Novel Lists of 7 ± 2 Known Items Can Be Reliably Stored in an Oscillatory Short-term Memory Network: Interaction with Long-term Memory

Ole Jensen and John E. Lisman[1]

Volen Center for Complex Systems
Brandeis University
Waltham, Massachusetts 02254

Abstract

This paper proposes a model for the short-term memory (STM) of unique lists of known items, as, for instance, a phone number. We show that the ability to accurately store such lists in STM depends strongly on interaction with the preexisting long-term memory (LTM) for individual items (e.g., digits). We have examined this interaction in computer simulations of a network based on physiologically realistic membrane conductances, synaptic plasticity processes, and brain oscillations. In the model, seven STMs can be kept active, each in a different gamma-frequency subcycle of a theta-frequency oscillation. Each STM is maintained and timed by an activity-dependent ramping process. LTM is stored by the strength of synapses in recurrent collaterals. The presence of preexisting LTM for an item greatly enhances the ability of the network to store an item in STM. Without LTM, the precise timing required to keep cells firing within a given gamma subcycle cannot be maintained and STM is gradually degraded. With LTM, timing errors can be corrected and the accuracy and order of items is maintained. This attractor property of STM storage is remarkable because it occurs even though there is no LTM that identifies which items are on the list or their order. Multiple known items can be stored in STM, even though their representation is overlapping. However, multiple, identical memories cannot be stored in STM, consistent with the psychophysical demonstration of repetition blindness. Our results indicate that meaningful computation (memory completion) can occur in the millisecond range during an individual gamma cycle.

Introduction

The previous paper in this series dealt with the short-term memory (STM) of lists of novel items. The current paper deals with the STM of novel lists of known items, as, for instance, a phone number. Each of these items (digits) already has a representation in long-term memory (LTM). Thus, when an item is presented, there is the potential for interaction with the preexisting memory of that item in LTM. It is this interaction that is examined here in networks having dual gamma/theta oscillations. These oscillations serve to multiplex the buffer so that up to seven memories can be simultaneously active, each in a separate gamma subcycle.
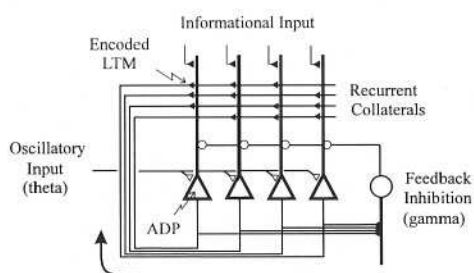
As we showed in the previous paper in this series (Jensen et al., this issue), the fidelity with which such a network stores *novel* items in STM is limited because there is no basis for error correction. The maintained firing that perpetuates short-term memory is caused by an activity dependent process intrinsic to neurons, the after-depolarization (ADP) (Storm 1989; Andrade 1991; Caeser et al. 1993; Libri et al. 1994). The ramp-like rise of the ADP provides the timing function necessary to make each memory active in the appropriate gamma cycle. In the presence of noise, such a sys-

[1]Corresponding author.

tem cannot maintain an STM accurately and cells will eventually fire in the wrong gamma cycle. Because an STM is encoded by the simultaneous firing of a subgroup of cells within a given gamma cycle, a cell that fires in the wrong gamma cycle would seriously degrade the memory. The ADP mechanism is not error correcting because the ADP is unique to each cell and information is therefore not redundantly stored. For novel items, there is no LTM in recurrent synapses about the item that can be used for error correction. Timing errors will therefore accumulate, and the STM of novel items will become degraded. In this paper we show that the situation is very different when there is a preexisting LTM representation for an item in LTM: Error correction is possible and the STMs can be stored accurately in their proper order.

The network that we analyze in this paper is shown in Figure 1. We assume that the network has already formed LTM for individual items (e.g., digits) through the process described in the previous paper (Jensen et al., this issue). This process involves fast *N*-methyl-D-aspartate (NMDA) channels that form associative connections between cells that fire in the same gamma cycle. These autoassociative LTMs are stored in the synaptic weights of the recurrent collaterals.

This paper focuses on answering several key questions about the network of Figure 1. One important question is whether a network of this kind can accurately maintain an STM in the presence of noise. As mentioned above, networks without LTM maintain STM quite inaccurately when challenged with noise. A second question is whether the net-
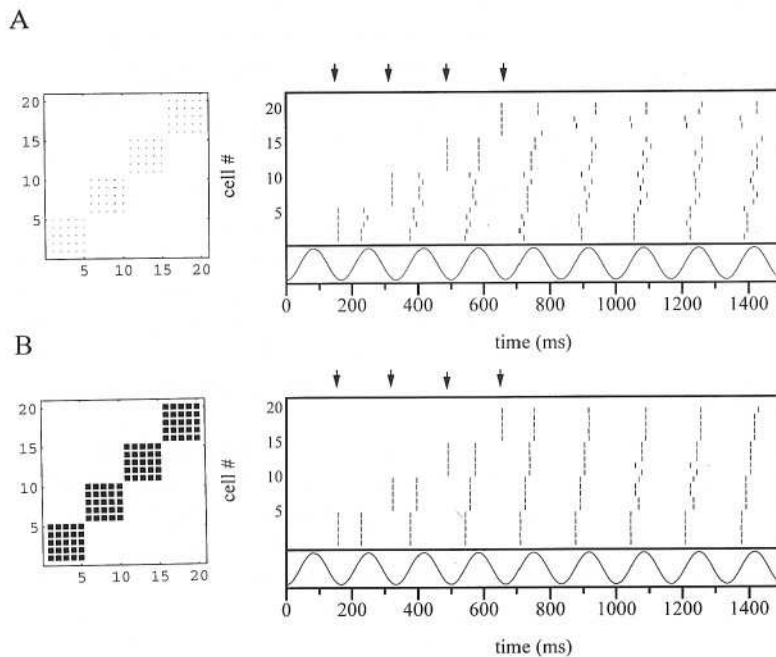
work will have properties similar to standard attractor networks, specifically, the ability to complete memories from degraded versions. An important underlying question that has not been addressed previously is whether meaningful neural computation can be performed *within* a gamma cycle. This is the time scale for computation that is required in the framework of our model. A third question regards coding. A key question is whether an overlapping code can be used or whether separate, nonoverlapping cell groups must be used to represent each item. The former is more efficient and would enable the network to store a larger number of memories. The simulations presented in this paper follow the methods described in the previous paper in this series (Jensen et al., this issue).
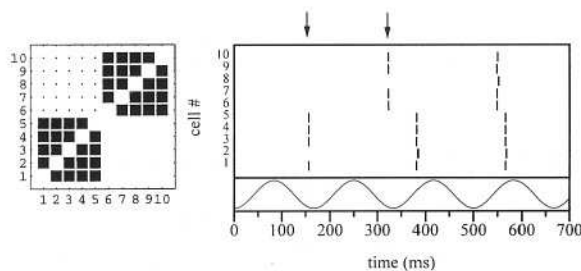
Results

The simulation in Figure 2B shows that memories introduced to the STM can be kept active and stable in the presence of noise, if they already have a representation in LTM. In the example, the simultaneous firing of cells 1 to 5 constitutes the first memory and cells 6 to 10 the second memory and so forth. The synaptic matrix to the left represents the four preexisting encoded patterns (see legend for explanation). In the previous paper (Jensen et al., this issue) in this series, we described how the network can encode such a representation. At about $t = 1050$ msec (Fig. 2B) memory 2 and 3 are degraded by the noise; some of the cells constituting the memories fire in the wrong gamma cycles. However, owing to the synaptic feedback, the memories recover their correct representation at $t = 1400$ msec. Novel memories without a representation in LTM cannot be maintained stable in STM in the presence of noise (Fig. 2A). The noise eventually causes the active cells to fire in incorrect gamma cycles, thereby degrading the STMs.

Attractor networks (Hopfield 1982; Amit 1989) can perform pattern completion, but the time it takes for such completion is generally not considered. Figure 3, shows that our network can perform pattern completion within a single gamma cycle. Prior to the simulation in Figure 3, two memory patterns were encoded (cells 1–5; cells 6–10). At time $t = 150$ msec, fragments (80%) of pattern 1 were presented to the system. One theta cycle later, the fragment is activated by

Figure 1: The integrated STM–LTM model. Pyramidal cells in the network have the ADP mechanism to store STM and recurrent collaterals with modifiable synapses. In this paper LTM for items have been encoded previously into these synapses. Gamma oscillations are generated by recurrent inhibition through an interneuron. Theta oscillation results from a 6-Hz input from an external source.

A

B

Figure 2: An STM can be faithfully maintained for items with a representation in LTM, even in the presence of noise. (A) In this simulation, no LTM is present. Four memory patterns consisting of cells 1–5, 6–10, and so forth are introduced to the network (arrows). The noise causes the firing to be delayed or advanced, and all synchronization is eventually lost thereby degrading the STMs. (B) Prior to the simulation, four memory patterns were encoded into LTM, as shown by the synaptic matrix. When these four patterns are introduced into STM, their representation remains stable despite the noise. Both in A and B we show the connectivity matrix of the network as a way of visualizing the information in LTM. The cell number (1–20) is specified on both axes; the size of the square at location i,j denotes the synaptic strength of the connection between cell i (y-axis) and cell j (x-axis). See Jensen et al. (this issue) for the equations and parameters defining the network.

the ADP pattern and completed 2 msec later. The way this completion occurs can be explained as follows: The cells that were fired by sensory input corresponding to the degraded pattern have an ADP that will cause these cells to fire on the second theta cycle. This firing triggers monosynaptic recurrent excitation and disynaptic feedback inhibition. Because the excitation occurs faster than the inhibition and because these pyramidal cells are selectively connected to the other cells that are part of the same memory pattern, these other cells will become active within a few milliseconds,

thereby completing the memory pattern correctly. Shortly thereafter, inhibition will arrive and end all firing in that gamma cycle. It can thus be seen how meaningful computation occurs within a gamma cycle.

Figure 4 demonstrates that overlapping memories can be actively maintained in the STM. Two overlapping memories, A and B (each represented by five cells), were encoded previously in LTM as illustrated by the synaptic matrix. The overlap consists of a cell (in this example cell 5, which constitutes a 20% overlap) that participates in the representation of both memory A and B. If memory A and B are both active in the short-term buffer, cell 5 has to fire twice within a theta cycle. The simulation in Figure 4 shows that recurrent collaterals makes this possible. The first time the cell fires is a result of the ADP and the input from the recurrent excitation from cells 1–4. The second time it fires is a result of the recurrent input from cells 6–9. Figure 4B specifically illustrates the synaptic current mediated by the AMPA conductance that enables cell 5 to be activated in two gamma cycles within the same theta cycle.

The simulations shown in Figure 5 were done to demonstrate the overall capabilities of a hybrid STM/LTM network. The task is to demonstrate that multiple, overlapping memories can be stored accurately in STM, even in the presence of noise, and
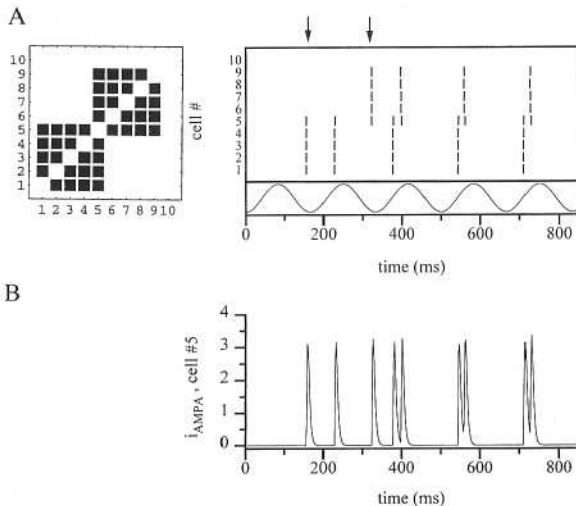

Figure 3: The integrated network is capable of pattern completion. Two memory patterns (cells 1–5, cells 6–10) were encoded in the synapses prior to the simulation shown, as illustrated by the synaptic matrix. When degraded versions of these patterns were presented to the network (arrows), the patterns were rapidly completed.

A

B

Figure 4: The integrated network is capable of storing two patterns in STM that overlap (a cell fires as part of both memories). (*A*) Two patterns were stored initially in the network as shown by the synaptic matrix. Both patterns have cell 5 in common, i.e., an overlap of ~20%. The overlap is also seen in the matrix: all cells are connected to cell 5. The recurrent collaterals enable cell 5 to take part in both patterns. (*B*) The recurrent synaptic AMPA current activating cell 5.

to show that the STM memory capacity is limited, as in our previous simulation (Lisman and Idiart 1995). Prior to the shown simulation, eight partially overlapping memories were encoded in LTM. The memories were 20% overlapping. During the shown simulation, we introduced memory items 1–7 into STM in successive theta cycles. Some of these memories were degraded when introduced. As seen, degraded patterns were completed and the seven memories remained stable despite the noise and the overlaps. When an eighth memory was introduced, a previously stored STM (#7) stopped being active. This demonstrates that the capacity of the STM buffer is limited to seven items.

Discussion

INTERACTIONS WITH LTM ALLOWS RELIABLE
STORAGE OF MULTIPLE STMS IN THE
MULTIPLEXED BUFFER

Humans can store novel lists of about seven known items in STM (Miller 1956) and recall the list items in correct order. In our simulation of STM (Jensen et al., this issue; Fig. 2A), memories

that do not have a representation in LTM are gradually degraded because noise disrupts the timing of the ADP. This timing is necessary to keep individual cells firing in the correct gamma cycle. A key finding presented here is that STM can be maintained reliably even in the presence of noise, provided that the network already contains a representation of each item in LTM. Because of the interaction of LTM with STM, timing errors can be corrected (Fig. 2B). This means that attractor principles apply to the short-term storage of list items in their proper order. This is remarkable, because there is in fact no LTM attractor that specifies which of the known items are on the list or their order.

The attractor property of the network for order occurs because errors in the timing of individual cells can be recognized on the basis of item information stored in LTM. This can be understood as follows: There are two types of excitatory drive on neurons during STM. The first is the ADP, an intrinsic membrane process that ramps up with time after the cell fires. The second is the synaptic input from the recurrent collaterals. Suppose a cell is part of a group that represents memory 2 and that it is supposed to fire in gamma subcycle 2. If the ADP in this cell is too large or a positive noise
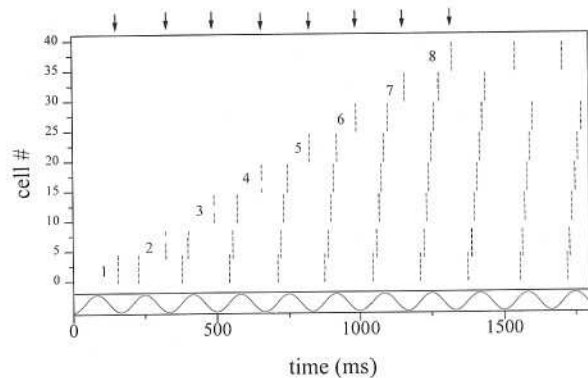
Figure 5: Demonstration that a hybrid LTM/STM network can stably store only seven memory patterns in STM in the presence of noise. Each memory consisting of five cells is introduced to the network on successive theta cycles (arrows). Several cycles later, an eighth memory is loaded causing the seventh to drop out. The simulation also shows the ability of the networks to maintain overlapping patterns active and to perform pattern completion. Patterns 1 and 2 overlap with respect to cell 5. Patterns 2, 3, and 4 are presented degraded to the network but were completed by the recurrents. Parameters as in previous simulations except for σ = 0.01.

fluctuation occurs, the cell will erroneously fire in subcycle 1. This will partially degrade memory 1, but as we will see, this degradation is only temporary. Now, consider what happens when we come to subcycle 2. On the basis of the ADP timing mechanism alone, the cell would not fire in the second subcycle, and this would partially degrade memory 2. However, because the cell receives synaptic inputs from the other cells that represent memory 2 it will fire in the second subcycle. This will not only prevent the degradation of memory 2 but will also reset the ADP timing ramp in that cell. As a result, when we come to the next theta cycle, the cell will tend to fire correctly in the second gamma subcycle. Thus, the fact that it fired in an incorrect gamma cycle in the first theta cycle does not lead to a permanent error in the phase of firing in subsequent theta cycles. In this way, the autoassociative LTM restores order information, even though no attractor information is available about order.

The idea of temporal segmentation of memories was put forward by Wang et al. (1990). Horn and Usher (1992) and Horn and Opher (1996) have suggested an oscillatory recurrent network separating memories in subsequent cycles. Our model is based on these principles, but we suggest a physiological realistic implementation of such a system.

We should emphasize that the network of Figure 1 can provide a short-term storage of lists but cannot store lists in LTM. Thus, a telephone number could be stored temporarily but not permanently. What the circuit can do through its STM properties (sample-and-hold) is to provide the repetitive input required to produce synaptic modification in some other network. The next paper in this series (Jensen and Lisman, this issue) shows how such input leads to LTM for lists in networks with slow NMDA channels.

OVERLAPPING CODE AND SPARSENESS

Our analysis of the hybrid STM/LTM network indicates that an overlapping code can be used to store memories in LTM, as in classic attractor network models. The existence of an overlapping code means that a given cell can fire in multiple theta cycles, as part of the representation of different memories. We do, however, have to require that the overlap is not too big. We have shown that the memory network can handle at least 20%

overlap among the represented memories. It has been pointed out by Rolls and Treves (1990) that memories in the brain are probably represented by a sparse neural code. In the case of the hippocampal memory storage, Treves and Rolls (1994) have estimated the sparseness to be $a = 0.02$ (a is defined in Jensen et al., this issue). Sparse coding has also been observed in recordings in the primate temporal cortex by Miyashita (1988).

PROPERTIES OF OVERLAPPING CODE; IMPLICATIONS FOR DETECTION OF THETA AND GAMMA IN SPIKE TRAINS

Consideration of cases where firing occurs in multiple gamma cycles makes it clear that it could be hard to detect theta and gamma frequencies by analyzing spike trains. Cells can fire in two or more gamma cycles, and there appear to be no restrictions about the gamma cycles in which firing occurs; for example, they could be neighboring subcycles or distant subcycles. Moreover, the gamma cycles in which firing occurs could change from theta cycle to theta cycle. For these reasons, firing occurs neither at gamma (20–60 Hz) or theta (5–8 Hz) frequency, even though these frequencies provide the organizational principles for controlling the timing of spike firing. Both frequencies are, of course, readily apparent in any signal that reflects groups of cells, such as field potentials or interneuron firing.

POSSIBLE EXPLANATION OF REPETITION BLINDNESS

Although partially overlapping memories can be stored in STM, two *identical* memories cannot be in the STM buffer at the same time. When cells fire together to represent a memory, the ADP will be reset and will lead to reexcitation of the memory on the next theta cycle. There is no basis for these cells firing again during the same theta cycle because the ADP takes time to build up (smaller parts of a memory overlapping with other memories can, however, be repeated within a theta cycle owing to recurrent excitation; Fig. 4). This implies that lists in which the same digit occurred twice would be impossible to hold in STM. At first this

seems to contradict our experience, but there is, in fact, psychophysical evidence to indicate that repetition is problematic. "Repetition blindness" has been well documented for visually presented stimuli that are presented rapidly and with precise control of interarrival times (MacKay 1969; Kanwisher 1987). For instance, if a sentence is presented in which a word appears twice, subjects only mention the repeated word once when asked to reproduce the sentence. When information is presented verbally, the same phenomenon occurs (repetition deafness), but it is less pronounced and requires a faster repetition rate (Miller and MacKay 1994). Obviously, in normal speech, repetition of a number can usually be detected, possibly because the first and second presentation of the number are spoken differently and so can somehow be encoded as "different" items.

BRAIN MECHANISMS OF STM

In this final section we discuss the literature as it relates to some of the biological underpinnings of our STM model. One important issue is the localization of the STM network. Current work suggests that STM or "working" memory is localized to prefrontal cortex (Goldman-Rakic 1995). One observation that points in this direction is that persistent firing, the hallmark of working memory, occurs in this region during the memory period in a delayed match-to-sample task. Abolishing this persistent firing with local cooling abolishes the behavior that depends on STM, a fact that distinguishes the frontal cortex from other regions that also show activity during the delay period. This delay is on the order of seconds.

Our model makes the assumption that the same network stores multiple STMs and does this through a multiplexing scheme based on theta and gamma oscillations. The presence of gamma oscillations in frontal cortex has been demonstrated by several methods (Llinas and Ribary 1993; Sanes and Donoghue 1993), but their link to memory storage remains unclear. Subthreshold oscillations in slice preparations of prefrontal cortex have been found by Llinas et al. (1991) and Gutfreund et al. (1995). Oscillations related to STM have been identified by Nakamura et al. (1992). They have found a 3.0- to 4.0-Hz oscillation in the activity of single neurons in the temporal lobe of monkeys performing a visual STM task.

Critical to maintaining order information in our model is the ramp-like rise of the ADP. Although there has been substantial progress in understanding the ionic basis of the ADP, the mechanism of the slow rise remains unclear. We favor the view that the afterpotential is actually the summation of a step-like ADP and a briefer after-hyperpolarization (AHP). According to this view, the slow rise of the ADP is actually attributable to the decline of the AHP. Because the AHP is triggered by a rise in intracellular Ca^{2+} that occurs after action potentials, it is straightforward to understand why timing ramps are reset by each action potential, as we have assumed [several types of AHP have been identified; the type required to account for the ramp-like rise of the ADP must have an >100-msec decay and would therefore be classified as a medium AHP according to Storm (1989)].

Although our work points to a mechanism by which multiple STMs could be stored by a single network, it is possible to take the contrary view that each STM is stored in its own network. Perhaps frontal cortex is actually a dense collection of independent networks, each devoted to a different memory. The observation that cells that remember the direction of a particular eye movement are localized to a small subregion of frontal cortex would support this view (Goldman-Rakic 1995). However, it is hard to imagine that such localization is general, because it would seem to severely limit the number of possible memories. Furthermore, there would still have to be interactions between the representations to store order information in STM. It thus seems highly desirable that the brain should be able to store multiple STMs in the same network. We have shown here that known brain mechanisms can accomplish this function.

References

Amit, D.J. 1989. *Modeling brain function: The world of attractor neural networks.* Cambridge University Press, Cambridge, UK.

Andrade, R. 1991. The effect of carbachol which affect muscarinic receptors was investigated in prefrontal layer V neurons. *Brain Res.* **541:** 81–93.

Caeser, M., D.A. Brown, B.H. Gahwiler, and T. Knopfel. 1993. Characterization of a calcium-dependent current generating a slow after-depolarization of CA3 pyramidal cells in rat hippocampal slice cultures. *Eur. J. Neurosci.* **5:** 560–569.

Goldman-Rakic, P.S. 1995. Cellular basis of working memory. *Neuron* **14:** 477–486.

Gutfreund, Y., Y. Yarom, and I. Segev. 1995. Subthreshold oscillations and resonant frequency in gunia-pig cortical neurons: Physiology and modelling. *J. Physiol.* **483:** 621–640.

Hopfield, J.J. 1982. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci.* **79:** 2554–2558.

Horn, D. and I. Opher. 1996. Temporal segmentation in a neural dynamic system. *Neural Comp.* **8:** 373–389.

Horn, D. and M. Usher. 1992. Oscillatory model of short term memory. In *Advances in neural information and processing systems 4* (ed. J.E. Moody, S.J. Hanson, and R.P. Lippmann), pp. 125–132. Morgan Kaufmann Publishers, San Francisco, CA.

Jensen, O. and J.E. Lisman. 1996. Theta/gamma networks with slow NMDA channels learn sequences and encode episodic memory: Role of NMDA channels in recall. *Learn. & Mem.* (this issue).

Jensen, O., M.A.P. Idiart, and J.E. Lisman. 1996. Physiologically realistic formation of autoassociative memory in networks with theta/gamma oscillations: Role of fast NMDA channels. *Learn. & Mem.* (this issue).

Kanwisher, N.G. 1987. Repetition blindness: Type recognition without token individuation. *Cognition* **27:** 117–143.

Libri, V., A. Constanti, M. Calaminici, and G. Nistico. 1994. A comparison of the muscarinic response and morphological properties of identified cells in the guinea-pig olfactory cortex in vitro. *Neuroscience* **59:** 331–347.

Lisman, J.E. and M.A.P. Idiart. 1995. Storage of 7±2 short-term memories in oscillatory subcycles. *Science* **267:** 1512–1515.

Llinas, R.R. and U. Ribary. 1993. Coherent 40-Hz oscillation characterizes fream state in humans. *Proc. Natl. Acad. Sci.* **90:** 2078–2081.

Llinas, R.R., A.A. Grace, and Y. Yarom. 1991. In vitro neurons in mammalian cortical layer 4 exhibit intrinsic oscillatory activity in the 10- to 50-Hz frequency range. *Proc. Natl. Acad. Sci.* **88:** 897–901.

MacKay, D.G. 1969. The repeated letter effect in misspellings of normals and dysgraphics. *Percept. & Psychophys.* **5:** 102–106.

Miller, G.A. 1956. The magical number seven, plus minus two: Some limits on our capacity for processing information. *Psychol. Rev.* **63:** 81–97.

Miller, M.D. and D.G. MacKay. 1994. Repeated words in computer-compressed speech are difficult to encode and recall. *Psychol. Sci.* **5:** 47–51.

Miyashita, Y. 1988. Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature* **335:** 817–820.

Nakamura, K., A. Mikami, and K. Kubota. 1992. Oscillatory neural activity related to visual short-term memory in monkey temporal pole. *NeuroReport* **3:** 117–120.

Rolls, E.T. and A. Treves. 1990. The relative advantages of sparse versus distributed encoding for associative neural networks in the brain. *Network* **1:** 407–421.

Sanes, J.N. and J.P. Donoghue. 1993. Oscillations in local field potentials of the primate motor cortex during voluntary movement. *Proc. Natl. Acad. Sci.* **90:** 4470–4474.

Storm, J.F. 1989. An after-hyperpolarization of medium duration in rat hippocampal pyramidal cells. *J. Physiol.* **409:** 171–190.

Treves, A. and E.T. Rolls 1994. Computational analysis of the role of the hippocampus in memory. *Hippocampus* **4:** 374–391.

Tulving, E., S. Kapur, F.I.M. Craik, M. Moscovitch, and S. Houle. 1994. Hemispheric encoding/retrieval asymmetry in episodic memory: Positron emission tomography findings. *Proc. Natl. Acad. Sci.* **91:** 2016–2020.

Wang, D., J. Buhmann, and von der Marlsburg. 1990. Pattern segmentation in associative memory. *Neural Comp.* **2:** 94–106.