Short Term Memory and Selection Processes in a Frontal-Lobe Model

Marius Usher
University of Kent
email: mu@ukc.ac.uk

Jonathan D. Cohen
Princeton University
email: jdc@princeton.edu

Abstract

We present a neural model that addresses the capacity of a frontal lobe system to hold up information for short periods of time and to perform response selection. In the model, reverberation states are sustained after stimulus offset, due to loops of recurrent excitation in neural cell assemblies and lateral inhibition is necessary to block an uncontrolled spread of activation. At high levels of inhibition the system performs response selection, and at lower levels it retains a number of units in active states, after stimulus offset. It is shown formally that such a system has capacity limitations: only a limited number of cell assemblies can be retained. Sequential presentation of a list of items is simulated and serial position curves characterised by recency are obtained. The model explains recency, list-length and presentation rate effects in immediate cued recall, as well as semantic effects and patterns of forgetting in Brown-Peterson type of experiments. A reduction in the strength of recurrent excitations explains the absence of lexical effects in tests of immediate memory for frontal lobe patients [1] and more extreme reductions result in impairments of response selection in dynamic aphasia [2].

Published in *Connectionst Models in Cognitive Neuroscience* (1999)
D. Heinke, G. W. Humphryes and A. Olson (Eds.). Springer-Verlag, pp. 78-91.

1 Introduction

Recent studies using a variety of methods, such as neurophysiological recordings [3, 4], functional imaging [5, 6, 7] and cognitive neuropsychology [1, 8, 9], have demonstrated the involvement of areas within the frontal lobes in memory activation over short periods of time, in particular for tasks that require semantic processing [1, 8, 6] and response selection [2]. Moreover these studies suggest a dissociation between short term memory (STM) characteristics within the frontal lobes and those found in the posterior areas; unlike the posterior areas, frontal lobe STM is mediated by attractor dynamics (neural reverberations) that can be sustained even during the processing of new, but irrelevant, information [4]. The aim of this paper is to present a biologically inspired model for active maintenance and response selection in the frontal STM system, and to test its predictions against behavioural data. Accordingly, it is proposed that while the posterior STM system is a modality specific (phonological, for verbal material) [10] and limited by a temporal decay window when articulation is prevented [11], the anterior system is more central and semantic [12] and uses active reverberations [13] which can sustain themselves and which are limited by displacement due to competing new information [14, 15].

2 Model

We assume that memory items in frontal lobe system are encoded within dedicated cell-assemblies [13]. Cells within each cell assembly, receive input from representations in the posterior system, and are connected among themselves by excitatory connections, which are formed in a slow learning process (not addressed here) corresponding to memory traces in LTM. [1] Accordingly, STM and LTM are not viewed as two physically distinctive stores, but rather STM can be thought of, as the active part of the long term memory [16, 17]. The recurrent weights stored in LTM are thus essential, in order to cause sustained states of activity which are the essence of STM, as demonstrated in neural models [18, 19, 20]. Lateral inhibition between units is also assumed in order to block an uncontrolled spread of activation [21]. Recurrent excitation and lateral inhibition also played a prominent role in neural models of STM developed by Grossberg[22] and in models by Taylor that address semantic effects in subliminal perception [23].

2.1 Model description

To begin, we assume a set of N units (cell-assemblies), and denote by x_i ($i = 1$, to N) their activation level. Each unit connects to itself by a recurrent excitatory connection of strength α (which reflects an effective interaction between cells within a cell assembly) and inhibits other units by inhibitory weights of strength β. In addition each unit might receive external sensory input (from posterior areas), denoted by I_i (Figure 1). The dynamics of such a system can be captured by:

$$\frac{dx_i}{dt} = -x_i + \alpha F(x_i) - \beta \sum_{j \neq i}^{N} F(x_j) + I_i + noise \tag{1}$$

where F(x) is the 'activation-function' which converts the input to a cell assembly into its activation response. Equation 1 reflects a passive decay of the activation toward the rest level (with a time constant, chosen as 1 [2]) in addition to excitatory self generated feed-back, lateral inhibition, external input and noise. The activation function is a sigmoid which keeps the activation variable within bounds, chosen here as:

$$F(x) = \quad 0, \quad for\ x < 0 \tag{2}$$

$$F(x) = \quad \tfrac{x}{1+x}, \quad for\ x > 0 \tag{3}$$

[1]Each assembly can be thought to encode a concept or a lexical item.
[2]This sets the time scale in equation 1 to that of synaptic currents.

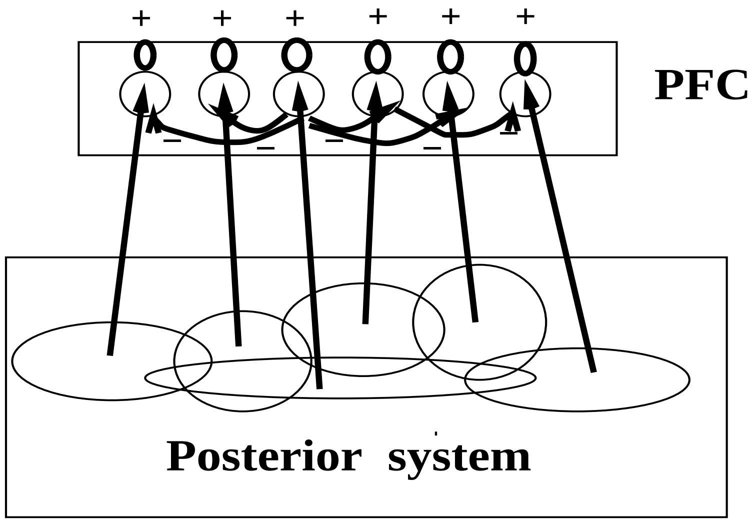

Figure 1: Model description. Each PFC cell assembly, is shown as a circle in the PFC box, which sends recurrent excitation to itself, inhibits the other PFC cell assemblies, and receives bottom up input from representations in the posterior system.

This activation function was chosen for two reasons. First, for small input ($x << 1$), the activation function is threshold-linear. Recent neurophysiological studies indicate that such a function describes more accurately the firing discharge of neural cells in response to current injection [24], for small values of current. Second, the saturation non-linearity is essential for large values of current ($x >> 1$) (due to refractory effects); equation 3 has a maximal activation of 1 which mimics the physiological bounds on firing rate.

2.2 Activating the system

The system is activated by clamping the inputs I_i for one or more units, for a fixed amount of time. Two situations will be addressed here. The first one, corresponds to tests of immediate memory, where following the period of stimulation, the inputs to all units are reset to zero ($I_i = 0$) and the system is allowed to evolve in accordance with Equation 1 for some time (corresponding to a short delay interval) after which the retainance of the input items is tested (it is assumed that a retained item can be always reported when cued for recall). In the following section we present formal solutions for the activated system after the offset of input, which we then apply to experimental data in immediate cued recall. The second situation (addressed in section 6) corresponds to sustained input activation to several units (corresponding to an ambiguous stimulus). Under this situation we will examine (using computer simulations) the ability of the system to solve the ambiguity and select a "response".

2.3 Formal solutions (in absence of input)

2.3.1 *Steady states*

Steady state solutions for equation 1 are obtained by requiring that $\frac{dx_i}{dt} = 0$ for all the units. The steady state solutions, in the absence of sustained inputs, $I = 0$, are characterised by:

$$x_i = \alpha F(x_i) - \beta \sum_{j \neq i}^{N} F(x_j) \tag{4}$$

Due to the symmetry of the equation (under permutation of units) we look for symmetric solutions, which are characterised by the number of active units, n, and their activation $x(n)$ (all other units are suppressed to zero):

$$x(n) = [\alpha - \beta(n-1)]F[x(n)] \tag{5}$$

Using the formula for the activation function (Equation 3) the activation of an n-active state is obtained:

$$x(n) = \alpha - 1 - \beta(n-1) \tag{6}$$

2.3.2 *Stability*

For a pair of units that are active together (with equal activation corresponding to equation 6) a fluctuation can arise due to noise, increasing the activation of one unit more than the activation of the other one and destabilising the solution; unless the fluctuation decays the activation of one unit will increase and the other will be suppressed. The condition for stability is obtained by subtracting the equations corresponding to Equation 1 for any pair of active units, say i and j. Denoting by x the difference in the activation of units i and j, $x = x_i - x_j$, and using the solution 6, one obtains:

$$\frac{dx}{dt} = x[-1 + \frac{\alpha + \beta}{(\alpha - \beta(n-1))^2}] \tag{7}$$

The fluctuation x will decrease only if the coefficient multiplying x is smaller than zero. Thus the condition for stability is:

$$\frac{\alpha + \beta}{[\alpha - \beta(n-1)]^2} < 1 \tag{8}$$

2.4 Capacity limitations

After a number of units, is transiently activated the system reaches a steady state solution, n units have an activation $x(n)$, which decreases with n due to the mutual inhibition, as shown in Equation 6. Notice that sustained solutions ($x(n) > 0$) exist only when $\alpha > 1$, requiring strong recurrent excitation. Under this condition and for weak inhibition $\beta << \alpha$ the activation of an n-state solution decreases linearly with n. Thus if, for example, $\alpha = 2$, and $\beta = .1$ up to 10 units can be activated together. As their number increases the activation corresponding to their steady state solution (equation 6) decreases continuously toward zero. The stability condition provides a further constraint, resulting in Equation 8. For any values of α and β the maximum n which satisfies equation 8

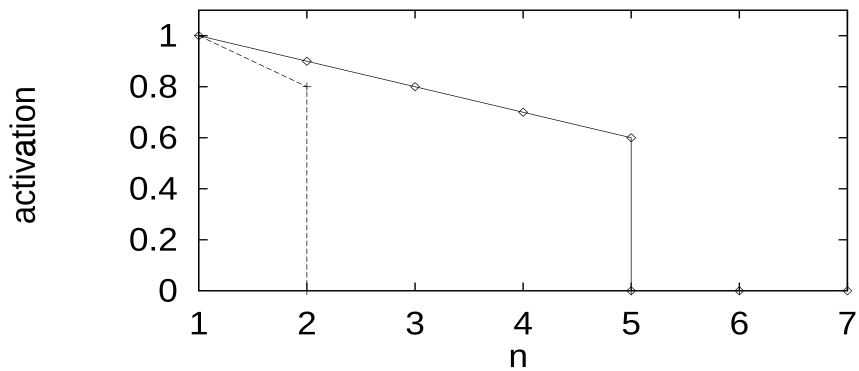

Figure 2: Stable solutions (Equations 6, 8), for $\alpha = 2$ and $\beta = .1$, with a maximum of n=5 co-active memory states (solid line), and $\alpha = 2$ and $\beta = .2$) with a maximum of 2 co-active states (dashed line).

can be found numerically. For the previous example, $\alpha = 2$ and $\beta = .1$, n=5 is the maximum number of active units satisfying a stable solution (solid line in Figure 2). Increasing the inhibition, β leads to an decrease of this critical number: $\beta = .2$ and $\alpha = 2$ results in maximal number of 2 (dashed line in Figure 2). The system can retain only up to this critical number of items (which depends on α and β). When more items are activated by the input stimuli, few of them are destabilised and decay as shown in Figure 3, for a system with $N = 9$ units. The excitation was chosen to be $\alpha = 2$ and two inhibition values were tested: a weak inhibition ($\beta = .1$) and an intermediate one [3] ($\beta = .2$). For intermediate inhibition, five of the 9 units received sensory input ($0 < t < 100$). As can be observed in Figure 3a, following the off-set of the input, only two out of those five units remain active. For weak inhibition, seven out of the nine units received input, and (as shown in Figure 3b), five out of those seven units remain active and the other two are suppressed to zero (in correspondence with Equation 8 and Figure 2). In the following section, the system described is applied to data in immediate cued recall[4].

[3] We retain the term, ‘strong inhibition’ for β values larger than .4 that result in a winner-take-all selection, further discussed in section 6.

[4] Cued recall is addressed since it maps naturally into the framework described and since it minimises additional processes such as serial order where we expect an important contribution from posterior STM systems [10].

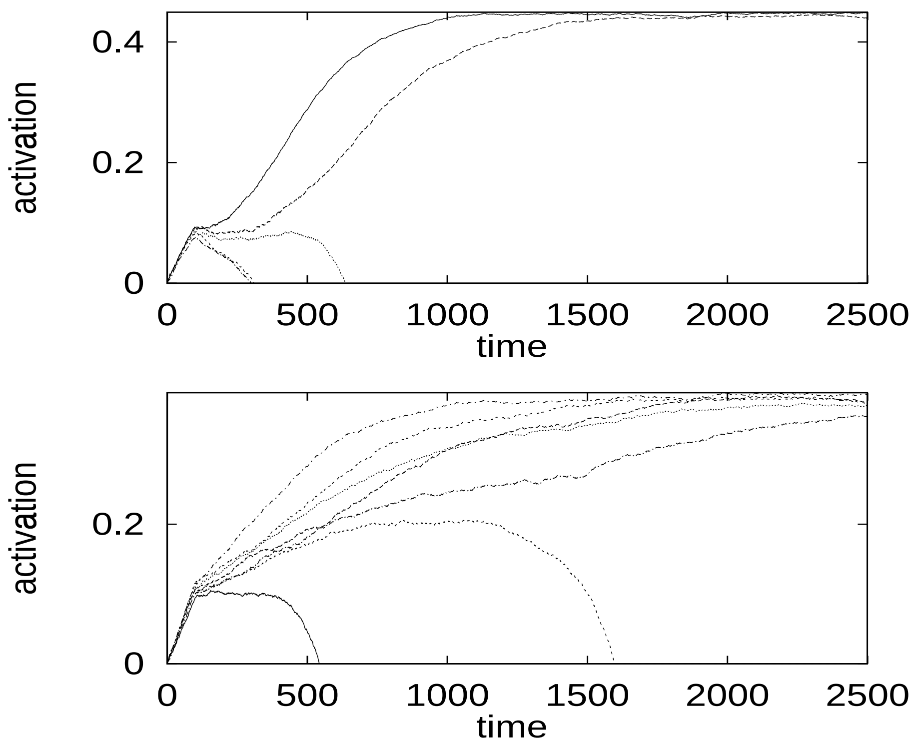

Figure 3: Simulation of the noisy system described by Equation 1; the x-axis represents iteration steps of the formula: $x_i = \lambda x_i + (1 - \lambda)[\alpha F(x_i) - \beta F(x_j) + I_i + noise]$ (for $\lambda = .99$, corresponding to Euler integration with $dt = .01$), for $\alpha = 2$ and simultaneous activation of items ($t < 100$). a) $\beta = .2$, b) $\beta = .1$.

3 Modelling cued recall

In cued recall, a list of n items is presented, sequentially, for retention, and immediately after that a cue to recall a specific item on the list is presented. The variable of interest is the probability of correct recall, as a function of the serial position of the cue-item on the list. Typical data [15] (illustrated in Figure 5b) shows a strong recency effect. Figure 4 shows the activation trajectories in a simulation where inputs were presented to the network sequentially. Each input was clamped, (to a value of .33) for a time window of 400 iteration steps while the input of all the other units was zero.

Notice first that, as more and more units are activated, this average activation decreases in a linear way in correspondence to equation 6. Second, at any moment in time, the activation of the last item presented is higher than the activations of the items presented before it. This is a consequence of the fact that the input favours the last presented items, breaking the symmetry of equation 1. When the number of inputs exceeds the capacity (in this case

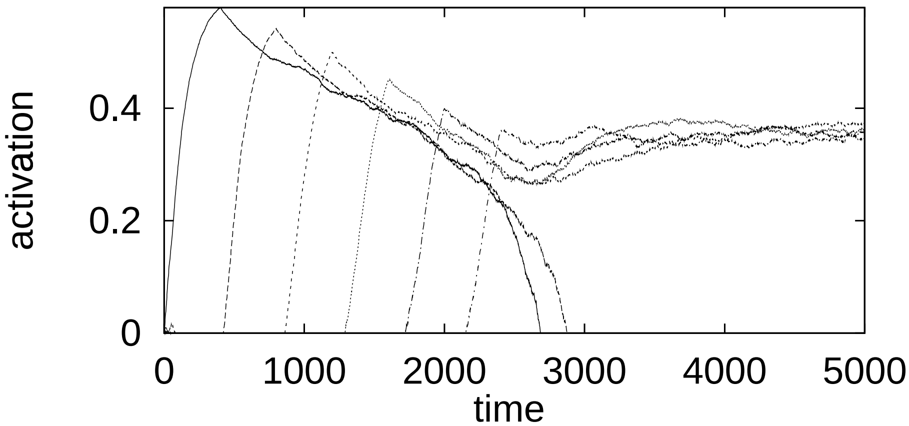

Figure 4: Activation trajectories for a sequential presentation of six items with $\alpha = 2$ and $\beta = .15$; only the last four items are retained.

4 items) some of the retained items are dropped. Earlier items have a higher chance of being discarded (Figure 4), since new items are initially activated at a higher level.

In order to obtain a serial recall curve we performed 500 such simulated trials. Serial position curves were obtained by measuring the fraction of trials in which an item corresponding to a specific position remained active at the end of the simulation, above some activation criterion (an arbitrary criterion, of .2 which does not affect the results, was chosen). The assumption is that if an item is still in active memory when a recall cue for it is given, recall is successful. The recency curve shares many characteristics found in experimental data [15] (see, Figure 5, below). Notice, in particular, the concave shape of the curve; performance decays only minimally at the first 2-3 items, after which it accelerates down.

3.1 Presentation rate effect

We tested the model, by computing recency curves for sequential presentation of objects, at two presentation speeds (400 and 800 time units per item). All other parameters were the same, and the simulation was run for such a time that there was always an additional of 2000 time units after the last presented item. The results of the simulation curves are presented in Figure 5a, and the data from Waugh & Norman (1965) in Figure 5b.

We observe, both in the data and in the model, that the slower presentation recency curve is a steeper sigmoid; the very recent items are recalled better, while the earlier ones are worse. Overall the total recall probability is rather similar (the two effects cancel each other out). In the model this is the outcome

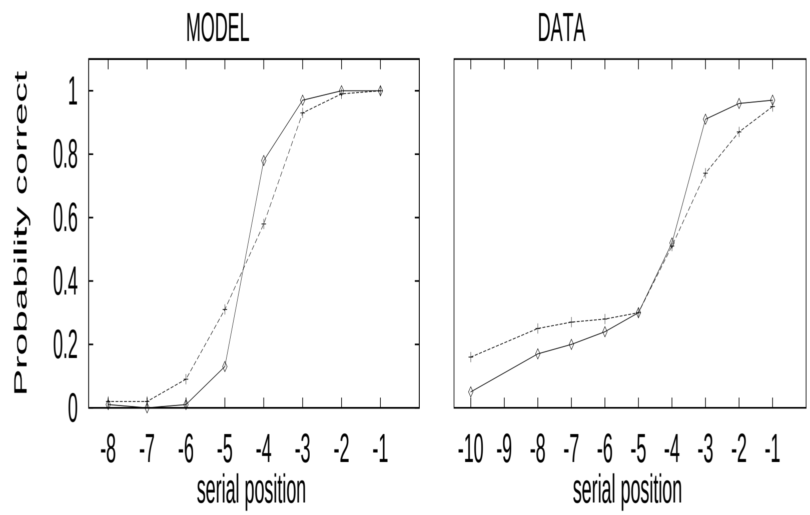

Figure 5: The effect of speed of presentation on cued recall. a) Model, b) Data (Waugh & Norman, 1965). Solid lines correspond to the slower presentation rates.

of the finite capacity limitation which is distributed among the various items. The reason for the difference is that, as the temporal window for presentation of each item increases, each new item has more time to be encoded at a stronger level than all the preceding items (Figure 4). As a result the activation of items in different serial positions are more strongly differentiated, leading to a less noisy process of selection for memory displacement. The simulation results, as well as the experimental data [15], indicate that the system is limited due to displacement of old items by new ones, and less so due to passive decay.

3.2 List-length and serial position effects

A prediction from this model is that serial-position interacts with list length. This is because no interference should take place in cued recall, until the capacity limitation of the system is exceeded. Three additional simulations were run with list length of 4, 5 and 6 items (time of presentation for each item was 400 time units, $\alpha = 2$ and $\beta = .15$). The predicted recency curves are shown in the left panel of Figure 6. Accordingly, we obtain perfect performance (at all serial positions) with a list length of 4 items. When list-length increases, the model predicts a gradual drop in recall probability which affects in particular the earlier items. Experimental data (with list-length of 3, 4 and 6 words), collected in our laboratory, is shown in the right panel of Figure 6.

4 Semantic effects in STM

Strong semantic effects have been often reported in immediate recall [25, 26, 27] and recognition [28, 29] of verbal material. Typical results show improved performance with lists of semantically related, relative to unrelated items (for

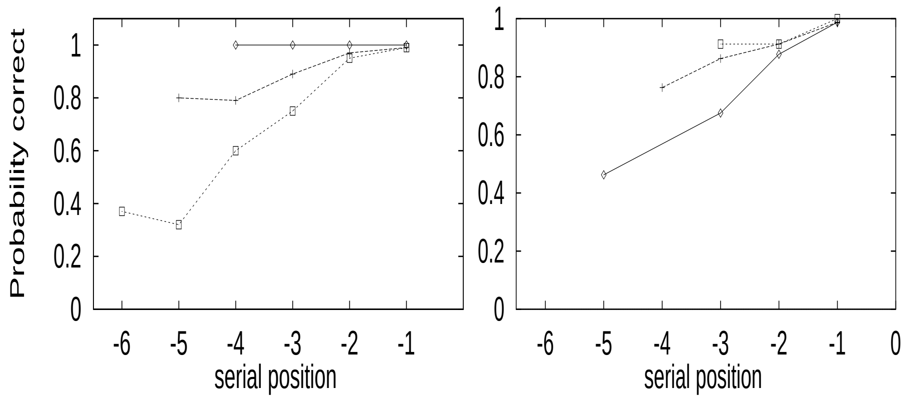

Figure 6: The effect of list-length on recency functions for cued-recall. Left: Model– Solid (upper) line: 4 items, dashed (middle) line: 5 items, dotted (lower), line: 6 items; Right– experimental data with list-length of 3, 4 and 6 (only up to four serial positions are tested).

example a list of 6 related items is better recalled than a list of 6 unrelated items [25].) One common interpretation of these results attributes the improvement to LTM retrieval [30]. This was supported by studies that compared free recall on lists of unrelated items (control) with lists where one cluster of related items was embedded either at the beginning or at the end of the list [31, 32]. Those studies showed much stronger improvement for middle-list cluster words than for end-list cluster words, relative to control lists (at the same serial positions). Since, items from middle serial positions are typically retrieved from LTM [15] this led to the conclusion that the semantic improvement is driven by LTM.

Here we show, however, that exactly these effects are obtained in our reverberation model without additional LTM encoding. The key assumption [25] is that semantically similar words support each other. In the model this is accounted for by assuming small excitatory links between units corresponding to associations between related words. As shown in Figure 7, while a list of unrelated items generates a typical recency curve (solid line with diamonds), a cluster-middle list shows a large improvement at the middle positions (dotted line with squares), and a cluster-end list shows an almost negligible improvement at the end-positions (dashed line with crosses). Thus characteristic features of the data [31] are reproduced. The explanation within the model is transparent. While unrelated items are typically displaced by more recent item, this happens to much less extent for semantically related items, which due to their mutual support, are relatively protected.

5 Forgetting in the Brown-Peterson paradigm

In the Brown-Peterson (BP) paradigm, subjects are given a number of items, and recall is tested after a delay filled with a distractor (e.g., counting-

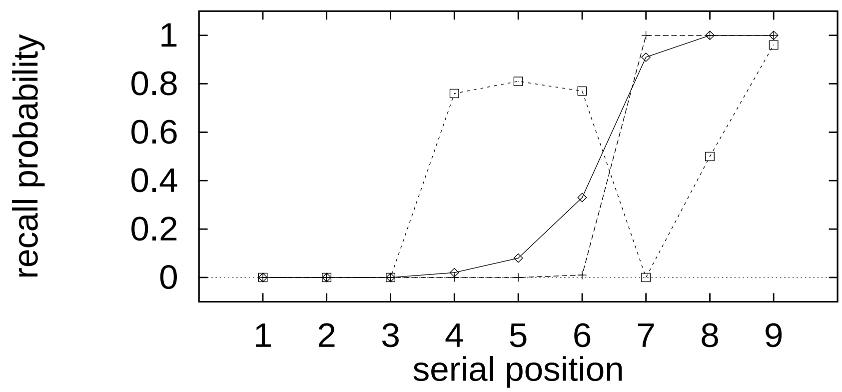

Figure 7: Serial position recall curves for i) unrelated items (solid line with diamonds), ii) cluster-middle list (dotted line with squares), and iii) cluster-end list (dashed line with crosses).

backward) task. At first the results suggested a decay mechanism, however following Murdock's results [33] who showed that the rate of decay is faster when there are more items to retain (Figure 8a), the decay mechanism was challenged [34]. Interestingly, with minimal auxiliary assumptions, the reverberation model can explain the dependence of the forgetting rate on the number of items to retain.

To model the BP experiments, we assume that after activating the units corresponding to the retention-set, subjects activate new items in the lexicon, one for each word/number they generate. The retention items are assumed to be encoded at a slightly stronger value of α (reflecting an attentional modulation; the difference was .09 in the simulations reported). Although they are relatively protected, some displacement of items from the retention set takes place when new items are activated. As shown in Figure 8b, this leads to a fast forgetting rate for larger response sets, where the items to be retained compete with each other over a finite retention capacity.

6 Winner-take-all and response selection

Until now we focussed on the ability of the system to retain information. However, a system based on lateral inhibition and recurrent excitation, (as described in Equation 1) can perform additional functions, usually associated with the frontal lobes, such as response selection [2, 6]. Accordingly when more than one unit receives input, the system can amplify the stronger input and suppress the weaker inputs, implementing a WTA network. This ability depends critically on the α and β parameters. To show this, we integrated numerically Equations 1 for the case when only two of the units receive an equal amount of input, $I_1 = I_2 = .1$ and independent noise. The value of the inhibition was relatively high (in order to perform response selection) $\beta = .45$, and two values

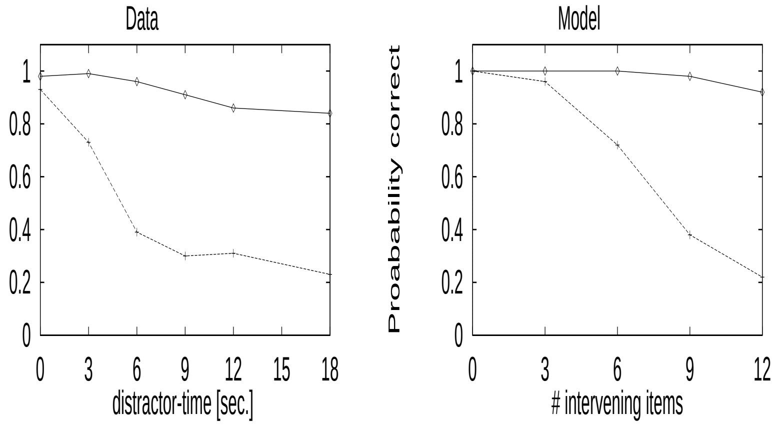

Figure 8: Probability of recalling the items in the response set (n=1: solid lines; n=3: dashed lines) as function of the delay time (or number of intervening items). Left: data from Murdock (1961). Right: model simulation ($\alpha = 2, \beta = .11, I = .33$).

of recurrent excitation were used: $\alpha = .8$ and $\alpha = 1.3$. The responses of the two units under both scenarios are shown in Figure 9. At low levels of recurrent

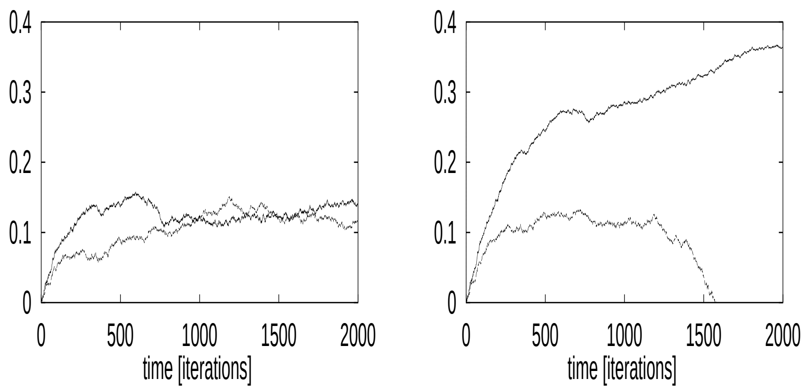

Figure 9: Response selection as function of the strength of recurrent excitation. Left column: weak recurrency ($\alpha = .8$); right column: strong recurrency ($\alpha = 1.3$)

excitation, the two units fluctuate around the same levels of activity. For the higher excitation, however, the system amplifies the noise fluctuations so that the ambiguity is solved in favour of one of the units.

It is interesting to note, that for weak recurrency ($\alpha = .9$) but larger inhibition ($\beta = 1$) and input ($I_0 = .3$) the time for selection depends on the size of the response set. As shown in Figure 10, in this situation the system performs easily a selection between two alternatives, but takes much longer (than the simulation run) to select among five of them. With increased excitation $\alpha = 1.2$ the time for selection does not depend on the response set (not shown). Thus weak recurrency mimics similar selection characteristics to those found in patients with dynamic aphasia [2].

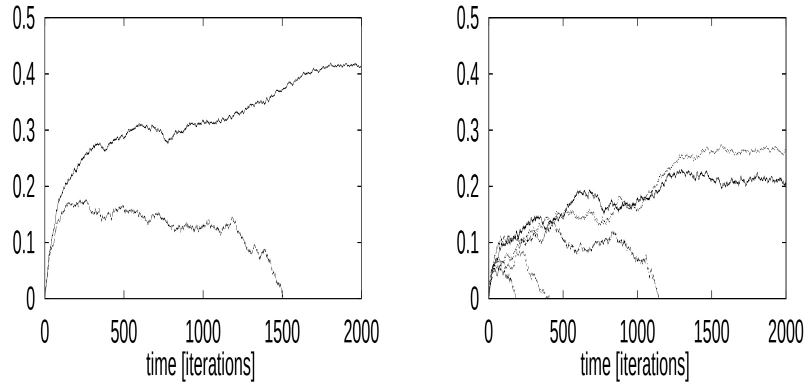

Figure 10: Response selection as function of the response-set Left column: small (n=2) response set; right column: large (n=5) response set. With weak recurrency ($\alpha = .9$), selection is much slower at large response sets.

7 Discussion

We presented a neural model that accounts for the retention of information for short periods of time under the form of neural reverberations, as reported in physiological recordings from the frontal cortex [3, 4]. The model explains effects of presentation rate, list length and semantic factors in immediate recall, as well as forgetting curves in the Brown-Peterson experiments, due to a displacement type of capacity limitation. The present model is limited in its range of application to the processes mediating short-term retainance and thus does not address primacy effects thought to reflect LTM contributions to recall [15]. Such effects could be addressed in an extended model that includes Hebbian learning and hippocampal storage [35]. Also the model addresses only item, and not order, memory. Additional components (such as a primacy gradient [36]) are needed in order to model serial order recall. Also, the model does not account for the phonological components in immediate memory [37]. Nevertheless, despite its simplicity and domain limitation, the model taps the components of active maintenance, characteristic of the frontal lobe system [3, 4] which seems to be sensitive to semantic, rather than to phonological, factors [1, 9]. [5] According to our approach, the short-term retainance of semantic contents is an essential part of context based control processes, mediated by the frontal lobe system [12].

The ability of this system to sustain more than one active representation following the offset of the stimulus, depends on the strength of the recurrent excitation α and of the lateral inhibition β. For the system described, the

[5] This is consistent with the reports of these studies, demonstrating a reduction of lexical and semantic effects in immediate memory of frontal lobe patients.

maximal span increases with the value of the excitation and decreases with the level of the inhibition. We assumed here that some small value of inhibition is necessary in order to prevent an unbounded spread of the activation across the whole system due to small overlap in the representations, as confirmed in our preliminary simulations of cell-assemblies with small overlaps. Since different tasks require different behaviour ('hold up few things in mind' or 'select one') it is appealing to assume that subjects have control on the level of excitation or inhibition. (This process might be mediated by other brain areas such as the Anterior Cingulate, which has been recently shown to activate in correlation with the level of response competition [38]). Nevertheless, those variable parameters might be within bounds that vary for every person and account thus for individual differences in span[6].

If patients that suffer from neurophysiological lesions in the frontal cortex have anomalously weak values of recurrent excitation, and if the network addressed is processing mainly semantic/lexical information, then such patients should suffer both in their ability to hold lexical and semantic information and in their ability to perform response selection. The former can explain the vanishing effects of word superiority in the STM span of frontal lobe patients [1], while the second could explain the deficit of response selection in patients with dynamic aphasia [2].

Acknowledgements Special acknowledgements are due to Henk Haarman for very stimulating discussions and to Dinkar Sharma for a critical reading.

References

[1] Martin R. C., Shelton J. R., & Yaffee L. S. Language processing and working memory, J. Memory and Language, 1994, 33, 83-111.

[2] Robinson G., Blair J., & Cipolotti L. Dynamic aphasia: an inability to select between competing verbal responses. Brain, 1998, 121, 77-89.

[3] Goldman-Racik, P. S. Working memory and the mind. Sci. Am. 1992, 267, 111-117.

[4] Miller, E. K., Erickson, C. A., & Desimone, R. (1996) Neural mechanisms of visual working-memory in prefrontal cortex of the macaque. J. Neurosci. 1996, 16, 5154-5167.

[5] Cohen J. D., Perlstein W. M,, Braver T. S., Nystrom L. E., Noll-DC et al. Temporal dynamics of brain activation during a working-memory task. Nature, 1996, 386 (6625), 604-608.

[6] Gabrieli J. D. E., Poldrack R. A., & Desmond J. E. The role of the left prefrontal cortex in language processing. Proc. Natl. Acad. Sci. USA, 1998, 95, 906-913.

[7] Posner M, & Pavese A. Anatomy of word and sentence meaning. Proc. Natl. Acad. Sci. USA, 1998, 95, 899-905.

[8] Romani C. & Martin R. A deficit in short term retention of lexical-semantic information. J. Exp. Psychol. Gen., in press

[6]We believe that this should map into complex span measures, such as the reading-span [39], which prevent reliance on phonological rehearsal.

[9] Shimamura A. Memory and frontal lobe function. In M. S. Gazzaniga (Ed.), The Cognitive Neurosciences, Cambridge, MIT Press, 1995.

[10] Warrington E. & Shallice T. The selective impairment of auditory verbal short-term memory. Brain, 1969, 92, 885-896.

[11] Baddeley A. D. Working memory. Oxford: Clanderon press, 1986.

[12] Cohen J. D., Braver T. S, & O'Reilly R. C. A computational approach to pre-frontal cortex, cognitive control and schizophrenia - recent developments and current challenges. Proc. Roy. Soc. London (B), 1996, 351, 1515-1527.

[13] Hebb D. O. The organization of behavior, New-York: Willey, 1949.

[14] Atkinson R., C. & Shiffrin R. M. The control of short-term memory. Scientific American, 1986, 225, 82-90.

[15] Waugh N., C. & Norman D., A. Primary memory. Psychol. Rev., 1965, 72, 89-104.

[16] Anderson J.R. The architecture of cognition. Cambridge, MA: Harvard, 1983.

[17] Cowan N., Activation attention and short-term memory, Memory and Cognition, 1993, 21, 162-167.

[18] Anderson J. A., Silverstein J.W., Ritz, S. W., & Jones, R.S. Distinctive features, categorical perception, and probability learning: Some applications of a neural model, Psychol. Rev. 1977, 84, 413-474.

[19] Amit, D. J., Brunel, N., & Tsodyks, M. V. Correlations of cortical Hebbian reverberations, J. Neurosci., 1994, 14, 6435-6445.

[20] Hopfield, J. J., Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Ac. Sci. USA, 1982, 79, 2554-2558.

[21] Keele S. W., & Neill W., T. Mechanisms of attention, In E.C. Cartrette & M. P., Friedman (Eds.), Handbook of perception, Vol. 9 , 3-47. New-York: Academic Press, 1978.

[22] Grossberg S. A Theory of human memory: self-organization and performance of sensory-motor codes, maps, and plans. In, Human Memory. Progress in Theo. Biol., R. Rosen & F. Snell, Eds., 1978.

[23] Taylor J. G. Breakthrough to awarness: a preliminary neural network of conscious and unconscious perception in word processing. Biol. Cybern. 1996, 75, 59-72.

[24] Ahmed, B., Anderson J. C., Douglas R., J. & Martin K., A., C. Comparison of current-discharge relationship of pyramidal neurons from the visual cortex in vitro and in the anesthetized cat. J. Physiol., London, 1996, 485, 10-11.

[25] Huttenlocher J. & Newcombe N. Semantic effects in ordered recall. J. Verb. Learn. Verb. Behav., 1976, 15, 387-399.

[26] Tulving E. & Patterson, R. D. Functional units and retrieval processes in free recall. J. Exp. Psychol. 1968, 77, 239-248.

[27] Hulme C. Maughan S, & Brown G. D. A., Memory for familiar and unfamiliar words, J. Memory and Language, 1991, 30, 685-701.

[28] Shulman H. G., Encoding and retention of semantic and phonemic information in STM, J. Verb. Learn. Verb. Behav., 1970, 9, 499-508.

[29] Raser G. A. Recoding of semantic and acoustic information in STM, J. Verb. Learn. Verb. Behav., 1972, 11, 692-697.

[30] Baddeley A. D., Retrieval rules and semantic coding in STM, Psychol. Bull., 1972, 78, 379-385.

[31] Craik F. I. M. & Levy B. A., Semantic and acoustic information in primary memory, J. Exp. Psychol., 1972, 86, 77-82.

[32] Levy B. A. & Baddeley A. D., Recall of semantic clusters in primary memory, Quart. J. Exp. Psychol., 1971, 23, 8-13.

[33] Murdock B. B Jr. The retention of individual items, J. Exp. Psychol., 1961, 62, 618-625.

[34] Melton A. W. Implication of short-term memory for a general theory of memory. J. Verb. Learn. Verb. Behav., 1963, 2, 1-21.

[35] McClelland J. L., McNaughton B. L., & O'Reilly R. C. Why there are complementary learning-systems in the hippocampus and neocortex - insights from the successes and failures of connectionist models of learning and memory. Psychol. Rev., 1995, 419-457.

[36] Page M., P., A. & Norris, D. The primacy model: a new model of immediate serial recall, in press.

[37] Shallice T. On the contents of primary memory. In P.M. Rabbitt & S. Dornic (Eds.), Attention & Performance, 5, 269-281, London: Academic Press, 1975.

[38] Carter C. S., Braver T. S., Barch D. M., Botvinick M. M., Noll D., & Cohen J. D. Anterior Cingulate gyrus dysfunction and selective attention deficits in Schizophrenia. Science, 1998, 280, 747-749.

[39] Daneman M. & Carpenter P. A., Individual differences in working memory and reading. J. Verb. Learn. Verb. Behav., 1980, 19, 450-466.