

## Spreading Activation Theories in Sentence Memory, Sentence Comprehension, and Speech Production\*

Takashi TSUZUKI

The concept of activation is the basis for recent empirical and theoretical studies in cognitive psychology, neuroscience, and artificial intelligence. Many theories and models in each of these areas use the construct of spreading activation as the main information-processing mechanism. This paper will discuss three topics of cognitive psychology which are related to spreading activation. These topics are sentence memory, sentence comprehension, and speech production—especially speech errors. The characteristics of different kinds of spreading activation theories will also be discussed.

Presented here is a review of spreading activation theories in the basic human memory system. According to Ratcliff, Hockley, & McKoon (1985), spreading activation theories in human memory can be grouped into two main classes; models which represent concepts as nodes in a network representation, and models which represent concepts in a distributed featural system.

In the network models for human long-term memory, activation is hypothesized to serve different functions. In a sentence verification task, activation is assumed to spread from two concepts in a knowledge structure. If the search by using

activation from two sources intersects, then information about the intersection becomes available for the decision process (Collins & Loftus, 1975; Quillian, 1967). To decide whether "a canary is a bird," activation spreads from the "bird" node and from the "canary" node and when an intersection is detected, a "true" response is produced.

The Collins and Quillian (1969) study is famous in early experiments of searching time. In this study it was shown that activation took time to spread from one node to the other node through a network representation.

The concept of activation is viewed in a different way by Anderson (1983) in his Adaptive Control of Thought (ACT\*) model. In the ACT\* model, activation is viewed as not needing significant time to spread as Ratcliff & McKoon (1981) found in their study. Instead, the level of activation at a node determines the rate at which decision processes can proceed. The levels of activation of the nodes in the network reflect their degree of association with the source nodes. When the source nodes change, spreading activation rapidly adjusts the levels of activation to achieve a new asymptotic pattern. It is proposed that the nodes corresponding to concepts form an interconnected network and that retrieval is performed by spreading activation throughout the network. The level of activation in the network determines the rate and the probability of correct recall.

In feature and neural models it is assumed that

---

\* I am grateful to Dr. Donald G. MacKay, my professor at the University of California, Los Angeles, and Gregory L. Kaufman MA. for their critical comments on this manuscript.

a mental state is a pattern of activation over a group of simple units. Recently these types of models have developed into the parallel distributed processing (PDP) models (McClelland & Rumelhart, 1986). These PDP models have in common the general assumption that input sets up a pattern of activation in the units of the system, and this pattern of activation leads to modification of the system. Once the system has been modified, partial input can be sufficient to interact with the modified system in order to reconstruct an activation pattern similar to the original input. McClelland & Rumelhart (1986) describe a distributed model of information processing. The PDP model consists of a large number of simple processing elements which send excitatory and inhibitory signals to each other through modifiable connections. Information processing is assumed to be the process where patterns of activation are formed over the memory units via their excitatory and inhibitory interactions. The internal representation of a processing event is the change in the interconnection's strength (weight) that results from the input encoding process. The representation of different events are superimposed on each other in the values of the weight matrix and activation patterns.

One typical view of the human memory system based on the concept of spreading activation through a network representation has been presented by Anderson (1985). Anderson (1985) describes the basic concepts and principles of human memory as follows. The meaning of a sentence can be encoded and represented as a hierarchical network of propositions. Short-term memory refers to the capacity for maintaining a limited amount of information in a specific state through activation. Activation refers to finite mental energy for processing information. Activation is assumed to increase the accessibility of information. Information can be used only when it is in an active state. The speed with which information can be processed in short-term memory is a func-

tion of its activation level. To be retrieved from long-term memory, information must be activated. Activation spreads along paths via a network of associations from the "currently active nodes" to the "to-be-retrieved nodes." The level of the spreading activation in a network representation (knowledge structure) determines how quickly that information can be examined and used. The level of activation which spreads to a knowledge structure depends directly on the strength of the connection through which the activation spreads and inversely on the number of paths which start from the same node. The inhibitory effect of competing connections on the amount of activation spread down a connection is referred to as associative interference. The strength of a knowledge structure increases with practice. But on the other hand, if the level of activation of a knowledge structure is low, due to either the low strength of each connection, or associative interference, errors of recall and recognition will occur. Though Anderson (1985) integrates many concepts in human memory based on spreading activation, and reports that his model is shown to predict important phenomena, his conceptualization about activation is too serial and restricted for understanding conscious processing when compared with the PDP model.

## **I. Sentence Memory as Viewed by both the ACT\* and the PDP Models**

Memory representation in the human cognitive system is one of the most important constructs, not only in cognitive psychology, but also in artificial intelligence. The discussion of sentence memory through these two cognitive models is appropriate in order to deal with the problem of memory representation and the spreading activation process. The following section consists of three parts. The first part compares the ACT\* model with the PDP model. It discusses both their similarities and differences. The second section

considers the application of each of these models in the human sentence memory process. And finally, in the concluding part of this section, the advantages and disadvantages of each of these two models are discussed.

### Similarities and Differences between the ACT\* Model and the PDP Model

There are both similarities and differences between these two types of models. One similarity is that both of these models have attempted to build a general theory of the human cognitive system. Both of them have tried to integrate the numerous experimental data which has been collected by psychologists, linguists, and neuroscientists and apply it to detailed programmable models for computer simulations.

A second similarity is related to spreading activation along memory representations, which is assumed to be able to change the accessibility of influenced memory units. Hinton, McClelland, & Rumelhart (1986) pointed out that it is hard to distinguish clearly between systems that use local representation and spreading activation together (e.g. the ACT\* model) and systems that use distributed representations (e.g. the PDP model). At the mathematical level, both of them follow very similar assumptions according to the spreading activation mechanism. The activation input to a unit consists of two parts: the input from outside the unit (external input), and the input from inside the unit (internal input). That is,

$$\begin{aligned} & \text{(the activation level of a given unit)} = \\ & \quad \text{(internal input)} + \text{(external input)} \end{aligned}$$

Where;

$$\begin{aligned} & \text{(external input)} = \\ & \quad \sum \{(\text{a matrix of connection strength}) * \\ & \quad \quad (\text{a vector of connected units})\} . \end{aligned}$$

Both models also assume that there is a decay factor which tends to pull the activation of whole units back to the resting level or "the asymptotic activation level" (Anderson, 1983).

The most obvious differences between these two models are in the form and level of the memory representations. In the ACT\* model, the basic form of the memory system is the cognitive unit which consists of a "unit node" and a set of specific "elements". Usually one element represents one concept or one word, and an abstract unit node integrates the lower levels of elements which interconnect with each other in a given network representation. On the other hand, the basic memory representation in the PDP model is called an "information-processing module" and it consists of a small set of several neuron like processing units. Information processing is assumed to depend on the patterns of activation which are formed over the units in the module. Therefore, in the PDP model, different kinds of information can be stored within the same module as different and distinct types of distributed activation patterns.

A second difference between these models is related to the learning process. In the ACT\* model, learning in the cognitive system involves the explicit formulation of rules and abstractions under the guidance of some explicit algorithm. The explicit formulation of rules occur in two ways. One way involves propositional knowledge using a semantic network representation; and the second way relates to procedural knowledge using the "if-then" type production system. In contrast to this, the PDP model uses a sophisticated mathematical expression called "the delta rule" which can extract regularities from a set of input data without the aid of complex rule-formulating mechanisms. The delta rule's name implies that the amount of learning is proportional to the difference between the actual activation achieved and the target activation provided as the goal state.

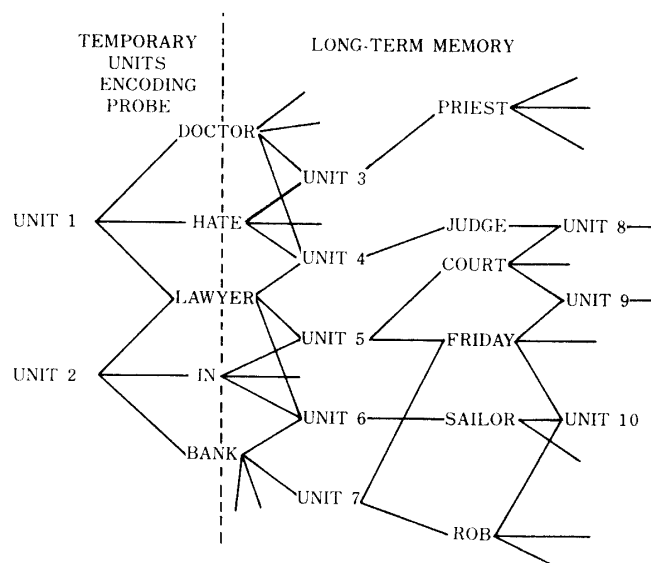
A third difference between these models is the information processing method which is closely related to the difference in memory representations. Although the information processing method of the ACT\* model is almost parallel in its activation system, its processing of the production system is

serial. Since the memory representation in the PDP model depends on the ensemble of simple distributed units, and this model uses the local learning rule between distributed representations, its information processing is completely parallel, even in the case of parsing complex sentences.

A final difference between these models is the manner in which constraints are used. One of the main properties of the PDP model is that its basic assumptions are appropriately restricted by the most recent findings in neuroscience. According to Gardner (1985), "as long as science remains relatively undeveloped, it may be necessary to carry out psychological experiments or computer simulations. But once the appropriate neuroscientific studies have been performed, a number of prior explanations should be rejected." In other words, neuropsychological data is especially helpful in guiding researchers to formulate the "theory of computation" (Kosslyn, 1987). Such a theory justifies positing a given computation through an analysis of the problems that must be solved, and the requirements for the solution to those problems. This type of theory specifies what must be computed and why. This computational theory is to be distinguished from the "theory of the algorithm," which specifies the particular steps actually used to achieve a computation. This algorithmic type of theory deals only with the details of how a computation is performed and it is not concerned with what the computation actually is. It might be possible to categorize the PDP model as a computational theory and the ACT\* model as an algorithmic theory.

### The Applicability of the ACT\* Model and the PDP Model to Sentence Memory

According to the ACT\* model, the memory trace of a given sentence is represented as a hierarchically integrated network structure which consists of a temporary active memory (working memory) network and a long-term memory (semantic mem-



**Figure 1.** A hypothetical network structure of a memory representation in the ACT\* model. Encoded in temporary active memory is the sentence: *The doctor hates the lawyer who is in the bank.* (From Anderson, 1983)

ory) network. Figure 1 is a hypothetical network structure of a sentence proposed by Anderson (1983). It shows an example of a piece of long-term memory in which unit nodes connect elements (words). In this example, temporary active memory has encoded a sentence, "The doctor hates the lawyer who is in the bank", and long-term memory involves the illustrated facts. For example, "The doctor hates the priest", "The bank was robbed on Friday", and so on.

Nodes corresponding to the main concepts are assumed to be sources of activation, and from these sources, activation spreads throughout the whole network. According to Anderson (1983), each link strength is calculated by the the word frequencies of connected elements. In the ACT\* model the integrated network structure, which consists of specific and abstract information, represents the meaning of an encoded sentence. Furthermore, spreading activation determines not only the structure of the encoded network but the strength of each node and each link in the interconnected network. In addition, the retrieval process is also performed through spreading activation along the network representation of the en-

coded sentence.

In the PDP model, the mechanisms of sentence processing are more complicated than that of the ACT\* model. The basic memory unit of the ACT\* model is an element which corresponds to one word or concept in a one-to-one relationship. In the ACT\* model, the purpose of spreading activation is mainly to make connections and to send information between semantically related words. In contrast to this, the basic unit of the PDP model is a set of distributed features. One active pattern of distributed features represents one word or one concept, but the correspondence between one memory representation of this model and one word is not on a one-to-one relationship. The same set of units are able to store different kinds of information using different activation patterns of basic features.

According to McClelland & Kawamoto (1986), when a sentence is presented, the PDP model determines the net input to each of the "sentence-structure units". Each sentence-structure unit represents the conjunction of two "microfeatures", which consist of a set of basic and abstract properties about words. Each sentence-structure unit has a modifiable connection to each of the "case-structure units" which represent the syntactical relationships between sentence-structure units. Based on the sentence-structure pattern and the current values of the connective weights of the relationships, an appropriate input to each case-structure unit is computed.

Figure 2 is an example of a sentence representation in the PDP model. The top line of this figure indicates the components (words) of the sentence, "The boy broke the window with the hammer." Be-

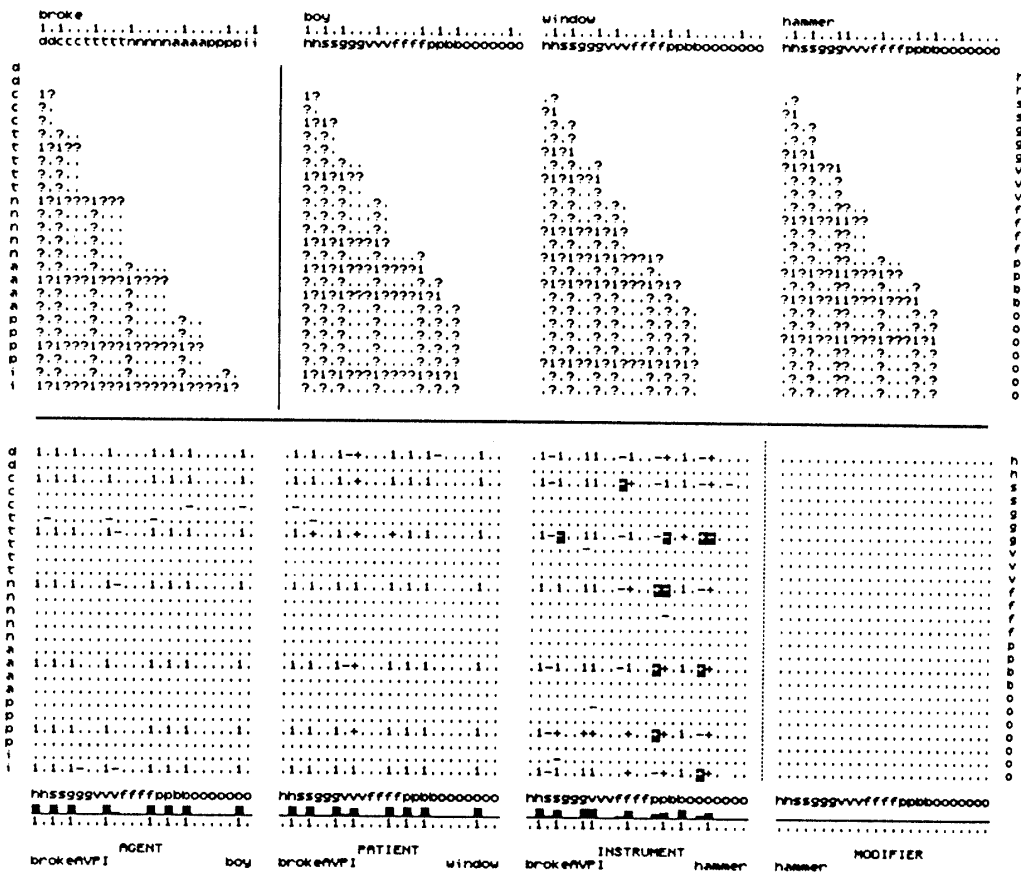


Figure 2. An example of a distributed memory representation in the PDP model. The sentence-structure units and the case-structure units are encoded *The boy broke the window with the hammer.* (From McClelland & Kawamoto, 1986)

low them the microfeatures of each component are shown, and below these are the conjunctive sentence-structure units for each component. Below the horizontal line, the activation pattern of the case-structure units for the syntactical roles of three nouns are shown.

As described above, McClelland & Kawamoto's application of the PDP model is directed mainly at the sentence comprehension process, but it is obvious that it can be applied to the sentence memory process without any major modification. In order to consider sentence comprehension and sentence memory using the PDP model, we have to construct a semantic data base by analyzing the detailed meaning of each word in the sentence. A similar semantic analysis of words is also required with the ACT\* model in order to construct semantic memory networks which define each word in sentences using other semantically related words. However the microfeatures of the PDP model are much more complicated and detailed than the semantic network of the PDP model. In studying sentence memory, including the efficiency of the encoding and retrieval processes, the efficiency in memory storage, and avoiding confusion between memory representations, many influential researchers have assumed that the word level of memory representations are more appropriate than semantically decomposed representations in order to construct simulation models (Anderson, 1983; Hayes-Roth & Hayes-Roth, 1977; Kintsch & van Dijk, 1978). But it may be possible that the mathematically well developed distributed representations proposed by the PDP model are indicative of a new perspective for viewing this important problem.

Many psychological researchers have shown that one's memory for sentence meaning is far superior to the retention of that sentence's wording (Brewer, 1975; Sachs, 1974). Although several recent experiments, using more sensitive measures than recall and recognition, (e.g. reaction time, confidence rating, and the saving rate of relearn-

ing,) have demonstrated that new distractor sentences, synonymous in meaning with previously studied sentences, could successfully be discriminated, even when the test was delayed (Anderson & Paulson, 1977; Hayes-Roth & Hayes-Roth, 1977; Kolers, 1976; Tsuzuki, 1987).

One of the essential properties of the human memory system seems to be the ability to form summarized representations that generalize from the details of specific experiences. However we also retain some specific details from the original information. Therefore, one major predicament for theorists of sentence memory is how to represent general and specific information. One solution to this dilemma may be to explain the explicit storage of both general and specific information. Anderson's ACT\* model may be a typical example of such a model. In this model, in order for generalization from specific to abstract information to occur, memory traces are stored in the production system, and mechanisms based on the production system induce summarized representations. The PDP model on the other hand, which relies on a sophisticated and local learning rule, can spontaneously account for empirical data that we store in abstract representations and is also able to enumerate specific experiences. Finally, as a solution to this dilemma, it is clear that the strategy of the PDP model is not only more reasonable, but also more flexible than that of the ACT\* model.

## Discussion

This new approach to understanding the human cognitive system is called "connectionism". According to McClelland (1985), its basic idea is that the computations carried out by a processing system are controlled by the connections among numerous simple processing units. The processing units themselves do very simple things. Usually they only update the level of activation they send to the other units based on a local learning rule according to signals they receive from other pro-

cessing units. The intelligence of the system, what it knows and what it can do with its knowledge, is determined by the weight matrix of the interconnections among the units. Ballard (1986) appraised connectionism from another point of view. Although the human neuron system's processing speed is extremely inferior to the modern computers' access time, every behavioral response can be realized in a few hundred milliseconds. Therefore the PDP system of neural units seems to be the only way to achieve these rapid response times.

In addition, there is one more important concept in studying the human cognitive system. This concept is called "modularity of mind" (Foder, 1983). It suggests that the human cognitive system might be constructed as a number of mostly separate information-processing devices, some constructed to deal with language, some constructed for visual processing, and so on. From this modular view, modules have evolved to perform particular forms of analysis in a hard-wired, rapid, and encapsulated manner. Gardner (1985) called the argument against the modular view a "horizontal" view, and he pointed out that Anderson supports this type of idea. Faculties like learning, memory, and perception are assumed to work in the same or similar fashion according to some general principles. Contrary to Anderson's theory, it is possible that an appropriate combination of connectionism, the modular view, and a suitable mathematical method is necessary to clarify the complex human cognitive system, including the sentence memory process. It seems reasonable to employ the bottom-up strategy in order to understand a complicated system like the human brain which is based on the constraints of the nervous system without assuming the existence of executive systems.

## II. Sentence Comprehension as Viewed by MacKay's Theory and the PDP Model

The purpose of this section is to compare Mac-

Kay's "node structure theory" (MacKay, 1987a, b) with the PDP model in two ways. These two ways are sentence comprehension in general; and sentence comprehension with specific phenomena, including the contextual resolution of ambiguity. This section is composed of four parts. The first part involves a general summary of MacKay's theory. The second part discusses the similarities and differences between these two theories. The third part contrasts them specifically in the case of the contextual resolution of ambiguity. And the fourth and concluding part discusses how MacKay's theory might be improved.

### General Properties of MacKay's Theory

According to MacKay (1987a, b), speech perception and production engage identical micro-processes which are performed by perception-production units called mental nodes. Mental nodes send top-down outputs to the muscle during production, and receive bottom-up inputs from sensory analysis during perception, including self-generated feedback (self-inhibition). Mental nodes become active during perception and production when we perceive language information (a word or

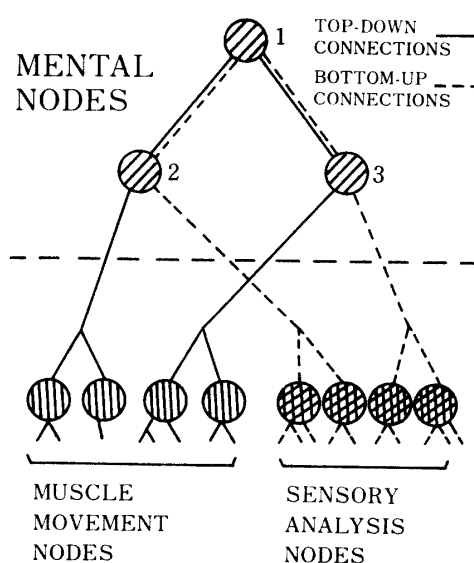


Figure 3. The relationship between mental nodes, sensory analysis nodes, and muscle movement nodes. (From MacKay, 1987a)

sentence) and produce it (cf. Figure 3).

Mental nodes have four processing properties: activation, priming, linkage strength, and self-inhibition. MacKay (1987b) assumes that *node activation* is all-or-none and self-sustained. Activation continues for a specific period, regardless of whether the sources of activation remain providing input (priming) or not. *Priming* means the transmission of a signal which increases sub-threshold activity from the active node to its connecting nodes. In order to become active, nodes must accumulate priming from the activating nodes which are connected to them. Although priming spreads through the connections between nodes in a parallel manner, priming is not self-sustained and begins to decay immediately after the input from the connected nodes ceases activation. *Linkage strength* is a long-term property of each connection determined by learning or practice. Linkage strength increases in proportion to the frequency of priming and activation through a particular connection. The given linkage strength determines how much and how quickly priming spreads at each connection. *Self-inhibition* is the inhibitory process which terminates the self-sustained activation of mental nodes. According to MacKay (1987a), self-inhibition is necessary in order to prevent the disruptive effects from bottom-up and top-down priming on mental nodes. MacKay (1987a, b) assumes mental nodes integrate perception and production in language processing, and that their dynamic characteristics such as activation, the spread of priming, linkage strength, and self-inhibition are quite simple. However these dynamic characteristics interact with one another in a complex manner and also depend on the current state of mental nodes and the results of prior experience.

### Comparison between MacKay's Theory and the PDP Model

The main similarity between these ideas is the

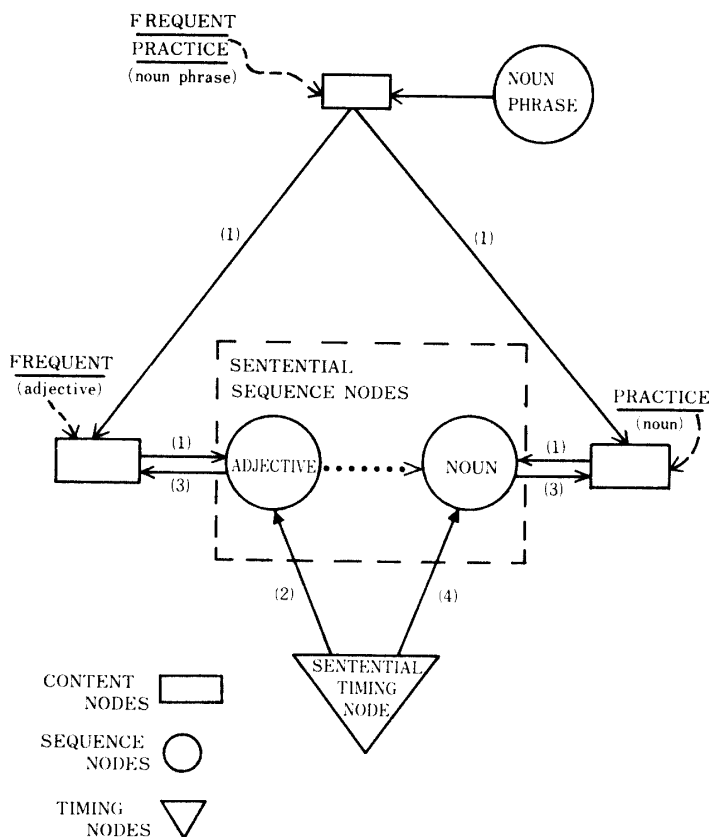
tendency to regard activation along internal representations as a critically important mechanism. However, the views about what activation actually is are completely different from each other. In MacKay's theory, activation refers to the all-or-none state of node accessibility and it does not spread through connections. Node activation and the spread of priming are distinguished clearly, and the accumulation of priming is necessary for node activation. An activating node increases the linkage strength of its connections and causes connected nodes to become primed. Linkage strength in turn influences how much and how rapidly priming can spread across the given connection. In addition, node activation is terminated by a period of reduced excitability (self-inhibition).

In contrast to this, the PDP model assumes that activation itself might spread via excitatory and inhibitory connections. The basic spreading activation mechanism of the PDP model follows: the activation input to a unit consists of external input and internal input. Instead of self-inhibition, it is a decay factor that pulls the activation level of each unit back to the resting level.

The first difference between these two theories is the level of the internal representation. In MacKay's theory, the basic form of a representation is the mental node, which consists of the content node, sequence node, and the timing node. Usually one content node represents one phrase or one word, and a superordinate node integrates the lower level nodes which are hierarchically interconnected to each other in a network structure. The basic representation of the PDP model is a small set of several simple units called the information processing module and its activation patterns.

Another difference relates to the learning rule. In MacKay's theory, learning in mental nodes involves the explicit formation of rules and production under the guidance of some explicit algorithms using sequence and timing nodes. Each linkage strength can store the results of learning





**Figure 4.** The order of top-down processes underlying activation of content nodes, sequence nodes, and the sentential timing node for producing the noun phrase *frequent practice*. (From MacKay, 1987a)

between each connection of the mental nodes in a bottom-up manner. Figure 4 illustrates the hypothetical content nodes, sequence nodes, and timing nodes. Unbroken lines in Figure 4 are excitatory, the broken lines represent the quenching mechanism, and the dotted lines represent the inhibitory relationship between sequence nodes. Similar connections and processes are postulated for all sequentially organized mental nodes. In contrast to this, the PDP model uses only a local learning rule (the delta rule).

A third difference between these two theories is in the method of information processing. Although the lower-level and bottom-up information processing in MacKay's theory is parallel in its priming mechanism, its higher-level and top-down mechanisms use serial processing mechanisms. In-

formation processing in the PDP model is completely parallel as already mentioned above.

### The Resolution of Contextual Ambiguity

A series of experiments using the lexical decision task was conducted by Swinney (1979) to investigate how ambiguous homonymous words are disambiguated. He had subjects listen to sentences like the following: "...The man was not suprised when he found several spiders, roaches, and other bugs in the corner of the room." Swinney was concerned with the homonymous word "bugs". Just after hearing the word, subjects were presented with a string of letters on a screen, and their task was to judge whether that string of letters made a word or not. For example, if they saw "s-e-w" they

## SPREADING ACTIVATION THEORIES

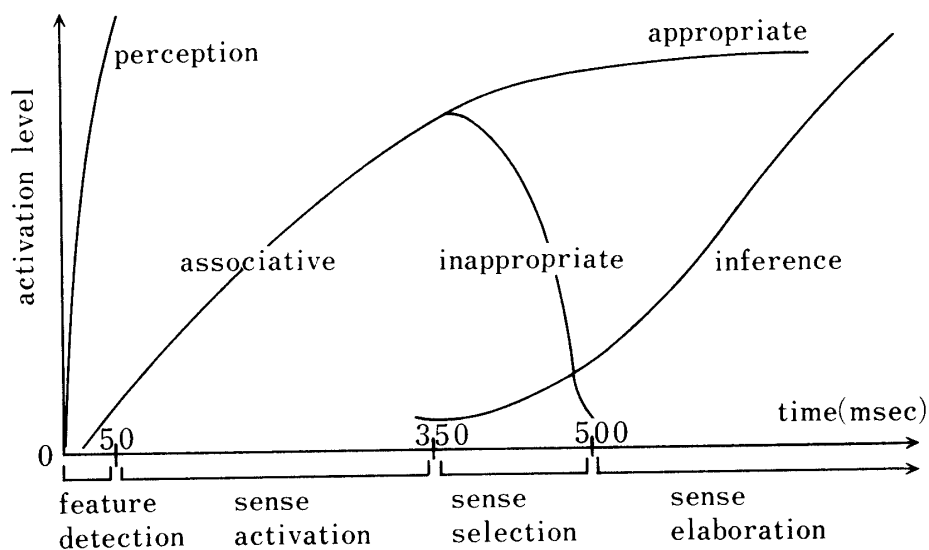
should say “yes”; but if they saw “s-i-w”, they should say “no”, as it is not a word.

The critical comparison involved having subjects judge words like “spy”, “ant”, or “sew”, following “bugs”. In the given context, the word “ant” is related to the primed meaning of “bugs”, while the word “spy” is not related to the context. The word “sew” is a neutral control condition. When the to-be-judged word was presented within 400 msec following the primed word “bugs”, Swinney (1979) found that recognition of both “spy” and “ant” was faster. Thus, the presentation of the homonymous word immediately accesses both of its meanings. When the to-be-judged word was delayed over 700 msec, Swinney (1979) found that there was facilitation for the contextually appropriate word “ant” only. It appears that a contextually correct meaning is selected at about this time and the other is inhibited. Thus, two meanings of a homonymous word are momentarily accessed, but a given context operates rapidly to select the appropriate meaning. Recently, similar results have been reported by other researchers (Davidson, 1986; Kintsch & Mross, 1985; Marcel, 1980; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Tsuzuki, 1988a).

Kintsch (1987) and Kintsch & Mross (1985)

proposed the following assumptions in order to explain these results (cf. Figure 5). They assumed that words were represented lexically by one or more word meanings in the mental lexicon. The first subprocess of word identification consists of the activation of plural word senses (“sense activation”). Next, the contextually appropriate meaning is selected from the activated word senses. Usually, this “sense selection” process provides no more than a possible candidate to be verified through contextual elaboration or inference. The network of the mental lexicon itself provides no help when this process selects one of the activated items. In short, Kintsch & Mross (1985) describe the sequence of processing as being: the sense activation stage, the sense selection stage, and finally the elaboration stage.

If the mental lexicon is supposed to be a kind of connected network with verbatim items at the nodes, and with links from each item to several other nodes, we can assume that the access of an item in the lexicon activates the corresponding node and that the energy of activating a node spreads along the links which lead from it. Furthermore, it can be assumed that if activation spreads through the lexical network, response thresholds for activated nodes are temporarily



**Figure 5.** The activation process of a homonymous word's sense. (Adapted from Kintsch, 1987)

lowered. As a result, accessing time for a lexical item, which is related to directly activated items, becomes shorter (Collins & Loftus, 1975; Morton, 1969). In the previously described example, we can assume the existence of connections between the nodes assigned to the target word ("bugs") and each of the semantically related words ("spy" and "ant"). It can be assumed that facilitation of "ant" in the "bugs" context can be effected by the activation of such internal connections.

According to MacKay (1987a), ambiguity can be said to occur when different content nodes in the same domain simultaneously receive comparable levels of priming from bottom-up connections. MacKay's explanation of Swinney's results is similar to the interpretation in studies which have been mentioned previously (Kintsch, 1987; Kintsch & Mross, 1985), except for his strict distinction between node activation and spreading priming. MacKay's theory requires no special-purpose mechanism between content nodes for accomplishing the either-or resolution of ambiguity. The "most-primed-wins" principle automatically resolves ambiguity in an either-or manner. With the most-primed-wins principle, only the most primed node becomes activated in a domain, including the domain of sentential sequence nodes. Thus, according to MacKay, the time course of disambiguation in sentence comprehension is explained as nonunique lexical priming among connected content nodes by given linkage strength, and automatic resolution of ambiguity (activation of a correct content node) under the most-primed-wins principle and the control of sentential sequence nodes.

By assuming complex information-processing mechanisms McClelland & Kawamoto (1986) reported that their PDP type simulation model works very well with ambiguous words. That is, it has little difficulty determining which reading to assign to an ambiguous word based on its context of occurrence as long as the sentential context was itself sufficient to disambiguate the meaning

of the word. But as described above, in order to consider sentence comprehension and the resolution of contextual ambiguity using the PDP model, it is necessary to construct a semantic data base of microfeatures by decomposing the meaning of each word. There may be an important point relating to the representative level of basic units; to perform a study about one particular aspect of human cognition like contextual ambiguity, the most appropriate level of basic representations for the given problem must be determined.

### How MacKay's Theory might be Improved ?

As MacKay (1987a) summarizes, one of the distinctive characteristics of his theory is the distinction between content nodes, sequence nodes, and timing nodes. The general processing properties of these nodes (e.g. priming and activation), their basic characteristics (e.g. hierarchical organization with respect to activation), and their long-term memory properties (e.g. linkage strength) are not especially new. However, the exact nature of the interactions among these processing mechanisms provides the basis for new theoretical generalizations. Like other current theories, MacKay's node structure theory follows the current trend toward a focus on dynamic, on-line, or real-time processing issues, in addition to static or structural issues. Although the integration of perception and production is important in MacKay's theory, the following discussion will be restricted to the perception and understanding of sentences in order to explain the critical issues.

MacKay's theory has been shown to possess not only advantages but disadvantages when compared with the PDP model. One of the advantages of MacKay's theory is the simplicity of the basic mechanism including the clear distinction between node activation and spreading priming and the assumption of self-inhibition instead of the decay parameter or inhibitory connections. However, it is very difficult to understand how the informa-

tion-processing mechanism actually works in a specific situation, because MacKay (1987a, b) does not use a mathematical method to describe his theory. The mathematical properties of priming in MacKay's theory may be similar to those of the ACT\* model's spreading activation mechanisms (Anderson, 1983) even though their conceptualizations are quite different from each other. Furthermore, the concept of timing nodes and sequence nodes are not clear from a computational point of view. MacKay (1987a) argues this problem as follows: "The theory's interface with neurophysiology remains largely unexplored. I currently have no solid answer to questions such as 'how are timing, sequence and content nodes instantiated in the brain?' The computational adequacy of the theory from a linguistic point of view also remains to be explored." Sequence nodes and timing nodes may correspond to some top-down mechanisms such as the production system (Anderson, 1983) or the Augmented Transition Network (ATN) grammar (Woods, 1970) which are able to control the bottom-up processing based on local rules. Thus, it might be necessary to consider the more detailed rules of top-down mechanisms which are related to sequence nodes and timing nodes. In order to construct a detailed interactional model from a descriptive general theory, like MacKay's, computer simulations have proven to be one of the best methods (e.g. Anderson, 1983; McClelland & Kawamoto, 1986; Tsuzuki, 1985, 1988b; Waltz & Pollack, 1985).

Although the PDP model is too parallel in information processing and too local with the learning rule, it uses reasonable assumptions for a cognitive theory in general and is considerably developed in its mathematical formulation. MacKay's theory, which is a mixture of parallel bottom-up processes and sequential top-down processes, should be developed as a programmable and computational model by using mathematical methods or algorithmic formulations. As a first step, the interaction between the spread of priming and the

transitional state of nodes should be formulated by using linear mathematics. Also the properties of sequential and timing nodes should be conceptualized as well-defined rules (i.e. the ATN grammar or the production system). In summary, the most important problem for both MacKay's theory and the PDP model is the introduction of an appropriate top-down mechanism by integrating it with well-defined bottom-up rules without losing theoretical economy and simplicity.

### III. Spreading Activation Theories of Speech Production and Speech Errors

In sections I and II of this paper, the general characteristics of spreading activation theories for sentence memory and sentence comprehension have been discussed. Since theories of speech production, which use the concept of spreading activation, can be regarded as one kind of application of these theories, this section will focus on these theories' explanations of speech errors, especially the work of MacKay (1987a, b) and Dell (1986).

The process of self-activation is the central concept in MacKay's node structure theory. According to MacKay (1987b), self-inhibition is the inhibitory process which terminates the self-activation of mental nodes and temporarily re-

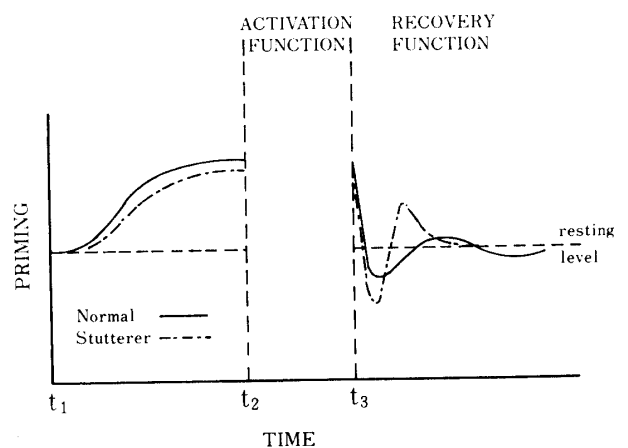


Figure 6. The priming, activation, and recovery functions for stutterers and nonstutterers. (From MacKay, 1987a)

duces their priming level to below the normal or resting state. Self-inhibition is a "built-in characteristic" of all nodes, but only nodes which have already been activated, rather than just primed, during the course of perception and production exhibit self-inhibition. Self-inhibition is the first stage in the cycle of recovery from activation. As summarized in Figure 6, this cycle includes four phases of excitement and inhibition. These stages follow the accumulation of priming; activation, self-inhibition, hyperexcitability, and a return to the resting level of priming.

In the node structure theory, self-inhibition is needed to prevent the destructive effects from internal feedback (bottom-up priming) on mental nodes. Since mental nodes receive both bottom-up and top-down connections, internal feedback can potentially cause reverberatory activation of mental nodes. During production, a superordinate node becomes activated and it primes a subordinate node via top-down connections. However, soon after the subordinate node becomes activated, priming of the superordinate node through the bottom-up connections works for perception using exactly the same nodes. Thus, self-inhibition is needed at every level to make sure that bottom-up priming from subordinate nodes does not lead to the reactivation of just activated nodes.

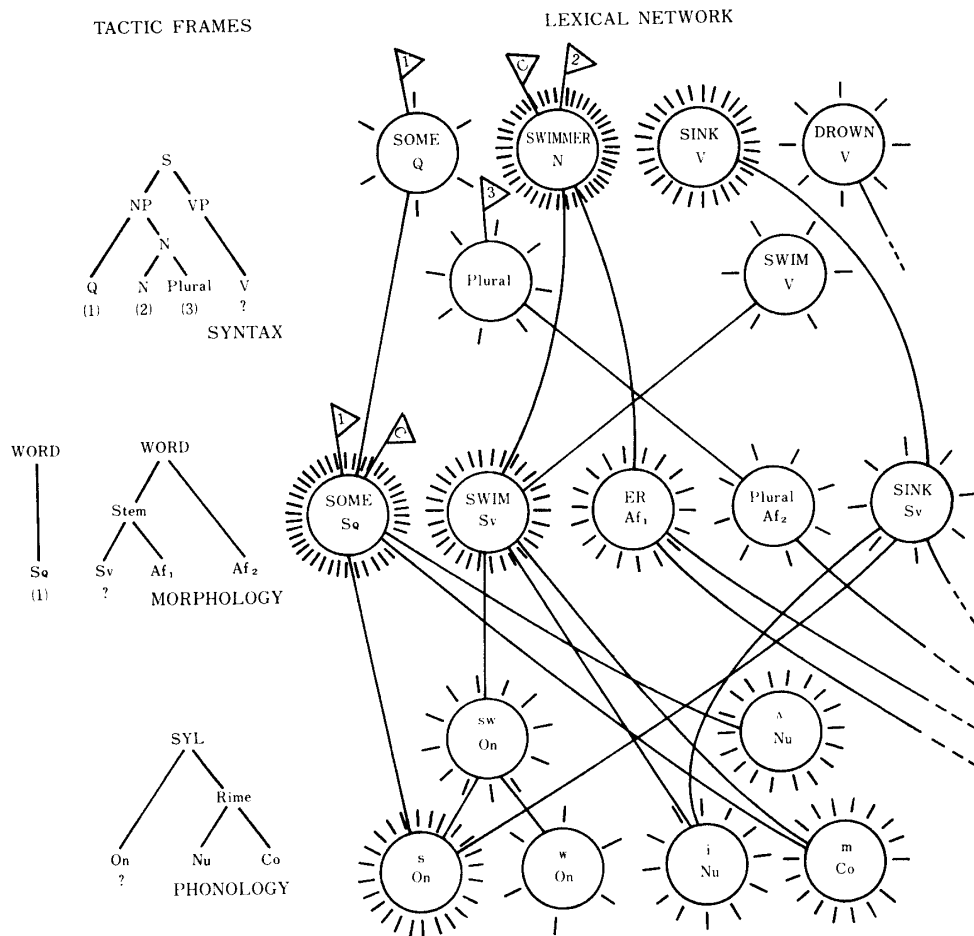
MacKay (1987b) uses feedback-induced stuttering as an example of the destructive effects of internal feedback. In MacKay's theory, the muscle movement units in stuttering display an abnormal recovery cycle as illustrated in Figure 6. This means that the rebound from self-inhibition comes earlier than normal, and rises to a higher level of priming. As a result, just activated nodes in their hyperexcitability stage are most primed at the time when the next node is to be activated. Heightened hyperexcitability alone may cause repetition errors, but amplifying the auditory feedback increases the possibility of stuttering.

Dell's (1986) work with speech errors focuses on sentence production. This involves the kinds of

errors that occur, the constraints on their form, and the conditions that precipitate them. Dell (1986) combines a spreading activation retrieval mechanism through a network representation with assumptions regarding linguistic units and rules (the frame-and-slot view of production). The processing assumptions in Dell's theory can be seen in Figure 7. At each level (syntactic, morphological, and the phonological level) a representation of the to-be-spoken sentence is constructed, and each representation is an ordered set of items found in the mental lexicon (long-term memory). The basic premise of this theory is that the tagged nodes, constituting higher representation, activate nodes that may be used for the immediately lower representation through a spreading activation mechanism. In Dell's theory, the rules of spreading activation are almost the same as Anderson's formulation for memory retrieval (Anderson, 1983). While spreading activation is working, the rules associated with the lower level nodes construct a frame of categorized slots. After a categorically defined slot is created in the frame, the insertion rules operate to fill that slot. The insertion rules refer to the activation levels of all nodes that are marked with the specified category and select the item whose node possesses the highest level of activation. The selection of an item is followed immediately by the reduction of its activation level to zero (to prevent it from being selected repeatedly) and the tagging of the node. Although this reduction of the node's activation level is similar to MacKay's idea of self-inhibition, MacKay's node structure theory is more developed than Dell's theory (cf. Figure 6).

In Dell's theory, in order to understand the speech-error phenomena, how the order information is presented must first be understood. A good deal of this ordering is accomplished by the rules that build the frame and the ordered set of categorically specified slots. According to Dell (1986), both an item's category and its activation level interact to determine its serial position in a repre-

# SPREADING ACTIVATION THEORIES



**Figure 7.** A moment in the production of the sentence *Some swimmers sink*. (Tactic frames [left] specify an ordered set of categorically labeled slots. The numbered slots have already been filled in; a numbered flag (an order tag) has been placed on each node in the lexical network that stands for an item filling a slot. The question mark indicates the slot in each frame that is presently being filled. The highlighting surrounding the nodes reflects activation level and the flag with a C on it marks the current node of each level. Each node is labeled by membership in some category: Syntactic categories for words are (Q)uantifier, (N)oun, (V)erb, plural marker, morphological categories for morphemes are (S)tem, (Af)fix; and phonological categories for sounds are (On)set, (Nu)cleus, (Co)da.) (From Dell, 1986)

sentation. From this kind of ordering mechanism, errors can be expected to follow categorical constraints. An error occurs when a wrong item is more activated than the correct one and is selected and tagged. However, for this wrong item to be selected, it must be a member of the same category as the correct item. Dell's theory assumes that the insertion rules, whose job is to pick out

the most highly activated node of a specified category, stick to their categories. The consequence is a strict categorical constraint on errors. For example, in morpheme errors, stems can replace other stems, and affixes can replace other affixes. For sound errors which occur at the phonological level, initial consonants replace other initial consonants, vowels replace other vowels, and

so on. Dell's theory explains speech errors statically within the categorical constraint using the frame-and-slot mechanism, but MacKay's theory is able to describe them from a more dynamic point of view, one which uses the recovery cycle hypothesis to describe what occurs between perception and production.

#### IV. Comparison among Spreading Activation Theories: Conclusion

Three differences in spreading activation theories will briefly be discussed here: the level of representations, basic assumptions about the activation mechanism, and control rules. First, as was argued in the first part of this paper, according to Ratcliff et. al. (1985), it is possible to consider two different types of conceptualizations for representations in spreading activation theories. The two types are: first, a hierarchical network; and second, distributed features. The theories of Anderson (1983), Dell (1986), and MacKay (1987a, b) belong to a type of hierarchical network representation which assumes that the knowledge structure between currently activated parts of the mental lexicon and superordinate nodes is integrated. In this kind of theory, the basic level of representation involves one word or concept which corresponds in a one-to-one manner with a real world item. In contrast to this, the PDP model assumes that the activation pattern of distributed features can represent different kinds of information. However, there is also a mixed type of representation between both classes (e.g. Waltz & Pollack, 1985).

The basic formulation of spreading activation is quite similar among theories which assume continuous activation levels, like the theories of Anderson (1983), McClelland & Rumelhart (1986), and Waltz & Pollack (1985). In each of these theories, the activation input to a unit consists of two parts, the external input and the internal input (self-activation).

The inhibitory mechanism is different in each theory. In Anderson's theory and the PDP model, there is a decay parameter which tends to pull the activation of the whole unit back to the resting level. In the PDP model and in Waltz & Pollack's model, it is assumed that there are both inhibitory connections and excitatory connections. There is also a view which assumes that there is a special inhibitory mechanism similar to self-inhibition in MacKay's theory, which might be the developed form of a simple decay factor in self-activation. Furthermore, some theories (e.g. Dell, 1986; MacKay, 1987a, b) distinguish the binary activation state of a node from the continuous transmission of a signal which increases the node's accessibility.

Each theory also assumes specific control rules of spreading activation in order to achieve each purpose. It may be possible to divide them into bottom-up rules and top-down rules. The bottom-up rule refers to a local rule or "built-in" mechanism of a basic unit, like the delta rule in the PDP model or self-inhibition in MacKay's theory. It also includes the basic spreading activation mechanism itself. The top-down rule includes the generation rule in Dell's theory, the production system in Anderson's work, and the sequence and timing node in MacKay's theory.

In summary, in order to construct a detailed interactional model from a descriptive theory, the computer simulation seems to be the best method. Except for MacKay's theory, all of the other theories which have been referred to in this paper, were translated into computer simulation models and the results were examined by comparing them with experimental data. As Dell (1986) argues, the theory suffers to some extent from a trade-off between explicitness and a desire to avoid unmotivated assumptions. In their account of psychological phenomena, some theories are explicit but contain several ad hoc assumptions. Where these assumptions are less in evidence, in the description of higher level processes, the

theories become much fuzzier. This is just another way of saying that there is a need for much more experimental research in the area of sentence memory, sentence comprehension, and speech production.

## REFERENCES

- Anderson, J. R. 1983 *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. 1985 *Cognitive psychology and its implications*. (2nd. Ed.) San Francisco: Freeman.
- Anderson, J. R., & Paulson, R. 1977 Representation and retention of verbatim information. *Journal of Verbal Learning and Verbal Behavior*, 16, 439-451.
- Ballard, D. H. 1986 Cortical connections and parallel processing: Structure and function. *The Behavioral and Brain Sciences*, 9, 67-120.
- Brewer, W. F. 1975 Memory for ideas: Synonym substitution. *Memory & Cognition*, 3, 458-464.
- Collins, A. M., & Loftus, E. F. 1975 A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407-428.
- Collins, A. M., & Quillian, M. R. 1969 Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 240-247.
- Davidson, B. J. 1986 Activation of semantic and phonological codes during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 201-207.
- Dell, G. S. 1986 A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283-321.
- Fodor, J. A. 1983 *Modularity of mind*. Cambridge, MA: MIT Press.
- Gardner, H. 1985 *The mind's new science*. Basic Books: New York.
- Hayes-Roth, B., & Hayes-Roth, F. 1977 The prominence of lexical information in memory representations of meaning. *Journal of Verbal Learning and Verbal Behavior*, 16, 119-136.
- Hinton, G. E., McClelland, J. L., & Rumelhart, D. E. 1986 Distributed processing. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.) *Parallel distributed processing: Explorations in the microstructure of cognition*. Vol 1. Cambridge: MIT Press.
- Kintsch, W. 1987 Understanding word problems: Linguistic factors in problem solving. In M. Nagao (Ed.), *Language and artificial intelligence: Proceedings of an international symposium on language and artificial intelligence held in Kyoto, Japan, 1986*. Amsterdam: North-Holland.
- Kintsch, W., & van Dijk, T. A. 1978 Toward a model of text comprehension and reproduction. *Psychological Review*, 85, 363-394.
- Kintsch, W., & Mross, E. F. 1985 Context effects in word identification. *Journal of Memory and Language*, 24, 336-349.
- Kolers, P. A. 1976 Reading a year later. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 554-565
- Kosslyn, S. M. 1987 Seeing and imaging the cerebral hemispheres: A computational approach. *Psychological Review*, 94, 148-175.
- MacKay, D. G. 1987a *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer.
- MacKay, D. G. 1987b Theoretical and empirical relations between perception, action, and cognition. In W. Prinz & O. Neumann (Eds.) *Functional relations between perception and action*. New York: Springer.
- Marcel, A. J. 1980 Conscious and preconscious recognition of polysemous words: Locating the selective effects of prior verbal context. In R. S. Nickerson (Ed.), *Attention and performance VIII*. Hillsdale, NJ: Erlbaum.
- McClelland, J. L. 1985 Putting knowledge in its place: A scheme for programming parallel



- processing structures on the fly. *Cognitive Science*, 9, 113-146.
- McClelland, J. L. & Kawamoto, A. H. 1986 Mechanisms of sentence processing: Assigning roles to constituents. In J. L., McClelland, D. E., Rumelhart, & the PDP Research Group (Eds.) *Parallel distributed processing: Explorations in the microstructures of cognition. Vol 2.* Cambridge: MIT Press.
- 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. 1986 A distributed model of human learning and memory. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.) *Parallel distributed processing: Explorations in the microstructures of cognition. Vol 2.* Cambridge: MIT Press.
- Morton, J. 1969 Interaction of information in word recognition. *Psychological Review*, 76, 165-178.
- Quillian, R. 1967 Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science*, 12, 410-430.
- Ratcliff, R., Hockley, W., & McKoon, G. 1985 Components of activation: Repetition and priming effects in lexical decision and recognition. *Journal of Experimental Psychology: General*, 114, 435-450.
- Sachs, J. S. 1974 Memory in reading and listening to discourse. *Memory & Cognition*, 2, 95-100.
- Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. 1982 Automatic access of meaning of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, 14, 489-537.
- Swinney, D. A. 1979 Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 645-660.
- Tsuzuki, T. 1985 A computer simulation for the network activation theory of sentence memory: A case for S-O-V sentences. *Bulletin of the Faculty of Education, the Department of Educational Psychology, Nagoya University*, 32, 21-47. (In Japanese with English summary)
- Tsuzuki, T. 1987 Effects of verbatim information on recognition of S-O-V sentences under incidental learning. *The Japanese Journal of psychology*, 58, 151-157. (In Japanese with English summary)
- Tsuzuki, T. 1988a Effects of context and processing time in the mental conversion of homophones from *Kana* to *Kanji*. Paper presented at the 52th Annual Convention of the Japanese Psychological Association, 615. (In Japanese)
- Tsuzuki, T. 1988b An interactive activation model of context effects in the recognition of phonophones. Paper presented at the 5th Annual Convention of the Japanese Cognitive Science Society, 102-103. (In Japanese)
- Waltz, D. L., & Pollack, J. B. 1985 Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science*, 9, 51-74.
- Woods, W. A. 1970 Transitional network grammars for natural language analysis. *Communications of the ACM*, 12, 591-606.

(1988年7月31日 受稿)

## 要 約

## 文記憶，文理解，発話産出における活性化拡大理論

都 築 誉 史

本論文では、文記憶，文理解，発話産出に関する活性化拡大理論についてレビューした。

活性化拡大理論は、内的表象に焦点を当てた場合、ネットワークモデルと特徴モデルとに2分することができる。特に後者は近年、並列分散処理（PDP）モデル（McClelland & Rumelhart, 1986）として発展してきている。

まず、文記憶について、ACT\*モデル（Anderson, 1983）とPDPモデルとを比較すると、連続的な活性化の拡大メカニズムに関しては、両者に共通点を見出すことができる。しかしながら、抑制メカニズムについては、ACT\*モデルは減衰パラメータのみによって制御されているのに対し、PDPモデルではそれに加えて、促進的結合と抑制的結合とが仮定されている。

次に、文理解について、多様な機能を有するメンタル・ノードのネットワークを仮定するノード構造理論（MacKay, 1987a）とPDPモデル（McClelland & Kawamoto, 1986）とを比較した。特に、文脈上の多義性を解消するプロセスに関して、Kintsch（1987）のモデルを介して両者を比較した。

最後に、スピーチ・エラーの観点から、発話産出に関する活性化拡大理論について検討した。ノード構造理論

によれば、スピーチ・エラーは、プライミングにおける活性化・自己抑制サイクルのゆらぎとして把握される。一方、Dell（1986）のモデルによれば、スピーチ・エラーは、階層的ネットワーク表象において、活性化の拡大によって起動されるフレーム・スロット規則の適用の失敗としてとらえられる。

上記のように、活性化拡大理論は様々な観点からモデル化がなされているが、内的表象の性質、活性化・抑制メカニズム、制御規則の3点から考察を加えることができる。また、活性化・抑制メカニズムと制御規則は、ボトムアップ規則とトップダウン規則に各々対応している。

全体のまとめとして、文の記憶，理解，産出といった問題を扱う場合に、どのレベルの内的表象が適切であるかについて議論し、さらに、PDPモデルのように、基本的にはボトムアップ規則のみを用いるアプローチと、他のモデルのようにトップダウン規則をも併用するアプローチとを比較検討した。そして、内的表象、活性化・抑制メカニズム、制御規則間の複雑な相互作用をモデル化する上で、コンピュータ・シミュレーションの意義について言及した。