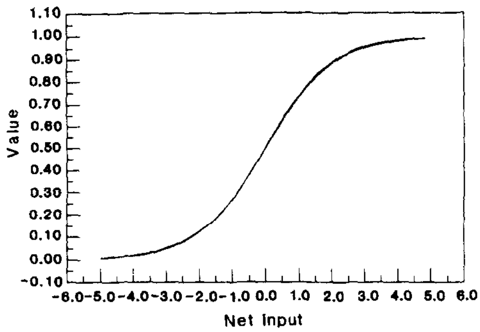

FIG. 3. The logistic function, a smooth nonlinear function that is frequently used in relating activations of units to their net inputs. This function is often used to set the activation of a unit to a value between 0 and 1, or to set the activation of the unit to 1 or 0 probabilistically, with the probability determined by the value of the function.

ring. In many models, these kinds of knowledge would be stored in tables. For example, information about *the* would be stored in a table called a lexicon, listing correspondences of letter strings and the linguistic objects they represent.

In connectionist models, the knowledge is stored in the connections among the processing units. This assumption works together with the assumptions connectionist models make about representations. An active representation on a set of units, together with the knowledge stored in connections, will give rise to new patterns of activation on the same or on other units.

Typically in connectionist models, connection strengths are real-valued. In models whose connections are set by assumption, it is typical to assume homogeneity of connection strengths as much as possible, to avoid excessive degrees of freedom. In models that learn, however, connection strengths are typically allowed to take on whatever values the learning process gives them; parsimony arises from the use of a homogeneous principle of learning.

Learning

If knowledge is in the connection weights, learning must occur through the adjustment of these weights. This weight adjustment process is assumed to occur as a by-product of processing activity. Some knowledge can in fact be built into connectionist models, in the form of initial connection strengths, before there has been any learning, but it is common to explore the limits of what can be acquired through connection strength adjustment with minimal prewiring. The initial architecture of the network serves to impose constraints on the learning process; these can in many cases greatly facilitate learning and generalization, if these constraints are appropriate to the problem the network is given to learn.

A wide variety of "learning rules" for tuning connections has been proposed. A recent review is provided by Hinton (1987). Generally, these rules state that the adjustment that is made to each connection should be based on the product of a presynaptic term, associated with the unit sending input through the connection, and a postsynaptic term, associated with the unit that is receiving input through the connection. For example, the *Hebb rule*, as used by J. A. Anderson (1977), makes the change in the strength of a connection proportional to the product of the activation of the sending unit and the receiving unit.[3]

Learning through connection strength adjustment is very different from learning processes in most other types of models. It is governed by simple mathematical expressions, and results in knowledge that is completely implicit, in that it is embedded inextricably in the machinery of pro-

[3] The Hebb rule is about the simplest connectionist learning rule, and it is limited in what it can do, so it has recently been somewhat less popular than other learning rules (but see Linsker, 1986a, 1986b, 1986c). Three learning rules frequently used in current connectionist models are the *competitive learning rule*, the *delta rule* or *least-mean-squared procedure*, and the *generalized delta rule* or *back propagation procedure* (see Hinton, 1987, for details).

cessing, and is completely inaccessible to introspection or report. However, it should be noted, that while the connection changes themselves are not accessible, the patterns of activation they make it possible to construct can be accessible to other parts of the processing system.

The Environment

Though it has been implicit in what I have said already, there is another aspect of connectionist models that deserves comment, namely, their *environment*. The environment consists of an ensemble of possible patterns that might be presented to the network. In most cases, these patterns are thought of as separate events, each one presented when the network is in a resting state, then left on until processing is complete. However, input patterns can have a richer temporal structure, or course; each event may consist of a sequence of events, or of a graded progression of input activations.

For networks with fixed connections, the environment simply defines the domain of inputs on which the network might be tested. For networks in which the connections are adjusted as a result of processing experience, however, the environment plays a crucial role in determining exactly what is learned. Thus models that aim to capture aspects of cognitive development through connectionist learning include among their assumptions a specification of the details of the experience that gives rise to the resulting developmental sequence. In many cases, these assumptions play a major role in determining the success or failure of the modeling effort.

The Spirit of the Thing

The connectionist framework is cast, not as a list of specific detailed assumptions, but as a set of *general principles* and some guidelines that provide weak constraints on the range of variants that fall within the scope of these principles. Indeed, as Rumelhart, Hinton, and McClelland (1986) noted, it is possible to build a von Neumann computer out of connectionist primitives, if they are organized in accordance with the von Neumann architecture. It thus becomes important to focus on the spirit of the connectionist framework. Generally, connectionist models of cognitive processes have been constructed expressly to exploit the capability for parallelism inherent in the approach, to make use of the graded capabilities of patterns of activation, and to capture the incremental nature of human learning in many tasks through the adjustment of connection strengths based on signals arising in the course of processing.

The Microstructure of Cognition

Finally, it is worth pointing out that the connectionist framework is not incompatible with other levels of description in cognitive science. Thus, there is nothing inconsistent with connectionist models in the claim that a cognitive system may traverse a sequence of states in a temporally extended cognitive task such as solving an arithmetic problem. According to the connectionist approach one would tend to view each such step in the process of solving the problem as a new state of the processing network. Indeed, Rumelhart, Smolensky, McClelland, and Hinton (1986) describe a network that performs a mental tic-tac-toe simulation, settling into a sequence of states representing the results of the successive mentally simulated moves made by each player.

There are important differences between conventional and connectionist models of sequential behavior. In connectionist models, the states need not be so discrete as they generally are in other models (Jordan, 1986; Rumelhart & Norman, 1982; Smolensky, 1986). Furthermore, the powerful constraint-satisfaction characteristics inherent in the connectionist framework are not typically exploited by conventional models of sequential processing. The idea that each step in a sequential process involves a massively parallel constraint satisfaction process seems like a promising starting place for a new way of thinking about the macrostructure of cognition.

The point that connectionist models characterize the microstructure of cognition applies not only with respect to time, but also with respect to the structure of the processing system and with respect to the description of the computational operations that the system is performing. Structurally, a processing system may consist of many parts, and for some purposes it may be adequate to describe its structure in terms of these parts and the flow of information between them. Computationally, too, it may often be useful and illuminating to describe what function a part of such a system computes without referring specifically to the role in this computation that is played by the specific units and connections. The claim is, though, that it will be necessary to delve more deeply than this to provide a full description of the mechanisms of cognition.

Are Connectionist Models Mere Implementations?

In allowing that there may be a macrostructure to thought, connectionists may seem to suggest that their models merely describe the implementation details of a processing system that would be best characterized more abstractly. However, we simply do not know exactly what level of description is the appropriate one for characterizing many behavioral phenomena. Those of us who have turned to connectionist models have done so because these models have seemed to provide exactly the right level of description for characterizing certain kinds of cognitive processes. Just where the bounds of usefulness of the connectionist framework may lie seems at this point to be one of the very open questions. Since there is little in cognitive psychology that we understand perfectly at this point, we are not at present in a position to say which aspects of cognition might be explainable without recourse to a model of the microstructure.

Psychological and Neural Modeling

Connectionist models are formulated for many different purposes. Some modelers are interested in characterizing actual neural circuitry. The framework is quite apt for this, and a growing group of researchers is pursuing this approach (Hawkins & Kandel, 1987; Gluck & Thompson, 1987; Rolls, in press; Zipser, 1986). This use of the connectionist framework is often a source of confusion in psychological circles, because most connectionist models in cognitive psychology are not aimed at this explicitly neural level. Instead, the aim of most connectionist models of psychological processes is to characterize processing at a level of description whose utility is assessable through behavioral, rather than neurophysiological, experimentation. This is certainly the case for the three papers that follow (Gluck & Bower, 1988; Elman & McClelland, 1988; Dell, 1988).

Thus, the connectionist models we are concerned with here remain functional characterizations of the mechanisms of thought. While it is true that they seem readily implementable in the brain, and this is often one source if inspiration in connectionist modeling, the models we are considering in these articles are all offered for their usefulness in characterizing aspects of cognitive processing, as it is revealed through experimental research.

Role of Connectionist Models in Empirical Research

In this section I consider what role connectionist models can play in the business of empirical psychological investigation. I begin with the question of the role of modeling in general, and then I turn to the particular merits of the connectionist framework. In what follows I have drawn heavily from my own experience with connectionist models, from the cascade model (McClelland, 1979) to a recent model for converting print to sound (Seidenberg & McClelland, 1987). These experiences are what lead me to invest my own energies in modeling, and they illustrate my own reasons for turning to modeling in general and connectionist modeling in particular.

The Role of Modeling: An Illustrative Example

To me, the central function of modeling is to make vague and complex ideas accessible and explicit, and precise enough to make their implications clear. My own history in turning to modeling illustrates this function. In the mid-1970s, my research on visual word recognition led me to the view that information processing might not involve a sequence of discrete stages, but a continuous flow of information through a series of processing levels (McClelland, 1976). When formulated in verbal terms, this idea had little force; its implications were obscure. However, it was possible to formulate an explicit information-processing model based on this assumption (the cascade model, McClelland, 1979) and from this model to derive several basic consequences. This model served, for one thing, to make the idea of an ensemble of continuous processes much more vivid and therefore, I believe, made it much easier to see that there was an alternative to the idea of discrete stages. Ten years later, it is difficult to remember the frame of mind of the mid-1970s, but I believe that at that time the idea of continuous processing was not widely considered and that the formulation of an explicit model in which that idea was embodied played a role in changing the way many of us think about mental processes. It certainly helped to consolidate my own thoughts on these matters.

Equally important, the cascade model led to the discovery of several implications of the idea that processing occurred in a system of continuous processes. Surprisingly, the cascade model revealed that additive effects of experimental factors could easily arise in a system of continuous processes. At the same time, the model demonstrated that in a system of processes in cascade, interactions of experimental factors did not necessarily indicate that the factors were influencing the same processing level.

Beyond these basic observations, the cascade model has played a role in leading to experimental studies investigating whether processing is continuous, as assumed by the cascade model, or discrete, as assumed by more traditional approaches (cf., Miller, 1982; Meyer, Yantis, Osman & Smith, 1985). The empirical picture that is emerging from these studies is quite rich and complex, and the circumstances under which there is continuous processing remain to be fully described. The point here is not to review this picture, but simply to note that the existence of the model, as a concrete embodiment of the idea of continuous processing in a multilevel system, has helped to stimulate an ongoing line of empirical research.

One point to take from this review of the cascade model is that the usefulness of a model is not simply a matter of its correctness. The cascade model illustrates how a model can make vague ideas precise and can allow the discovery of the implications of these ideas; the correctness of such a model can then be examined, once the implications have been made explicit. Of course, for a model to be interesting, it has to have some motivation; there must be some reason to suppose that the principles that it embodies are worth exploring. The point is that the cascade model and other models of complex processes should be taken as tools that help us understand the implications of possible assumptions that might be made about the characteristics of information-processing systems.

This goal of clarifying the implications of ideas often leads to deliberate simplification and elimination of detail, so that the consequences of the central assumptions can be made as clear as possible. The cascade model is certainly a case in point: it assumes a unidirectional flow of processing in a completely linear, multilayer system. Similar simplifications are often invoked in other fields of science, when some complex phenomenon needs to be understood. Sejnowski (1986) illustrates this point by describing the modeling of magnetism in iron. Ferromagnetism is modeled by replacing the complex structure of iron with a set of oriented point particles in a two-dimensional lattice, with each particle influencing the orientation of the nearest neighboring particles. He notes that such a model is successful if it exhibits large-scale qualitative phenomena (such as phase transitions) that are actually seen in the more complex objects (real iron bars) that are being modeled. The simplification is crucial because it would be impossible to model magnetism while taking the structure of iron into account in all of its details.

Why Connectionist Models?

Given this view of the role of modeling, one can ask, what framework is the best to use? To me the answer is simply whatever framework appears to be the most useful. The connectionist framework is useful for capturing certain kinds of assumptions about the way in which information processing occurs. Two principles that have motivated a good deal of my own explorations of connectionist models are (1) the idea that processing in a multilayered processing system is continuous, so that information accumulates gradually over time and is propagated as it builds up, and (2) the idea that this kind of continuous processing may be interactive, so that influences can be bidirectional, flowing both from higher to lower levels and from lower levels to higher levels. These ideas are well captured in the connectionist framework. They are generally not captured well in highly symbolic processing frameworks, in which the objects manipulated are discrete tokens that stand in an all-or-none fashion for some mental object.

Of course, there has been considerable recognition of the need for continuous, dynamic information processing within the traditional symbolic framework. A number of spreading-activation models of memory have been proposed (e.g., J. R. Anderson, 1983), as have a number of activation-based production systems (Thibadeau, Just, & Carpenter, 1982). In such systems, the effects of experience are, at least in part, a matter of gradual parameter adjustment. Many times this sort of model can capture some aspects of the assumptions mentioned above in a way that allows their implications to be explored.

In spite of the fact that some features of connectionist models can be captured in other frameworks, I have found the connectionist framework more workable for many applications, in part because there is less extra apparatus extraneous to the essential character of the ideas under exploration. This is not an argument in principle that the connectionist framework is the best for all purposes, but it is an argument that it may often offer a better match between the assumptions one wishes to explore and the tools the framework offers for exploring them. To illustrate, consider J. R. Anderson's (1983) demonstration that the interactive activation model of visual word perception (McClelland & Rumelhart, 1981) can be embodied in ACT*'s production system formalism. In this demonstration, Anderson made several extensions of the basic ACT* architecture specifically to capture the interactive activation process, and left unutilized major aspects of the production system architecture. The connectionist framework seems much more apt for capturing the essential assumptions of continuous, interactive processing that Rumelhart and I wished to explore in the interactive activation model.

Beyond this general point that the connectionist framework is often the most appropriate, there are other reasons why it has proved extremely fruitful to work within this framework. Models developed in this framework have been useful in three ways:

(1) These models have led to new interpretations of basic phenomena in the literature.

(2) They have provided unified accounts of what had previously been seen as highly disparate or even contradictory phenomena.

(3) They have clarified the relevance of certain kinds of evidence for adjudicating basic questions about the character of the information-processing system.

I will briefly consider each of these three points in turn.

New Interpretations

There are several examples of longstanding phenomena that have been given new accounts within the framework of connectionist models. Here I will focus on one theme that illustrates this, the discovery through connectionist models that sensitivity to the regularities of language might not require an explicit rule-formulation mechanism, together with the subsequent discovery that sensitivity to the exceptions to these regularities might not even require an explicit lexicon.

Perceptual facilitation of letters in pseudowords. Before the interactive activation model, perceptual facilitation for letters in pronounceable pseudowords (e.g., *mave*) had been reported several times (Baron & Thurston, 1973; McClelland & Johnston, 1977; Spoehr & Smith, 1975). The phenomenon had variously been attributed to familiarity of subword spelling patterns; to the application of spelling-to-sound conversion mechanisms; and to the use of a system of orthographic rules. Neither of the first two accounts appeared adequate, since pronounceability is not critical (Baron & Thurston, 1973), and letter-cluster frequency did not correlate with degree of perceptual facilitation in some experiments (McClelland & Johnston, 1977), though it did in others (Rumelhart & McClelland, 1982). The third account was never given a formulation explicit enough to allow a detailed comparison with data.

The interactive-activation model offered a different interpretation. It attributed perceptual enhancement to partial activation of word units by pseudowords. For example, *mave* produces partial activation of several words that share three letters with it (*gave, save, have, male, mate, mare,* and others) and these in turn produce feedback activation which ends up enhancing perception of all of the letters in *mave*.

It must be stressed that this is a radically different kind of interpretation than the others. It differs from them in two ways. First, it proposes that the same mechanism that accounts for perceptual enhancement of words over pseudowords also applies to the advantage of pronounceable pseudowords over random letter strings. Other approaches either ignored the perceptual advantage for words over pseudowords (though a small advantage for words over carefully matched pseudowords was typically found) or attributed it to a separate, lexical mechanism. Second, it suggests that evidence of sensitivity to regularities of language—in this case, perceptual facilitation for wordlike stimuli but not for letters in random strings—need not be taken as evidence for the explicit extraction of these regularities in a system of rules or familiar subunits. The model simply had a set of word-detector units, yet it could account for effects of orthographic structure on the perception of items that were not words.

Eliminating word detectors. The word detectors in the interactive activation model seemed at the time to be necessary, at the very least to account for the perceptual advantage of items which were themselves familiar, relative to others that were not. However, J. A. Anderson (1977; J. A. Anderson, Silverstein, Ritz, & Jones, 1977) pointed out that familiarity effects of particular patterns fall naturally out of a simple autoassociative model like the one shown in Fig. 1, in which there is no single unit representation for a particular stimulus item. Rather, changes to the connections in the network occasioned by experience with a set of patterns results in connection strengths that tend to enhance or sustain familiar patterns, and to complete or rectify distorted versions of such patterns, in much the same way that the interactive-activation model can enhance letter sequences that form familiar words.

Such observations suggested that performance on familiar items in a variety of tasks might be the result, not of the formation of a specific unit for each item, but of the changes in the pattern of connection weights among an ensemble of units in a network. Further, such models suggested that generalization—that is, the extension of performance on familiar items to novel items—might also be explicable in terms of the way in which an autoassociative network deals with patterns similar to those that gave rise to the changes in the pattern of connection strengths (see Knapp & Anderson, 1984; McClelland & Rumelhart, 1985).

This insight from connectionist modeling led to the development of a model of learning the past tense of English (Rumelhart & McClelland, 1986). The past tense is interesting because it is quite regular, and speakers seem to know the regular pattern, in that they can apply it to novel words; while at the same time, there are many exceptions. The model was interesting in that it captured both the regular correspondence and the exceptions, without having special units for specific exception words nor rules. Instead, there was simply a network of connections from a set of units for representing the base form of an input word, to a set of units for representing the word's past tense.

The network used in the model is shown in Fig. 4. The main part of the network—the part where the learning takes place—is called a pattern associator, and it has been used in a number of connectionist models (J. A. Anderson et al., 1977; Kohonen, 1977). The pattern associator consists of two sets of units, one to represent the root form of the word and one to represent the past-tense form, and connections from each unit in the first set to each unit in the second set. Learning occurs by exposing the network to patterns representing the root forms of words, allowing the net to compute a pattern representing the past tense form based on this input and the existing connection strengths, and then adjusting the strengths of the connections to reduce the difference between the actual output and the desired output. The adjustment occasioned by a single pattern presentation is assumed to be small, so that learning occurs rather gradually, through repeated trials.

In a model of this type, if there is a strong regularity in the mapping from input

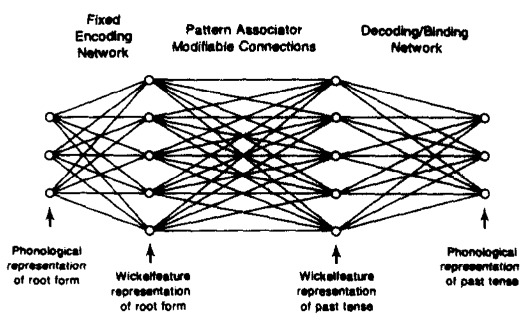

FIG. 4. The structure of the network used in modeling acquisition of the past tense. (From "On Learning the Past Tenses of English Verbs" by D. E. Rumelhart and J. L. McClelland, 1986, J. L. McClelland, D. E. Rumelhart, and the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the Microstructure of Cognition*. (Vol. 2, p. 123). Cambridge, MA: MIT Press. Reprinted by permission.)

to output, then the network will pick up on it; the changes produced in learning each pattern pair will work together synergistically, and there will be positive transfer among the patterns. If some of the pairs diverge from the regular pattern, they will be relatively more difficult to learn than regular patterns, and at certain stages in learning, they can be regularized.

This simple model, then, illustrates a truly novel interpretation of the "regularization error" frequently reported in the child language literature, the finding that children in the 3- to 5-year age range often say "taked" instead of "took" or "goed" instead of "went." Such errors have previously been taken as evidence that the child has discovered the rule, and have motivated a search for a characterization of the mechanisms that formulate, evaluate, and modify such rules. The past-tense model suggests an entirely new line of explanation.

The past-tense model does not account for all aspects of human performance in formulating the past tenses of words in English. Pinker and Prince (in press) and Lachter and Bever (in press) have amply documented these shortcomings. Thus, the model cannot be taken as a correct model of English past-tense formation in its present form. I mention it because it has served an important role, in spite of these shortcomings. It has brought a new kind of interpretation of lawful behavior into consideration, and it has stimulated a number of subsequent modeling efforts which transcend its limitations (Seidenberg & McClelland, 1987; Sejnowski and Rosenberg, 1987). A problem with the past-tense model is that it has no intervening layers of units between the input and the output. This limitation has been overcome by the development of the back-propagation learning algorithm (Rumelhart et al., 1986). More recent models make use of this algorithm to train multilayer networks. Whether such models will be able to meet the empirical challenges leveled by Pinker and Prince and by Lachter and Bever is a matter that is currently under investigation.

I hope this section illustrates the fundamental point that modeling can be worthwhile, even if a model that is developed in the course of exploring a set of issues is not fully consistent with the facts. The model may serve as a stepping stone in the development of a new way of thinking, even if it is only partly correct. It is a part of the ongoing process of working toward an understanding of the facts.

Bringing Phenomena Together

Another role that models can play is to bring together into a single, coherent picture a large body of phenomena which might otherwise appear to be disparate and unrelated. My experience with connectionist models has been extremely gratifying in this regard. Several of the connectionist models I have examined have helped to bring disparate or even apparently contradictory phenomena together into a single, coherent account.

The TRACE model of speech perception is one example of a connectionist model that provides a unified account for several widely disparate phenomena. The model covers phenomena ranging from categorical perception of phonemes in minimal context (e.g., /ba/ or /pa/) to cue trade-off studies examining the effects of varying several cues to phoneme identity, to studies of the effects of context in phoneme identification and the time course of these effects, to studies of the time course of spoken word recognition, to studies of the segmentation of sequences of phonemes into words. To my knowledge, no prior model has provided a treatment of even half of the sets of phenomena on this list. In TRACE, they all emerge from a homogeneous set of assumptions about an interactive-activation process between units representing hypotheses about the features, phonemes, and words present at different points in time within a temporally unfolding stream of speech.

Models can sometimes go beyond relating the previously unrelated to actually reconciling the previously contradictory. The interactive-activation model of visual word perception provides two cases in point. In the literature on visual word recognition, some studies have found reliable effects which might be attributable to familiarity of specific letter clusters (Rumelhart & McClelland, 1982), and other studies have not (McClelland & Johnston, 1977). At first (after the Rumelhart–McClelland experiment was finished but before the modeling work was done) these findings seemed simply contradictory to me. However, in the course of developing the model and applying it to the data, Rumelhart and I discovered that the specific experimental design used by McClelland and Johnston (1977) tended to counteract a tendency that the model had to favor pseudowords with more frequent letter clusters. Similarly, the finding reported by Johnston (1978) that contextual constraint did not appear to control accuracy of letter perception also appeared puzzling at first, especially in that others had found evidence supporting such an effect under other conditions (Broadbent & Gregory, 1968). However, the apparent contradiction disappeared when we discovered how the model worked on Johnston's high- and low-constraint items.

The details of the way in which these apparent contradictions were resolved by the model are somewhat involved, so I will not review them fully here. Rather, I will focus on just one part of the matter, namely, the reasons the model does not predict an effect of constraint in Johnston's (1978) experiment. I do so to illustrate the critical role the simulation model itself played in our ability to understand the overall pattern of results.

When a word is presented to the interactive-activation model, it sends activation to letter detectors, which in turn send activation to word detectors. Word detectors receive bottom–up activation in proportion to the number of letters they share with the pattern of activation at the letter level. Consider, in this light, what happens when the word *clue* is shown, and the subject is to be tested on the first letter, in a forced-choice with alternatives *c* and *b*. (It is always the case in these experiments that both alternatives form a word with the context letters, as *b* does with *_lue*.) The input will activate the letters *c*, *l*, *u*, and *e*. These in turn will activate the word unit for *Clue* and, to a lesser extent, the word units for other words such as *blue* and *glue*. These other words compete for activation with *clue* to some extent, but in this case there are only two of them (we disregard the very infrequent word *flue*). In contrast, consider the item *cake*, again with the first letter tested in a forced choice between *c* and *b*. This input will activate the letters *c*, *a*, *k*, *e*, and these in turn will activate the word *cake*, and, to a lesser extent, other words not beginning in *c*. In this instance there are about 10 such words. One might expect, then, that these partially activated competitors would tend to overwhelm *cake*, thereby reducing the feedback support for the *c* considerably, compared to the case of *c* in *clue*. Since such feedback is responsible for facilitation of perception of letters in words, according to the model, we expected that the model would in fact produce larger facilitation for *c* in *clue*, compared to *c* in *cake*.

These expectations turned out to be incorrect for two reasons. First, the inhibitory connections among word units allowed the word unit receiving the most excitation (in this case, *cake*) to keep other words from becoming very active, thereby attenuating the impact of the number of competitors. Second, this story leaves out the fact that items like *cake* produce partial activations of words other than *cake* that actually support the *c*. Thus *cake* activates *coke*, *cafe*, *cage*, *care*, *case*, and *cave*, as well as several words that do not begin with *c*; *clue*, on the other hand, activates only one other word (*club*) beginning in *c*. In general, Johnston's low-constraint words (items like *cake*) activate more friends of the target letter *c* as well as more enemies of the target letter, with the result that the expected disadvantage of high-constraint words is neutralized.

Once these points became clear, it became apparent that Johnston's findings were not in fact inconsistent with the interactive-activation model or with the results of other studies in which constraint effects were found (see McClelland & Rumelhart, 1981, for further details). In this case, the model was crucial for discovering that there was no inconsistency.

Relevance of Data

Beyond leading to new kinds of interpretations and unifying disparate phenomena, connectionist models have also helped to clarify what data might be relevant to deciding certain basic questions. This point can easily be made with reference to the cascade model. The model showed that factors affecting the asymptotic activations of units at any level of processing would interact with factors affecting either asymptotic activation or processing rate at any other level. This discovery indicates that finding an interaction between experimental factors cannot be taken as unequivocal evidence, as it once was routinely, that the two factors are directly affecting the same processing stage.

Another case of this can be found in the TRACE model of speech perception. As Elman and I point out in our paper in this collection (Elman & McClelland, 1988), top–down effects in TRACE tend to manifest themselves in simulations primarily as bias rather than as sensitivity effects in a signal detectability analysis. Thus a failure to find that context alters sensitivity does not rule out a top–down effect of context in perception. This is a key point about interactive models, and it has often been misunderstood. A second fact about TRACE is that it takes time in the model for contextual effects to influence processing at lower levels. Thus the finding that context effects often do not appear when subjects are induced to respond very shortly after a critical stimulus (Fox, 1982, 1984; Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979) is not inconsistent with the notion that contextual influences arise from feedback from higher to lower levels (see McClelland, 1987, for further discussion).

Both a failure to find sensitivity effects in a signal detection analysis and the failure to find context effects at short time intervals have been taken as evidence against the view that higher processing levels feed back activation to lower levels. These arguments make sense in some theoretical frameworks, as does the argument that additive effects of two factors on reaction time indicates that the factors affect different stages of processing. What the connectionist models have made clear is that the pattern of allowable inference can depend on basic assumptions which may at the least be open to question.

Can Connectionist Models Stimulate Research?

For me, a considerable part of the appeal of connectionist models lies in their power to evoke new interpretations of old findings, to provide a coherent account of a disparate body of known facts, and to help us understand more clearly the implications of empirical results for basic questions about the nature of processing. These contributions seem to me sufficient to warrant continued exploration of connectionist models as a useful adjunct to standard empirical investigation.

However, there is a further role that a theoretical framework can play, and that is to stimulate research. The question arises then, can the connectionist framework lead to new directions for empirical research or can it only give us new ways of thinking about the phenomena that have already been discovered by research emerging from other frameworks?

It is the burden of the collection of papers that follows here to argue that it can. Each describes a series of experiments that grew out of an attempt to test implications of a model constructed within the connectionist framework. Dell (1988) describes tests of the predictions of a model of speech production that accounts for the errors that subjects make in speaking. Elman and McClelland (1988) describe tests of a single but central prediction of the TRACE model of speech perception. Gluck and Bower (1988) describe tests that attempt to distinguish a simple adaptive network model of learning to make valid predictions in context from a Bayesian model. In all three cases, it will be apparent that the models have played a central role in the formulation of the question that led to the research.

It must be emphasized that the particular experiments all are tests of the particular models that motivated them, and so they cannot be taken really as tests of the connectionist framework in general. Indeed the framework cannot really be tested as such, nor can it directly stimulate experimentation. It is a formal system in which to construct models, not a model itself. The usefulness of a particular model is relatively easy to assess, but the usefulness of a framework is a matter that will only become clear in the long run. For the moment it appears that the connectionist framework provides a valuable set of tools for constructing models that are apt for a wide range of phenomena. Only the future can tell to what extent it will continue to be useful as work goes on. It seems most likely that the framework will prove its greatest worth as a stepping stone on the path to some future framework that we will only begin to grasp as we pursue research that combines both theoretical and empirical investigations.

REFERENCES

ANDERSON, J. A. (1977). Neural models with cognitive implications. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension* (pp. 27–90). Hillsdale, NJ: Erlbaum.

ANDERSON, J. A., SILVERSTEIN, J. W., RITZ, S. A., & JONES, R. S. (1977). Distinctive features, categorical perception, and probability learning: Some applications of a neural model. *Psychological Review*, **84**, 413–451.

ANDERSON, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard Univ. Press.

ANDERSON, J. R., & BOWER, G. H. (1973). *Human associative memory*. New York: Winston.

ATKINSON, R. C., & SHIFFRIN, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol 2.). New York: Academic Press.

BARON, J., & THURSTON, I. (1973). An analysis of the word-superiority effect. *Cognitive Psychology*, **4**, 207–228.

BERWICK, R. (1985). *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.

BROADBENT, D. E., & GREGORY, M. (1968). Visual perception of words differing in letter diagram frequency. *Journal of Verbal Learning and Verbal Behavior*, **7**, 569–571.

CLARK, H. H., & CHASE, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, **3**, 472–517.

DELL, G. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, **27**, 124–142.

ELMAN, J. L., & MCCLELLAND, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, **27**, 143–165.

ESTES, W. K. (1988). Toward a framework for combining connectionist and symbol-processing models. *Journal of Memory and Language*, **27**, 196–212.

FELDMAN, J. A. (1981). A connectionist model of visual memory. In G. E. Hinton & J. A. Anderson (Eds.), *Parallel models of associative memory* (pp. 49–81). Hillsdale, NJ: Erlbaum.

FELDMAN, J. A. (1986). *Neural representation of conceptual knowledge* (TR 189). Rochester, NY: University of Rochester, Department of Computer Science.

FELDMAN, J. A., & BALLARD, D. H. (1982). Connectionist models and their properties. *Cognitive Science*, **6**, 205–254.

FOX, R. A. (1982). Unpublished manuscript, Vanderbilt University.

FOX, R. A. (1984). Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human Perception and Performance*, **10**, 526–540.

GLUCK, M., & BOWER, G. (1988). Evaluating an adaptive network model of human learning. *Journal of Memory and Language*, 27, 166–195.

GLUCK, M. A., & THOMPSON, R. F. (1987). Modeling the neural substrates of associative learning and memory: A computational approach. *Psychological Review*, **94**, 176–191.

GROSSBERG, S. (1978). A theory of visual coding, memory, and development. In E. L. Leeuwenberg & H. F. J. M. Buffart (Eds.), *Formal theories of visual perception*. New York: Wiley.

HAWKINS, R. D., & KANDEL, E. R. (1987, July). Paper presented at a symposium on Parallel Distributed Processing, Oxford, England.

HINTON, G. E. (1987). *Connectionist learning procedures* (Tech. Rep. CMU-CS-87-115). Pittsburgh: Carnegie-Mellon University, Department of Computer Science.

HINTON, G. E., MCCLELLAND, J. L., & RUMELHART, D. E. (1986). Discussion and preview. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. I). Cambridge, MA: Bradford Books.

JOHNSTON, J. C. (1978). A test of the sophisticated guessing theory of word perception. *Cognitive Psychology*, **10**, 123–154.

JORDAN, M. I. (1986). Attractor dynamics and parallelism in a connectionist sequential machine. In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*. Amherst, MA.

KNAPP, A., & ANDERSON, J. A. (1984). A signal averaging model for concept formation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **10**, 616–637.

KOHONEN, T. (1977). *Associative memory: A system theoretical approach*. New York: Springer.

LACHTER, J., & BEVER, T. G. (in press). The relation between linguistic structure and associative theories of language learning—A constructive critique of some connectionist learning models. *Cognition*.

LINSKER, R. (1986a). From basic network principles to neural architecture: Emergence of spatial opponent cells. *Proceedings of the National Academy of Sciences (USA)*, **83**, 7508–7512.

LINSKER, R. (1986b). From basic network principles to neural architecture: Emergence of orientation-selective cells. *Proceedings of the National Academy of Sciences (USA)*, **83**, 8390–8394.

LINSKER, R. (1986c). From basic network principles to neural architecture: Emergence of orientation columns. *Proceedings of the National Academy of Sciences (USA)*, **83**, 8779–8783.

MARCUS, M. (1980). *A theory of syntactic recognition for natural language*. Cambridge, MA: MIT Press.

MASSARO, D. (1988). Some criticisms of connectionist models of human performance. *Journal of Memory and Language*, **27**, 213–234.

MCCLELLAND, J. L. (1976). Preliminary letter identification in the perception of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 80–91.

MCCLELLAND, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, **86**, 287–330.

MCCLELLAND, J. L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. London: Erlbaum.

MCCLELLAND, J. L., & JOHNSTON, J. C. (1977). The role of familiar units in perception of words and nonwords. *Perception and Psychophysics*, **22**, 249–261.

MCCLELLAND, J. L., & RUMELHART, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, **88**, 375–407.

MCCLELLAND, J. L., & RUMELHART, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, **114**, 159–188.

MCCLELLAND, J. L., RUMELHART, D. E., and the PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. II). Cambridge, MA: Bradford Books.

MEYER, D. E., YANTIS, S., OSMAN, A. M., & SMITH, J. E. K. (1985). Temporal properties of human information processing: Tests of discrete versus continuous models. *Cognitive Psychology*, **17**, 445–518.

MILLER, J. (1982). Discrete vs. continuous stage models of human information processing: In search of partial output. *Journal of Experimental Psychology: Human Perception and Performance*, **8**, 273–296.

MINSKY, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 211–277). New York: McGraw-Hill.

NEWELL, A., & SIMON H. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

PINKER, S., & PRINCE A. (in press). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*.

ROLLS, E. T. (in press). Information representation, processing, and storage in the brain: Analysis at the single neuron level. In *Neural and molecular mechanisms of learning*. Berlin: Springer-Verlag.

RUMELHART, D. E., & MCCLELLAND, J. L. (1982). An interactive activation model of context effects in letter perception: Part II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, **89**, 60–94.

RUMELHART, D. E., & MCCLELLAND, J. L. (1986). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. II). Cambridge, MA: Bradford Books.

RUMELHART, D. E., & NORMAN, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science*, **6**, 1–36.

RUMELHART, D. E., HINTON, G. E., & MCCLELLAND, J. L. (1986). A framework for parallel distributed processing. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. I). Cambridge, MA: Bradford Books.

RUMELHART, D. E., MCCLELLAND, J. L., and the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. I). Cambridge, MA: Bradford Books.

RUMELHART, D. E., SMOLENSKY, P., MCCLELLAND, J. L., & HINTON, G. E. (1986). Parallel distributed processing models of schemata and sequential thought processes. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. II). Cambridge, MA: Bradford Books.

SCHANK, R. C. (1981). Language and memory. In D. A. Norman (Ed.), *Perspectives on cognitive science* (pp. 105–146). Norwood, NJ: Ablex Publishing.

SEIDENBERG, M. S., & MCCLELLAND, J. L. (1987, November). *A distributed, developmental model of visual word recognition and naming*. Paper presented at the 28th Annual Meeting of the Psychonomics Society, Seattle.

SEJNOWSKI, T. J. (1986). Open questions about computation in cerebral cortex. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. II). Cambridge, MA: Bradford Books.

SEJNOWSKI, T. J., & ROSENBERG, C. R. (1987). Parallel networks that learn to pronounce English text. *Complex Systems*, **1**, 145–168.

SHRAGER, J., HOGG, T., & HUBERMAN, B. A. (1987). Observation of phase transitions in spreading activation networks: Behavioral change as topology of spreading activation network. *Science*, **236**, 1092.

SMOLENSKY, P. (1986). Information processing in dynamical systems: Foundations of harmony theory. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. I). Cambridge, MA: Bradford Books.

SPOEHR, K., & SMITH, E. (1975). The role of orthographic and phonotactic rules in perceiving letter patterns. *Journal of Experimental Psychology: Human Perception and Performance*, **1**, 21–34.

STERNBERG, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II*. Amsterdam: North-Holland (Reprinted from *Acta Psychologica*, 1969, **30**, 276–315).

SWINNEY, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, **18**, 645–659.

TANENHAUS, M. K., LEIMAN, J. M., & SEIDENBERG, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, **18**, 427–440.

THIBADEAU, R., JUST, M. A., & CARPENTER, P. A. (1982). A model of the time course and content of reading. *Cognitive Science*, **6**, 157–204.

WAUGH, N. C., & NORMAN, D. A. (1965). Primary memory. *Psychological Review*, **72**, 89–104.

WILLIAMS, R. J. (1986). The logic of activation functions. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. I). Cambridge, MA: Bradford Books.

WOODS, W. A. (1970). Transition network grammars for natural language analysis. *Communications of the ACM*, **13**, 591.

ZIPSER, D. (1986). Biologically plausible models of place recognition and goal location. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. II). Cambridge, MA: Bradford Books.

(Received August 17, 1987)
(Revision received November 18, 1987)