Title
Context-dependent Recognition in a Self-organizing Recurrent Network

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 19(0)

Authors
Myung, In J.
Kim, Cheongtag
Levy, William B.

Publication Date
1997

Peer reviewed

Context-dependent Recognition in a Self-organizing Recurrent Network

In J. Myung and Cheongtag Kim
Department of Psychology
Ohio State University
Columbus, OH 43210-1222
{myung.1, kim.315}@osu.edu

William B. Levy
Department of Neurological Surgery
University of Virginia Health Sciences Center
Charlottesville, VA 22908-0420
wbl@virginia.edu

Abstract

Cognition of an object depends not only upon the sensory information of the object but also upon the context in which it occurs, as demonstrated in many psychology experiments. Although there has been considerable amount of research in cognitive science that demonstrates the importance of context, seldom has this research concerned specific computational mechanisms for learning and encoding of context. As context is largely an integration of the past up to the present, some form of information about the past stimuli must be abstracted and stored for a certain period of time so as to be used in the interpretation of the present stimulus. In this modelling approach we explore such mechanisms. In particular, we describe an unsupervised, sparsely connected, recurrent network that creates its own codings of input stimuli on ensembles of network units. Moreover, it also self-organizes itself into a short-term memory system that stores such codings. Simulations demonstrate the context-dependent recognition performance of the network.

Introduction

The study of context-dependent recognition has long been a focus of cognitive psychology, for example, in visual and auditory perception (e.g., Labov, 1973), speech perception (Maslen-Wilson, 1975), word recognition and recall (Sweeney, 1979), and sentence processing (Tyler & Maslen-Wilson, 1977). The top panel of Figure 1 shows examples of context-dependent recognition. Here the surrounding letters (numbers) determine the interpretation of the middle letter (number) (Hunt & Ellis, 1974). The bottom panel illustrates yet another context effect, which we might call a "mental hysterisis" effect. In this figure ambiguous pictures in the middle of the series are perceived differently, as a man's face or as a woman, depending upon whether the series is viewed from left to right, or in the opposite order, respectively (Fisher, 1967). These examples demonstrate that the same sensory stimulus gives rise to multiple, internally evoked interpretations of that stimulus, depending upon the context in which it occurs. In contrast to such numerous demonstrations of the importance of context, specific computational mechanisms for learning and encoding of context has yet to be well understood. As context is largely an integration of the past up to the present, a minimal condition for a system to exhibit context-dependent recognition is that some form of information about the past stimuli be abstracted and preserved for a period of time so as to be used in the interpretation of the present stimulus. In this modelling approach we explore such mechanisms.

Hebbian Learning and Assembly Coding

According to the Hebb rule (Hebb, 1949), two neurons that tend to fire repeatedly in close temporal proximity are likely to develop an excitatory synaptic connection between them so that when one of the pair gets excited later by an input, the other will also get excited through the connection. A congregation of such interconnected, mutually excitatory neurons, all induced by the same input, is called a cell assembly which itself is a distributed internal representation of the input. We might then view the activation pattern of the cell assembly as the *neural code* for the input stimulus. On the other hand, in psychology of perception, what an observer perceives internally for an externally presented input is called a percept. To the observer the percept is his/her mental interpretation or meaning of the input. How

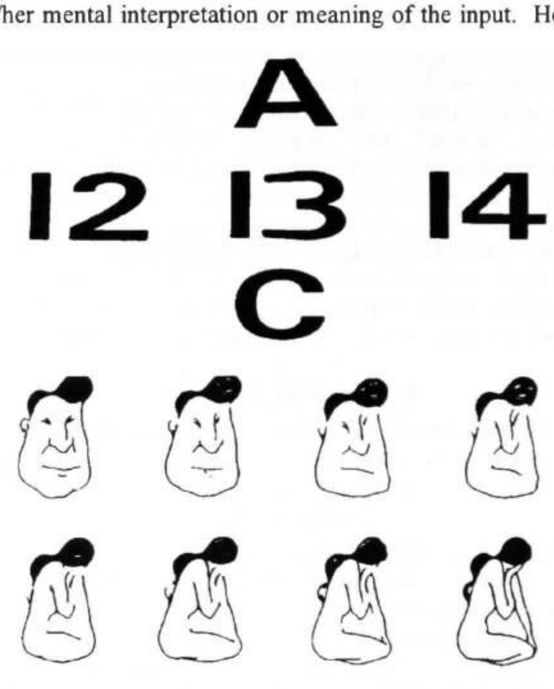

Figure 1. Examples of context-dependent recognition (After Hunt & Ellis, 1975; Fisher, 1967).

might percepts then be represented in the cognitive system? We make the standard and relevant assumption, as other researchers (Anderson & Murphy, 1986; Elman, 1990), that for a given sensory input, the percept is an activation pattern that has arisen in the cognitive system as a result of the input, that is, the neural code of that input.

In this paper we describe a neural network that finds its own neural codes for input stimuli without external teaching signals or internal error-correcting signals but rather, self-organizes itself into a short-term memory system that maintains internal representations of inputs in the network for a short-period of time. In other words, our network seems to meet the forementioned condition necessary for context-dependent recognition performance.[1] The following simulations bear out this hypothesis, showing that the network indeed exhibits such performance, in particular, context-dependent recognition of ambiguous objects and a mental hysterisis effect.

Simulations

The Network

The network, similar to the one in Levy, Wu and Baxter (1995), is a McCulloch-Pitts based recurrent network with asymmetric, sparse random connections. Network connectivity, which is defined as the probability of any two neurons being connected, is 0.05. A portion of the network units called "external" units is used to encode input stimuli. The rest called "internal" units receive no inputs directly but instead, they are activated through their connections to the external units as well as recurrent connections among themselves. The activation pattern over the internal units that has arisen as a response to an input is interpreted as the internal representation or percept of the input stimulus. The external units are also recurrently activated, but they are not counted during decoding. The firing (z_j = 0 or 1) of any given unit at time t is determined by

$$z_j(t) = 1 \text{ if either } \frac{\sum w_{ij}(t-1) z_i(t-1)}{\sum w_{ij}(t-1) z_i(t-1) + K_I \sum z_i(t-1) + K_R \sum_i x_i(t)} \geq \theta$$

$$\text{or} \quad x_j(t) = 1$$

and $z_j(t) = 0$ otherwise. In the equation w_{ij} ($0 \leq w_{ij} \leq 1$) is the connection weight between units i and j, θ ($0 < \theta < 1$) is a threshold, K_I (=0.007) and K_R (=0.0032) are delayed- and feedforward-inhibition constants, respectively, and x_i (= 0 or 1) is the activation of input unit i. The connection weight is modified according to a Hebbian-type, associative rule (Levy, 1982)

$$w_{ij}(t+1) = (1-\varepsilon) w_{ij}(t) + \varepsilon \, z_j(t) \, (z_i(t-1) - w_{ij}(t))$$

where ε (=0.02) is a learning rate.

Context-dependent Recognition of Ambiguous Objects

Methods. The network consisted of 90 external units and 710 internal units, and two categories of artificial stimuli were used as training inputs. A total of nine input vectors of size 90 were generated as follows. Each input vector was made up with 10 ON units (x_i=1) and 80 OFF units (x_i=0). No overlap was allowed between any two of the nine "orthogonal" vectors. Four of these inputs were designated as category A stimuli indicating A1, ..., A4, and another four as category B stimuli indicating B1, ..., B4. The remaining input denoted as W was the ambiguous stimulus that was included in both categories. In other words, category A consists of five stimuli, {A1,..., A4, W (≡ A5)} and category B consists of five stimuli, {B1,..., B4, W (≡ B5)}. Note that because of the non-overlapping coding scheme, all nine stimuli were *equally* dissimilar or similar to one another and consequently, the present distinction between the two categories was completely arbitrary. The network was trained in blocks of 50 stimuli that belonged to the same category, either A or B, and that were randomly selected with equal probability (0.2) from the five category stimuli. The network received 100 blocks of training, of which a half was category A blocks and the other half was category B blocks. The two types of blocks were randomly mixed. The idea behind this training procedure was that during training, stimuli from the same category would be associated with one another so "learned" similarities between them would be developed and represented in the internal units. The neural code for a given input stimulus was defined as the pattern of activations over the 710 internal units arising in response to the input during training. As activation patterns varied across repeated presentations of the same input, the neural code was computed as an average activation across the last five presentations of that input. At the end of the 100th block, the connection weights were fixed and network performance was evaluated.

Results. To test network performance, a test stimulus in a given context was presented as an input to the network, and then the resulting activation pattern over the internal units was compared to the forementioned neural codes. To obtain recognition probability, an *ensemble* of 100 independently trained networks was created, and each network received 100 training blocks with different initial weights and

[1] Here we distinguish between temporal context and spatial context. The present study primarily concerns modeling of *temporal* context. An example of modeling of spatial context is Anderson and Murphy's (1986) brain-state-in-a-box model of context-dependent classification.

randomization of input stimuli. Performance of each network was tested separately as follows. A test input was deemed to be recognized as one of the nine stimulus members whose neural code had the greatest pattern similarity[2] to the activation pattern over the internal units responding to the test input. Using this definition, recognition probability was estimated as a proportion of the 100 networks in which the test input was recognized as a particular stimulus member.

Context-dependent recognition performance of the network is shown in Figure 2. The lightness of each rectangle represents recognition probability of a test input (ordinate) as a stimulus member (abscissa). The simulation included three conditions: without-context, with-context, and mixed-context. The left panel shows the result from the without-context condition in which the ambiguous stimulus (W) was presented to the network for recognition in the absence of any context. That is to say, because stimulus W was the first item of the input sequence, it could be a member of category A (A5) or category B (B5). As expected, the network was undecided between the two stimulus members, specifically, recognition probabilities being estimated as 0.52 and 0.48 for A5 and B5, respectively. The panel also shows that for the unambiguous stimuli in the input sequence such as A4 and A3, the network had no difficulty recognizing them correctly. A key demonstration of context-dependent recognition is shown in the middle panel for the with-context condition. Here, the ambiguous stimulus W was presented following four category A stimuli, A1 through A4. In this case, about two thirds of the time (i.e., probability of 0.67), the network recognized W as a member of the category of the preceding stimuli, that is, A5. The right panel shows the result from the mixed-context condition. Here, a random mixture of category A and B stimuli was presented followed by stimulus W (so context information was inconclusive). As expected, recognition of the ambiguous stimulus was blurred, with recognition probabilities of 0.54 and 0.46 for A5 and B5, respectively. To summarize, these results together confirm that the network can learn and use context to interpret ambiguous stimuli.

Analysis of Firing Patterns. To gain further insight into the underlying mechanisms of context learning in the network, we performed a detailed analysis of firing patterns. Recall that under Hebbian associative learning, "cell" assemblies are to be formed during training in the network. Indeed this prediction was confirmed. An analysis of between-unit correlations of firing activities of the 710 internal units can be summarized into the following three

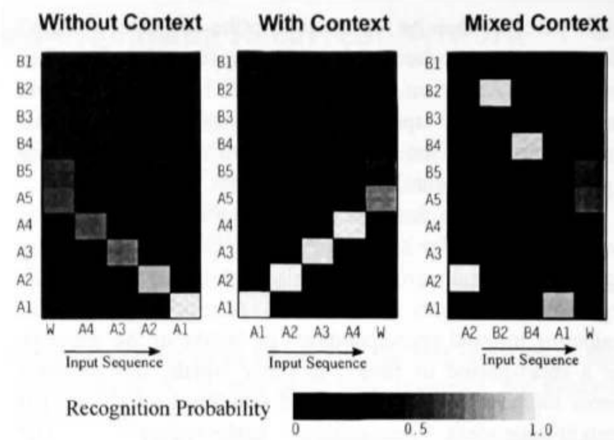

Figure 2. Simulation of context-dependent recognition of ambiguous objects. Recognition probabilities for the two critical cells in each condition are as follows: P(A5|W)=0.52 and P(B5|W)=0.48 for Without-context; P(A5|W)=0.67 and P(B5|W)=0.33 for With-context; P(A5|W)=0.54 and P(B5|W)=0.46 for Mixed-context.

observations.

First, we identified twenty-nine such assemblies. These mutually exclusive assemblies, whose sizes varied from a maximum 49 to a minimum 7, account for about 91% of all internal units. Cell firing of the units belonging to the same assembly is highly correlated with mean correlations ranging from 0.74 to 0.93. In other words, network units within an assembly tend to fire in unison.

Second, the analysis shows that each of the twenty nine assembles has its own firing selectivity, responding exclusively to a particular input stimulus, but never to other stimuli. In this sense then, we might view these assemblies as *internal representations* of the input stimuli. Whenever a particular stimulus is presented as an input to the network, the corresponding assembly of internal units would respond by turning themselves on — in unison — thus recognizing the stimulus.

Third, the analysis also reveals that there are three to four assemblies responding to a given input stimulus and, most interestingly, that each of the assemblies was bestowed with its own *temporal* selectivity. In other words, these assemblies fire at different time lags, one following another, after the input is presented to the network. To illustrate, suppose that stimulus A1 is presented to the network at time t. Then, assembly AS1 would fire at time (t+1), assembly AS2 at time (t+2), and assembly AS3 at time (t+3). This allows the network to maintain an internal representation of that stimulus for a few time steps by keeping it circulating among the assemblies. An important implication of this observation is that the self-organizing dynamics of the network somehow rendered itself into the creation of a time-delayed circuit that mimics short-term memory!

Figure 3 shows network firing during testing, that

[2] Pattern similarity between two vectors, **a** and **b**, is defined as $sim(\boldsymbol{a},\boldsymbol{b}) = \sum_{i=1}^{n} a_i b_i / \sqrt{\sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2}$, $(0 \le a_i, b_i \le 1)$.

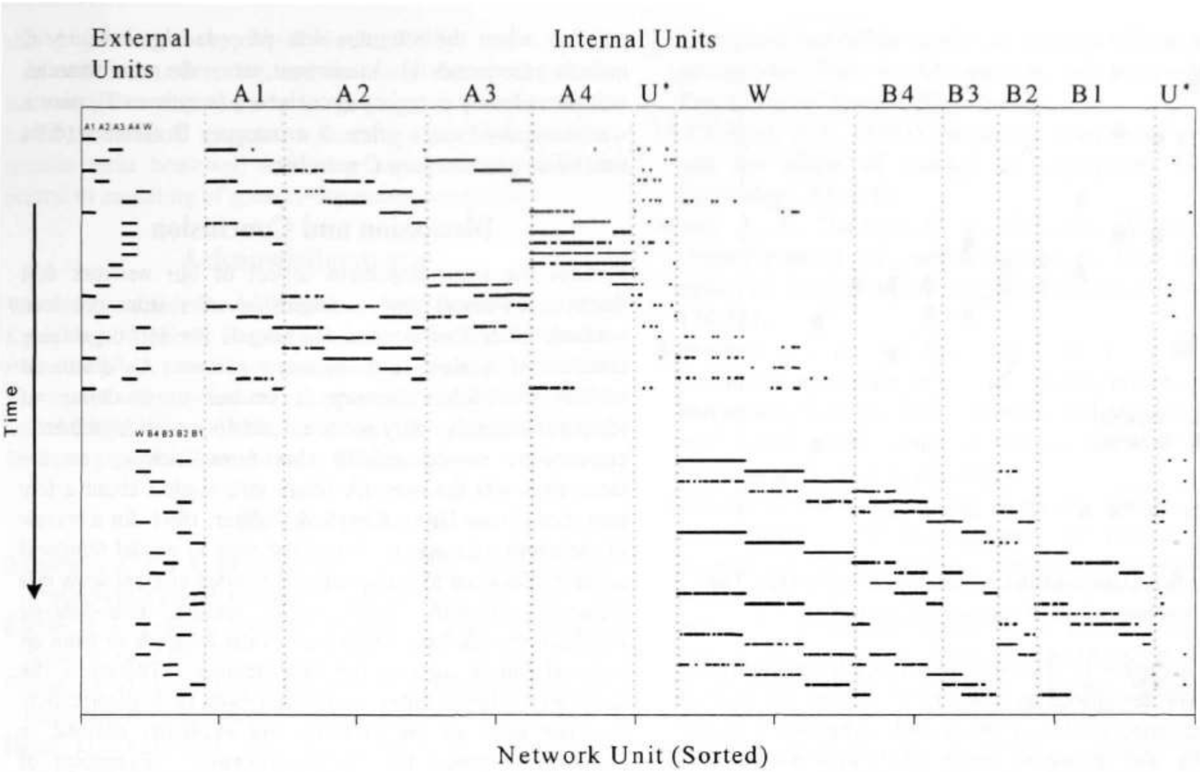

Figure 3. An example of network firing. Network units are sorted according to their assembly membership.

illustrates the above observations. In the figure each dot denotes firing of a particular network unit at a given time. Shown in the left-most block is firing of the 90 external units encoding the nine input stimuli. For example, the figure indicates the first seven inputs being A1 → A3 → A2 → A1 → A2 → W → A4. Firing of the 710 internal units, grouped according to their assembly membership, is shown in the right blocks of the figure. The dark horizontal lines indicate firing of assemblies of different sizes. Also note the stimulus-specific selectivity of these assemblies. For example, the three assemblies in the block denoted by A1 respond exclusively to stimulus A1, the next four to A2, the next three to A3, etc. The block U* includes the units that apparently belong to none of the assemblies.

Figure 3 also illustrates the short-term memory characteristic of the network, for example, by the successive firing pattern of the three assemblies associated with stimulus A1. Note that every time this stimulus is presented as an input to the network and then removed, these assemblies continue to fire for a few more time steps, meaning that information about that stimulus remains in the network for a while. The significance of this mechanism for context learning is obvious. That is, at any given time, firing activities of the network include not only the activity that has arisen as a response to the present stimulus but also activities pertaining to the past few stimuli. It is this mechanism that enables the network to interpret the present stimulus *in the context of* the past stimuli, thus exhibiting context-dependent recognition.

Mental Hysterisis Effect

In this section we describe results from another simulation of the mental hysterisis effect using a network similar to the one used above.

Methods. The network consisted of 84 external units and 616 internal units. A total of twenty input vectors that mimic the kind of stimuli used in studies of the mental hysterisis effect were generated as follows. The first input vector of size 84 was made up with eight ON units (x_i=1, i=1,...,8) and 76 OFF units (x_i=0 , i=9,...,84). The second input vector of the same size was created from the first vector by shifting its eight ON units to the right by four units (i.e., x_i=1, i=5,...,12; x_i = 0, i = 1,...,4, 13,...,84). Similarly, the remaining eighteen vectors were created by successively shifting the eight ON units to the right by four units. This overlapping-coding scheme means that a pair of successive input stimuli would be more similar to each other than to other stimuli, thus mimicking the face-women

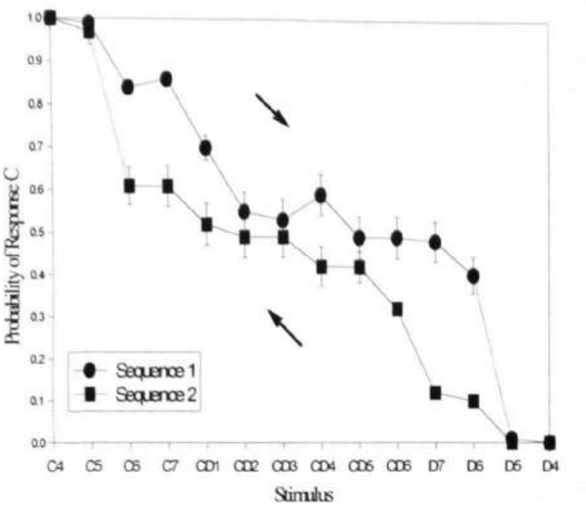

Figure 4. Simulation of the mental hysterisis effect. The error bars represent 95% confidence intervals.

pictures in Figure 1. The first seven inputs indicating C1, ..., C7 were designated as category C stimuli; the next six inputs indicating CD1,..., CD6 as both categories C and D; and finally, the remaining seven inputs indicating D7, ..., D1 as category D stimuli. In other words, category C contained thirteen stimuli, and six of them were also members of category D, and vice versa. The network was trained in blocks of 65 stimuli of either category C or D. Each block was made up with five repeated presentations of the same sequence of 13 inputs, that is, either {C1,...,C7, CD1,..,CD6} for category C blocks or {CD1,...,CD6, D7,..., D1} for category D blocks. The network received 100 training blocks. The neural code for each input stimulus was defined using the 616 internal units similarly as in the previous simulation.

Results. Network performance was evaluated by presenting to the network a test stimulus embedded in one of two sequences and then comparing the resulting activation pattern over the internal neurons to the neural code of that stimulus. Sequence 1 consisted of 20 stimuli {C1 → ...→ C7 → CD1 → ...→ CD6 → D7 →...→ D1}, and sequence 2 was the reverse of sequence 1. Recognition probability was estimated, again as in the previous simulation, using an *ensemble* of 100 independently trained networks. For each of these networks, an input stimulus was assumed to be recognized as a category C member if the activation pattern over the internal units was more similar to the neural code of the prototypic stimulus of category C (i.e., C4) than the neural code of the prototypic stimulus of category D (i.e., D4). Figure 4 shows recognition probability curves obtained for each sequence. The non-overlapping feature of the two curves reveals a mental hysterisis effect. That is, note that stimulus CD4 was recognized more often as a category C member (60% time) than as a category D member when the stimulus was preceded by category C members (sequence 1). In contrast, when the same stimulus was preceded by category D members (sequence 2), now it was recognized more often as a category D member (60% time) than as a category C member.

Discussion and Conclusion

Perhaps the most important aspect of our network that deserves comment and distinguishes ours from previous network models of context learning is the self-organizing creation of a short-term memory system. As discussed earlier, short-term memory is central to modeling of temporal contexts. Any recurrent network with asymmetric connections would exhibit short-term memory as the recurrency lets the network retain information about a few past stimuli (see Hertz, Krogh, & Palmer, 1991, for a review of recurrent networks). A popular way to model temporal contexts has been to set up what are called context units in a recurrent network. Such units receive time-delayed feedback signals from other part of the network or from an external source such as the experimenter. Therefore, the function of these context units, and even their connectivity to other units of the network, are explicitly defined or manually imposed by the experimenter. Examples of network models based on this approach are Jordan's network (1989), Elman's recurrent network of speech perception (1990), McClelland and Rumelhart's (1981) interactive activation models of letter perception (Rumelhart & McClelland, 1982; McClelland, 1991), and Wang and Arbib's (1990) sequence learning network, that uses a dual-neuron oscillator design for short-term memory, rather than a delayed feedback circuit. In contrast, our network creates its own context units, which preserve information about the recent past stimuli, through the locally adaptive process without any external mediation. Note the self-organizing nature of the network: the specific function and roles of the context units are not given *a priori*, but instead, we let the network develop its own representation of context.

Another aspect of our network that deserves comment is sparse connectivity. Because each unit is connected to a very small portion (5%) of the network, the unit would develop a representation of local events, reflecting only the activities of the units which it is connected to. Furthermore, if a subgroup of these interconnected units happens to fire in close temporal proximity, then Hebbian associative learning assures that this group of units will form an ensemble, through the reinforced, mutually excitatory connections. This way, the network produces ensembles that are locally sensitive but largely independent of one another. As discussed earlier, such ensembles were crucial to context-dependent recognition in our network. Recently, similar networks to this one have been successfully applied to simulations of other context-dependent phenomena, including sequence disambiguation (Minai, Barrows & Levy, 1994), sequence prediction (Levy & Wu, 1996), and short cut finding (Levy, Wu & Baxter, 1995).

In conclusion, the main contribution of the present study is its demonstration that a self-organizing recurrent network based on Hebbian associative learning can create its own context units bestowed with short-term memory, that is central to modeling of context-dependent recognition.

Acknowledgements

We wish to thank David Balota, Jerome Busemeyer, John Kruschke, Ira Hirsh, Caroline Palmer, and Elke Weber for their helpful comments on earlier versions of this manuscript. This work was supported by NIH MH48161 and MH00622 to WBL, and by the Department of Neurosurgery, Dr. John A. Jane, Chairman.

References

Anderson, J. A., & Murphy, G. L. (1986). Psychological concepts in a parallel system. *Physica, 22D*, 318-336.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*, 179-211.

Fisher, G. H. (1967). Preparation of ambiguous stimulus materials. *Perception and Psychophysics, 2*, 421-422.

Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. Wiley.

Hertz, J., Krogh, A., & Palmer, R. G. (1991). *Introduction to the Theory of Neural Computation* (pp. 163-187). Addison-Wesley.

Hunt, R. R., & Ellis, H. D. (1974). Recognition memory and degree of semantic contextual change. *Journal of Experimental Psychology, 103*, 1153-1159.

Jordan, M. (1989). Serial order: A parallel, distributed processing approach. In J.L. Elman & D.E. Rumelhart (eds.), *Advances in Connectionist Theory*. Erlbaum.

Labov, W. (1973). The boundaries of words and their meanings. In C. J. Bailey and R. W. Shuy (eds.) *New ways of analyzing variation in English*. Washington, D. C.: Georgetown University Press.

Levy, W. B. (1982). Associative encoding at synapses. *Proceedings of the fourth Annual Conference of the Cognitive Science Society*, pp. 135-136.

Levy, W. B., & Wu, X. (1996). The relationship of local context codes to sequence length memory capacity. *Network: Computation in Neural Systems, 7*, 371-384.

Levy, W. B., Wu, X., & Baxter, R. A. (1995). Unification of hippocampal function via computational and encoding considerations. *Proceedings of the Third Workshop on Neural Networks: from Biology to High Energy Physics. International Journal of Neural Systems, 6*, (supplementary issue), (D. J. Amit, P. del Guidice, B. Denby, E. T. Rolls & A. Treves, Eds.), Singapore: World Scientific Publishing, 71-80.

Maslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science*, 226-228.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception, Part I: An account of basic findings. *Psychological Review, 88*, 375-407.

McClelland, J. L. (1991). Stochastic interactive processes and the effect of context on perception. *Cognitive Psychology, 23*, 1-44.

Minai, A. A., Barrows, G. L., & Levy, W. B. (1994). Disambiguation of pattern sequences with recurrent networks. *INNS World Congress on Neural Networks, IV*, 176-181.

Rumelhart, E. 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*, 60-84.

Sweeney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645-660.

Tyler, L. K., & Maslen-Wilson, W. D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 683-692.

Wang, D., & Arbib, M. A. (1990). Complex temporal sequence learning based on short-term memory. *Proceedings of the IEEE, 78*, 1536-1546.