# Temporal Information Transformed into a Spatial Code by a Neural Network with Realistic Properties

Dean V. Buonomano* and Michael M. Merzenich

Neurons exhibit a wide range of properties in addition to postsynaptic potential (PSP) summation and spike generation. Although other neuronal properties such as paired-pulse facilitation (PPF) and slow PSPs are well characterized, their role in information processing remains unclear. It is possible that these properties contribute to temporal processing in the range of hundreds of milliseconds, a range relevant to most complex sensory processing. A continuous-time neural network model based on integrate-and-fire elements that incorporate PPF and slow inhibitory postsynaptic potentials (IPSPs) was developed here. The time constants of the PPF and IPSPs were estimated from empirical data and were identical and constant for all elements in the circuit. When these elements were incorporated into a circuit inspired by neocortical connectivity, the network was able to discriminate different temporal patterns. Generalization emerged spontaneously. These results demonstrate that known time-dependent neuronal properties enable a network to transform temporal information into a spatial code in a self-organizing manner—that is, with no need to assume a spectrum of time delays or to custom-design the circuit.

The elements in most neural network models consist of simple interconnected units that take the weighted sum of their inputs and generate an output by means of an activation function (*1*). These elements are meant to represent the summation of fast excitatory and inhibitory PSPs (EPSPs and IPSPs, respectively) and spike generation. Indeed, these models have been effective in performing complex computations and have provided many insights into how the nervous system processes information. Neurons, however, exhibit many additional properties, such as PPF, paired-pulse depression, voltage-dependent excitatory currents, rebound facilitation, bursting, and slow IPSPs and EPSPs (*2*). To date, few models have incorporated these properties, in part because the role of these properties in information processing is unclear. One possibility is that these properties contribute to the processing of temporal information in the range of tens to hundreds of milliseconds. Most network models have dealt primarily with tasks in which information is encoded in the spatial patterns of the inputs (*1*, *3*); yet, the nervous system must also extract information from the temporal features of input patterns. Speech recognition, frequency discrimination, music perception, and motion processing are a few tasks in which information is encoded in the temporal domain (*4*). It is not known how even a simple task such as discriminating between taps or tones presented at different intervals is solved, because such a solution cannot be based on spatial information from the inputs.

Keck Center for Integrative Neuroscience, University of California at San Francisco, San Francisco, CA 94143, USA.

*To whom correspondence should be addressed.

To address whether time-dependent neuronal properties may underlie temporal processing, we used a neural network composed of integrate-and-fire elements that incorporated PPF and slow IPSPs in addition to fast EPSPs and IPSPs (*5*) (Fig. 1). We focused on PPF and slow IPSPs because they have been described in some detail in cortical neurons (*6*–*8*) and can be incorporated efficiently into integrate-and-fire units. Furthermore, PPF may be particularly relevant to temporal processing because EPSP amplitude provides temporal information about recent spike occurrence. The time constants of the slow IPSPs and PPF were based on empirical data (*6*, *7*) and were the same for all elements. Excitatory (Ex) and inhibitory (Inh) elements were incorporated into a randomly connected circuit representing cortical layers 4 and 3 (*9*).

The simplest task studied was interval discrimination: two pulses were presented on the same input channels, with different intervals between them (Fig. 2). Each pulse represented a brief input such as a tap, tone, or flash. The first pulse of a stimulus initiated a set of excitatory and inhibitory interactions in the network (Fig. 2). Because of time-dependent changes imposed by PPF and slow IPSPs, the network is in a different state at the arrival of the second pulse. Thus, even if the second pulse is identical to the first some units will have different probabilities of firing depending on the interpulse interval. Indeed, between 25 to 50% of the Ex3 units exhibited interval-sensitive responses. These units can be used to encode temporal information. To demonstrate this in a more quantitative manner, we added an output layer to the network and trained it to recognize interval-specific patterns produced in layer 3 by five different stimuli (80-, 130-, 180-, 230-, and 280-ms intervals). All Ex3 units were connected to a number of output units equal to the number of stimuli being discriminated. A supervised learning rule was used to train each output unit to respond to a given stimulus (*10*). The output layer and the supervised learning rule are not meant to be part of a realistic simulation of temporal processing but a method to determine whether the activity pattern in the network can discriminate between different intervals. With the exception of changes in connection weights between the Ex3 and output units during training, there was no plasticity in the connection weights or time constants at any level of the network. After training, output units spiked (spikes are represented in yellow in Fig. 2) in response to the appropriate stimulus, which demonstrates that activity patterns produced in the Ex3 units contain sufficient information to code for temporal intervals (Fig. 2).

In addition to performing temporal discriminations, a biologically plausible model

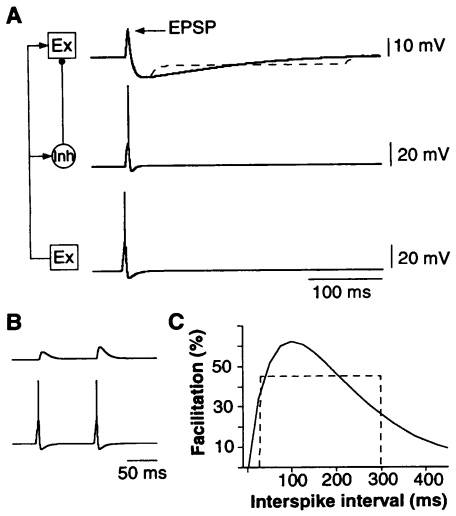


**Fig. 1.** Integrate-and-fire elements that incorporate slow IPSPs and PPF. Traces represent the voltages of the simulated integrate-and-fire elements. (**A**) Slow IPSP. By triggering a spike in the lower excitatory unit (Ex), a suprathreshold EPSP is elicited in the inhibitory unit (Inh), producing a fast EPSP followed by a slow IPSP in the upper Ex unit. For illustrative purposes, the connection strengths were increased. (**B**) PPF. The second of two consecutive spikes in an Ex unit will produce a larger EPSP in the postsynaptic unit. (**C**) PPF function. The time course and magnitude of the facilitation were estimated from empirical data from CA1 pyramidal neurons (*5*). For control experiments, the time-varying profiles of the slow IPSP and PPF were transformed into a step function from 30 to 300 ms, as shown by the dashed lines in (A) and (C).

must exhibit generalization in the temporal dimension. The network was therefore tested with intervals varying from 30 to 330 ms, and interval tuning curves were constructed for the output units (Fig. 3A). Although each output unit had been trained to respond to one of five stimuli, each exhibited a tuning curve centered around its trained interval. With these tuning curves, the network can represent any given interval between 30 and 300 ms by using a population code—that is, by using a combination of these units. Note that output unit 5, trained at the 280-ms interval, was significantly worse because a 280-ms interval approaches the limit of the time constants of the network. To demonstrate that the ability of the network to perform temporal discriminations was a result of the time dependency of the PPF and the slow IPSPs, we performed a control experiment in which PPF and slow IPSPs followed a step function (dashed lines in Fig. 1). In these control simulations, the output units were unable to discriminate among any of the trained intervals.

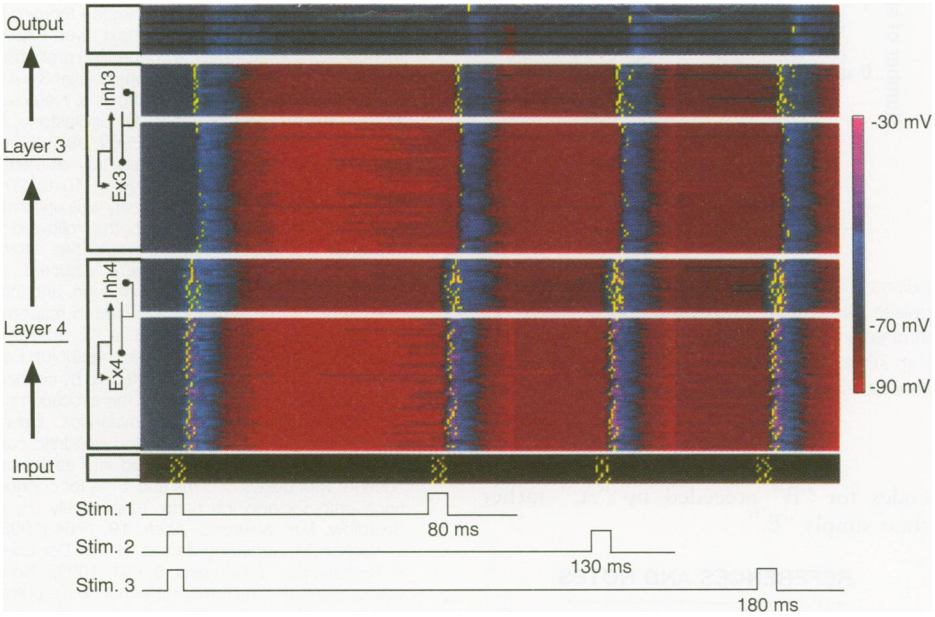


**Fig. 2.** Activity plot of the network in response to three different double-pulse stimuli. The four main blocks (Ex4, Inh4, Ex3, and Inh3) represent a sample of the units of the specified layer and unit type. Each horizontal line within each block represents the voltage of a given unit in time. Spikes are represented in yellow. Each stimulus consists of two pulses in which the second pulse was given at either 80, 130, or 180 ms. Each pulse consisted of a 5-ms burst of spikes in 50 of the input fibers. The plots of the shorter stimuli are overlaid on top of those of the longer stimuli. The output layer shows the response of three output units after they were trained to discriminate activity patterns of the Ex3 units elicited by the 80-, 130-, and 180-ms intervals.

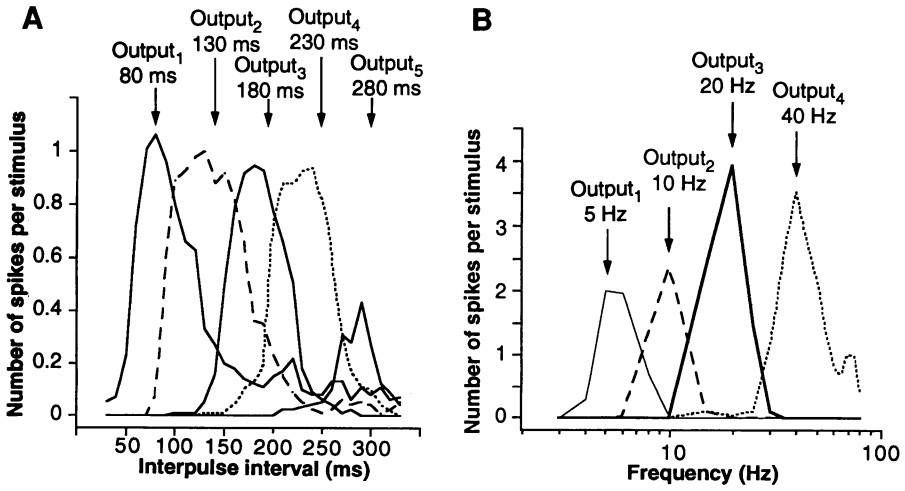


**Fig. 3.** Discrimination and generalization of temporal patterns. (**A**) To analyze the ability of the network to generalize, we tested it with double-pulse stimuli for intervals between 30 and 330 ms (10-ms steps), and interpulse interval tuning curves were then constructed. (**B**) The same network was also trained to discriminate between four stimuli consisting of 400-ms trains of pulses at frequencies of 5, 10, 20, and 40 Hz. After training the output units to recognize frequency-specific patterns, we tested the network with 20 frequencies (3 to 80 Hz). Tuning curves were constructed from 100 and 50 presentations of each stimulus for (A) and (B), respectively.

To examine the ability of the network to discriminate complex temporal patterns, we tested it with three tasks: frequency, random pattern, and phoneme discriminations. In the frequency discrimination task, each output unit was trained to one of four frequencies (5, 10, 20, and 40 Hz). After training, the network was tested with a range of frequencies. Each output unit exhibited a tuning curve centered around the frequency on which it was trained (Fig. 3B). The second task consisted of training the network to discriminate between four stimuli, each consisting of four pulses with randomly assigned interpulse intervals. Again, each output unit responded preferentially to the stimulus it was trained on (Fig. 4A); thus, each output unit was driven by a population of stimulus-specific Ex3 units. The third task consisted of using synthetic phonemes. Speech perception is a preeminent example of a task that relies on temporal cues. An important cue for discriminating voiced and unvoiced phonemes (that is, /ba/ and /pa/) is the voice-onset time (VOT; the time between air release and vocal cord vibration). /Ba/ tends to have a VOT of less than 30 ms, whereas /pa/ has a VOT of more than 30 ms (*11*). We trained a network with two outputs to discriminate /ba/ and /pa/ by training it with the two shortest (10 and 20 ms) and the two longest (70 and 80 ms) VOTs, and then we tested it with intermediate values (*12*). After training, the output units exhibited a response curve qualitatively similar to that observed psychophysically (*11*) (Fig. 4B).

We have shown that by using elements with realistic neuronal properties, temporal processing emerges as a result of state-dependent changes imposed on network dynamics. Without the need to change any model parameters, the network was able to perform interval, frequency, and complex pattern discrimination and to generalize to similar temporal patterns. We expect that increasing the complexity of the elements by including other neuronal properties and incorporating plasticity will further improve performance. A common form of associative plasticity, known as Hebbian plasticity, establishes that synaptic strength increases if both the pre- and postsynaptic elements are coactive. However, simulations that incorporated Hebbian plasticity made clear the difficulties of generalizing Hebb's rule to continuous-time networks with time-varying inputs. Hebbian plasticity leads units

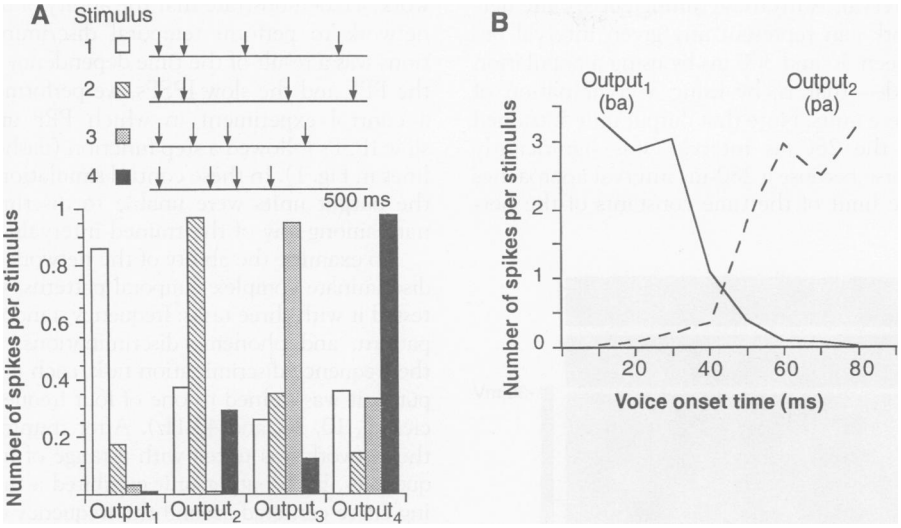


**Fig. 4.** (**A**) Discrimination of a random sequence of pulses. The network was trained to discriminate between four stimuli, each composed of four 5-ms pulses (arrows) with randomly chosen interpulse intervals between 50 and 250 ms. (**B**) Phoneme discrimination. The two output units were trained to discriminate 10- and 20-ms VOTs (perceived as /ba/) from 70- and 80-ms VOTs (perceived as /pa/). The network was then tested at all VOTs. Tuning curves were constructed from 100 presentations of each stimulus.

to respond to the most frequent synchronous input patterns, which in our tasks corresponded to the first pulse, which was common to all stimuli. Thus, the network became more responsive to the first pulse and less to the later pulses that contained temporal information. Hebb's rule is well suited to reinforce simultaneous activity coming from spatially distinct inputs, but it remains to be determined how well it will generalize to continuous-time networks that need to extract temporal as well as spatial information.

Previous models that have dealt with temporal processing have often assumed the existence of delay lines or elements with a spectrum of different time constants (*13*) or thresholds (*14*). These models often require an ad hoc architecture or do not generalize well to more complex temporal stimuli. Our simulations suggest that known time-dependent neuronal properties (not limited to PPF and slow IPSPs) with fixed and equal time constants permit a randomly connected network to transform temporal information into a spatial code (a place code). This transformation will occur at each layer of a network and will thus be amplified throughout layers. The general framework that emerges is that temporal combination-sensitive neurons (*15*) may arise as a result of time-dependent changes in network state—that is, if stimulus "A" then "B" is presented to an animal, "A" will produce a change in cortical network states as a result of time-dependent neuronal properties and stimulus "B" will then produce a pattern of activity that codes for "B" preceded by "A," rather than simply "B."

## REFERENCES AND NOTES


1. J. A. Anderson and E. Rosenfeld, Eds., *Neurocomputing* (MIT Press, Cambridge, MA, 1988); J. A. Anderson, A. Pellionisz, E. Rosenfeld, Eds., *Neurocomputing 2* (MIT Press, Cambridge, MA, 1990); D. E. Rumelhart and J. L. McClelland, *Parallel Distributed Processing* (MIT Press, Cambridge, MA, 1986).
2. R. S. Zucker, *Annu. Rev. Neurosci.* **12**, 13 (1989); E. H. Buhl, K. Halasy, P. Somogyi, *Nature* **368**, 823 (1994); D. A. McCormick, in *Synaptic Organization of the Brain*, G. M. Shepard, Ed. (Oxford Univ. Press, New York, 1990), p. 33.
3. Models that have focused on temporal tasks often have taken the unrealistic approach of collapsing time into a spatial dimension at the input level.
4. J. Gibbon and L. Allan, *Ann. N.Y. Acad. Sci.* **432** (1984); P. Tallal, A. M. Galaburda, R. R. Llinas, C. von Euler, *ibid.* **682** (1993).
5. Integrate-and-fire elements were based on those used in previous models [D. V. Buonomano and M. D. Mauk, *Neural Comput.* **6**, 38 (1994); F. Wörgötter and C. Koch, *J. Neurosci.* **11**, 1959 (1991)]. Briefly, each element was represented by a single compartment with a leak conductance ($g_{leak}$ = 0.1 μS, SD = 0.01). In addition to the leak current, each element has an after-hyperpolarization current, as well as excitatory and inhibitory synaptic currents. All synaptic currents had an instantaneous rise and an exponential decay ($\tau$ = 4 ms for all excitatory synapses and $\tau$ = 80 ms for all inhibitory synapses). We have modeled the γ-aminobutyric acid A ($GABA_A$) and $GABA_B$ fast and slow IPSP components as a single current with a fast onset and a slow decay. The threshold for all Ex and Inh units was −40 and −50 mV (SD = 4), respectively. PPF was simulated with an α function in which the time since the last spike modulated the strength of the synaptic conductance. The α function exhibited a peak conductance increase of 70% at 100 ms, which produced a maximal EPSP facilitation of approximately 60%.
6. T. Nathan and J. D. C. Lambert, *J. Neurophysiol.* **66**, 1704 (1991).
7. D. A. McCormick, *ibid.* **62**, 1018 (1989); R. J. Douglas and K. A. C. Martin, *J. Physiol. (London)* **440**, 735 (1990).
8. R. Creager, T. Dunwiddie, G. Lynch, *J. Physiol. (London)* **299**, 409 (1980).
9. The integrate-and-fire elements of Fig. 1 were incorporated into a circuit similar to a previously suggested neocortical circuit [R. J. Douglas and K. A. C. Martin, *Neural Comput.* **1**, 480 (1989)]. In particular, we simulated layers 4 and 3. The network was composed of 100 inputs and 150 and 250 elements in layers 4 and 3, respectively. In keeping with known empirical data, 20% of the elements in each layer were inhibitory and 15 to 20% of the connections onto each element type were inhibitory. The Ex units projected forward to both the Ex and Inh units in the next layer. Experimental data suggest that the connection probability between excitatory cells in neocortical layer 3 is 4 to 8% [A. Mason, A. Nicoll, K. Stratford, *J. Neurosci.* **11**, 72 (1991); A. M. Thomson, D. Girdlestone, D. C. West, *J. Neurophysiol.* **60**, 1896 (1988)]. Given that we were simulating relatively few elements, we used connection probabilities of 10 to 20% (between all unit types). Connectivity was uniformly random, and connection strengths followed a Gaussian distribution. For the simulations shown here, connection strengths were adjusted so that approximately 80% of the Ex4, Inh4, and Inh3 units and 25% of the Ex3 units fired in response to the first pulse (in general, 6 to 12 active inputs were necessary to reach the threshold). Performance was not qualitatively affected by connection strengths as long as neither the excitation nor inhibition dominated network dynamics. Delays representing axonal, synaptic, and dendritic conduction delays were incorporated into each connection, with delays of 3 ms and 1 ms for connections onto Ex and Inh units, respectively [H. A. Swadlow, *Soc. Neurosci. Abstr.* **19**, 1704 (1993); C. Welker, M. Armstong-James, H. van Der Loos, R. Kraftsik, *Eur. J. Neurosci.* **5**, 691 (1993)]. Noise was presented in the form of jitter during the pulses and by spontaneous activity in the input channels. In the simulations presented here, the spontaneous activity was set at 0.2 Hz, and the interval tuning curves were tested with spontaneous activity levels from 0 to 10 Hz. Over this range, there was a graceful degradation of the tuning curves, with the longer intervals being affected more.
10. The learning rule for the Ex3 to output unit connections established that at the end of each stimulus, recently active connections to the appropriate output unit were strengthened, and active connections to other output units were weakened. Connections were also decreased if an inappropriate output unit fired during a stimulus.
11. G. Ogden and D. Massaro, *Psychol. Rev.* **85**, 172 (1978); C. Wood, *J. Acoust. Soc. Am.* **60**, 1381 (1976).
12. Phonemes were synthesized with SynSen (for Macintosh). We then converted the spectrograms of the phonemes into spike trains by assigning each input unit of the network to a given frequency of the spectrogram. For this task, the connections from the inputs to the layer 4 units followed a Gaussian function in order to establish some spatial topography.
13. V. Braitenberg, *Brain. Res.* **25**, 334 (1967); J. W. Moore, J. E. Desmond, N. E. Berthier, *Biol. Cybern.* **62**, 17 (1989); S. Grossberg and N. A. Schmajuk, *Neural Networks* **2**, 79 (1989); D. W. Tank and J. J. Hopfield, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1896 (1987); W. E. Sullivan, *J. Neurophysiol.* **48**, 1033 (1982). The nervous system does use a spectrum of time delays for binaural sound localization; however, these delays are less than 1 ms [C. E. Carr, *Annu. Rev. Neurosci.* **16**, 223 (1993)].
14. P. S. Antón, G. Lynch, R. Granger, *Biol. Cybern.* **65**, 407 (1991).
15. D. Margoliash, *J. Neurosci.* **3**, 133 (1983).


16. This work was supported by National Institute of Mental Health fellowship F32 MH10431 and NIH grant NS-10414. We thank A. Doupe, S. Lisberger, H. Mahncke, M. Mauk, K. Miller, and J. Raymond for helpful comments and discussions.

29 June 1994; accepted 2 December 1994