

# Effects of long-term representations on free recall of unrelated words

Mikhail Katkov,<sup>1</sup> Sandro Romani,<sup>2</sup> and Misha Tsodyks<sup>1</sup>

<sup>1</sup>*Department of Neurobiology, The Weizmann Institute of Science, Rehovot 76100, Israel;* <sup>2</sup>*Center for Theoretical Neuroscience, Columbia University, New York, New York 10032, USA*

Human memory stores vast amounts of information. Yet recalling this information is often challenging when specific cues are lacking. Here we consider an associative model of retrieval where each recalled item triggers the recall of the next item based on the similarity between their long-term neuronal representations. The model predicts that different items stored in memory have different probability to be recalled depending on the size of their representation. Moreover, items with high recall probability tend to be recalled earlier and suppress other items. We performed an analysis of a large data set on free recall and found a highly specific pattern of statistical dependencies predicted by the model, in particular negative correlations between the number of words recalled and their average recall probability. Taken together, experimental and modeling results presented here reveal complex interactions between memory items during recall that severely constrain recall capacity.

In the classical free-recall paradigm (Binet and Henri 1894; Murdock 1960, 1962; Roberts 1972; Standing 1973; Murray et al. 1976) subjects are presented with lists of unrelated words and are prompted to recall them in an arbitrary order. Participants typically cannot accurately recall more than five words. For comparison, people can successfully recognize thousands of items after a brief exposure (Standing 1973). Furthermore, more words can be recalled with appropriate cues (Tulving and Pearlstone 1966), indicating that the bottleneck for the recall is information retrieval from memory rather than a complete loss of information. It was proposed that the average number of items that can be recalled exhibits sublinear power-law scaling with the number of items in memory:  $N_{\text{rec}} \sim L^\alpha$ , where  $L$  is the number of presented items, with  $\alpha = 0.3\text{--}0.5$  (Murray et al. 1976). In another experimental paradigm, subjects were instructed to freely recall words with different first letters (“a” to “z”). Also for this paradigm, subjects could only recall a small fraction of all the words they can reasonably be estimated to know (Murray 1975). Across the different initial letters, the number of recalled words exhibits a similar power-law scaling with the estimated number of familiar words that begin with the corresponding letters (Murray 1975), indicating that a universal mechanism of information retrieval from memory mediates recall.

Most existing models of free recall proposed in the psychological literature ignore the effects of words’ identity as expressed in their neuronal representation on long-term memory and focus on describing associations between the items established during acquisition and rehearsal. These models are characterized by a large number of free parameters that are tuned to fit the recall data (Raaijmakers and Shiffrin 1980; Howard and Kahana 2002; Laming 2009; Polyn et al. 2009; Lehman and Malmberg 2013). We recently proposed a basic mechanism of information retrieval (Romani et al. 2013) where memorized items are assumed to be encoded (represented) by randomly assembled uncorrelated neuronal populations (Hopfield 1982; Amit 1989). The goal of the model is to identify general principles of information retrieval from long-term memory and thus it did not include effects of serial position of memory items during presentation that are specific

to free-recall paradigm. The retrieval mechanism proposed (Romani et al. 2013) specified the general limits on retrieval capacity that arise due to the randomness of long-term neuronal representations of items. In particular, the model robustly accounts for power-law scaling of retrieval capacity mentioned above, without having to tune a single free parameter (Romani et al. 2013).

We believe that the mechanism proposed by the model (Romani et al. 2013) plays a crucial role in all experimental paradigms that involve information retrieval from long-term memory, including free-recall paradigm. The goal of the current contribution is to derive statistical properties of recall that emerge due to random memory representations and to confront them with the large data set of recall experiments performed in the laboratory of M. Kahana (data courtesy of M. Kahana, see Materials and Methods). More specifically, we derived model predictions regarding the relation between the probability to recall an item and the size of its neuronal representation, and statistical interactions between items of different recall probability presented in the same list. We analyzed the experimental data in the way suggested by the model and compared the results with model predictions. We also extended our model to include the effects of order presentations specific to free recall data, and showed that the predicted effects are retained in the extended model.

## Results

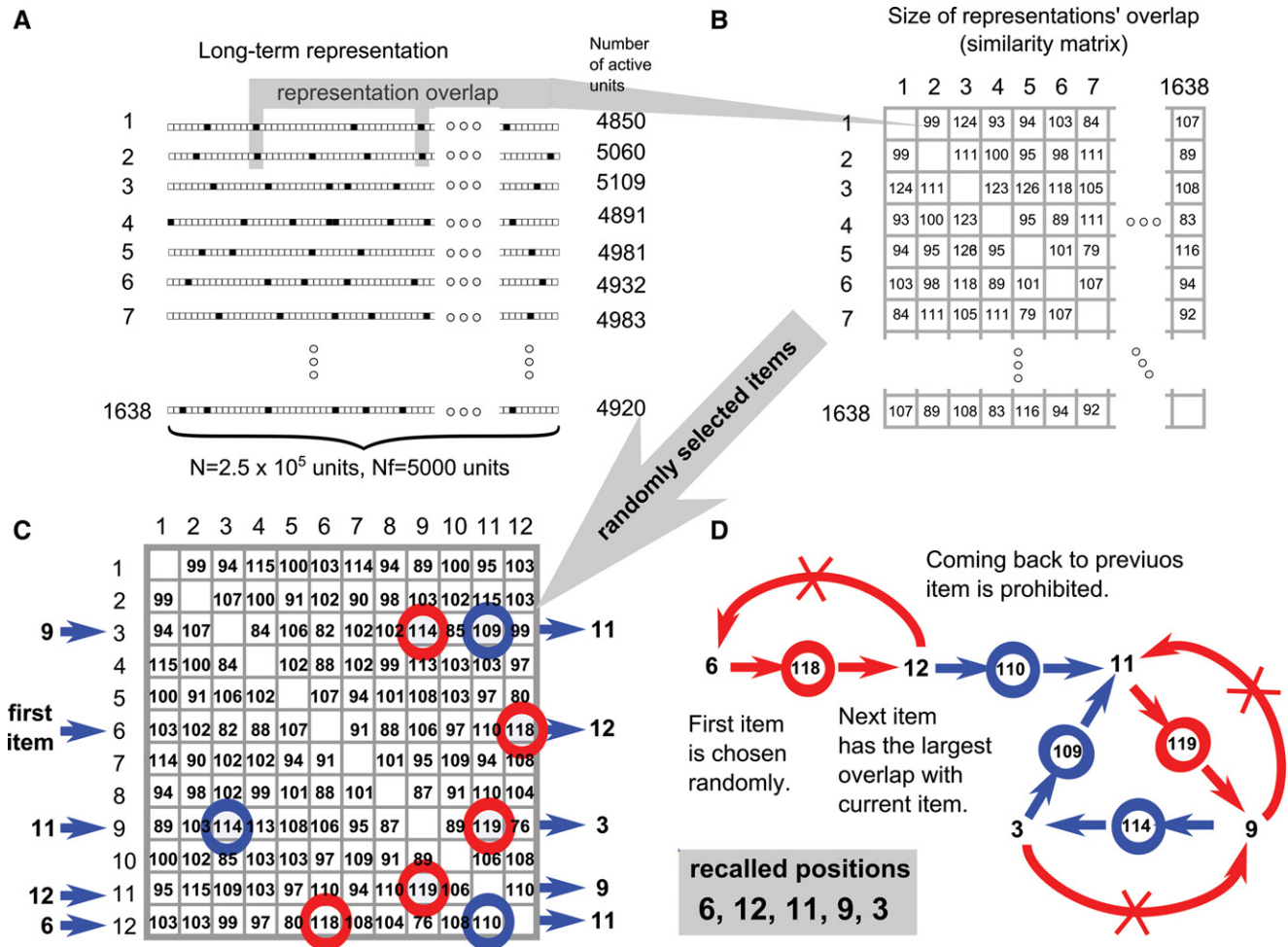
### Retrieval model simulations

The retrieval of an item in the model (Romani et al. 2013) is mediated by the activation of the corresponding neuronal population that encode this item in the long-term memory network (Hasselmo and Wyble 1997; Gelbard-Sagiv et al. 2008), and each retrieved item acts as an internal cue for the next item (Raaijmakers and Shiffrin 1980; Russo et al. 2008; see Fig. 1D for a schematic representation of the model). The size of the intersection

© 2015 Katkov et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first 12 months after the full-issue publication date (see <http://learnmem.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

**Corresponding author: misha@weizmann.ac.il**

Article is online at <http://www.learnmem.org/cgi/doi/10.1101/lm.035238.114>.



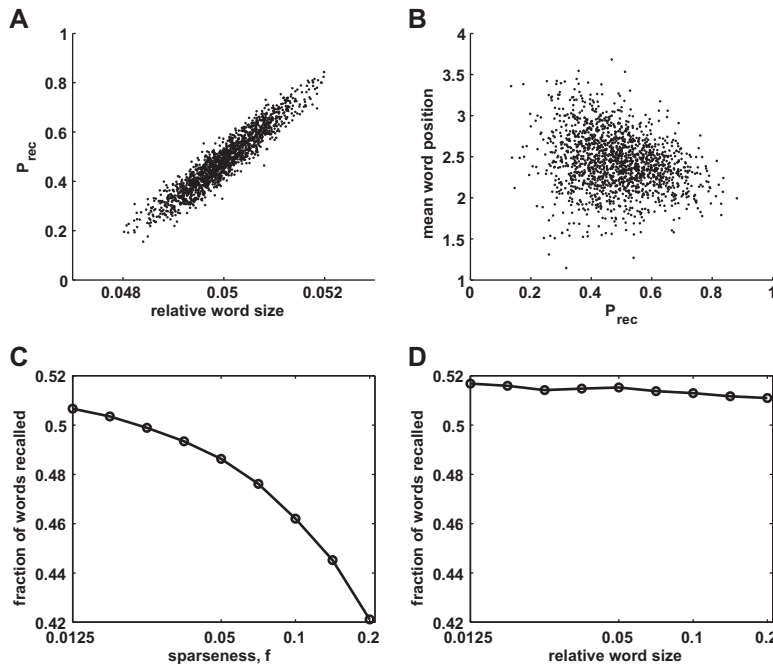
**Figure 1.** Schematic description of the retrieval model. (A) Long-term representations of items in memory dedicated network. Each black square represents a network unit participating in the encoding of the item. On average  $f = 2\%$  of units are active for each item. Due to the random nature of encoding the number of encoding units varies (examples are shown on the right side of the image). Some units simultaneously participate in encoding more than one word. (B) Similarity matrix. Each entry shows the number of units participating in encoding of pairs of corresponding words. (C) Similarity matrix for items chosen for a single trial. This is a corresponding submatrix of the full similarity matrix shown in B. Red circles indicate maximal similarity in a given row; blue circle indicated second maximal similarity in a row. (D) Schematic representation of the retrieval process. The first item is retrieved randomly. Next an item having largest or second largest similarity (see C) is selected (see Materials and Methods).

between populations (i.e., the number of neurons that represent both items) defines the “similarity” between corresponding items (Raaijmakers and Shiffrin 1980; Hills et al. 2012). In contrast to Howard and Kahana (2002) model, encodings of items are not required to be orthogonal, leading to differences in similarities between items (see Fig. 1B). The dynamics of retrieval is described by a sequence of active populations, with the next population chosen to be the one that has a maximal intersection with the currently activated one, not counting just “visited” item. The retrieval process is completely deterministic, hence the network eventually enters into a cycle, repeating the pattern of visited populations, and no more items can be retrieved (Fig. 1D; similar recall termination was also proposed in Laming 2009).

To derive testable implications of our model, we imitated the experimental protocol by generating a large pool of randomly generated memory representations and simulated multiple recalls of randomly assembled lists of 16 items (see Materials and Methods for more details). We then computed the intrinsic probability for a word to be recalled ( $P_{\text{rec}}$ ) when considering all trials where it was presented. We observed that across the pool of words,

$P_{\text{rec}}$  of a given item is to a large degree determined by the number of neurons that form its neuronal representation via a monotonically increasing relation (correlation coefficient is 0.94, Fig. 2A). Such a monotonic dependency emerges because larger neuronal groups statistically have larger intersections with other representations. Moreover, we observed that easier items (the ones with bigger representations and hence larger  $P_{\text{rec}}$ ) tend to be recalled earlier than more difficult ones (correlation coefficient is  $-0.24$ , Fig. 2B). This reflects the tendency of easy items to suppress the recall of more difficult ones and hence to limit the overall recall performance. As shown in the mathematical analysis of recall capacity of the model (Romani et al. 2013), it is precisely the variability in the sizes of the neuronal representations of memory items that is responsible for the decrease of the performance with increasing average relative size of neuronal representations,  $f$  (see Fig. 2C).

Comparing the results shown in Figure 2A and C, one might detect an apparent paradox: on the one hand, items with larger representations have a higher probability to be recalled, but on the other hand, increasing the average size of items



**Figure 2.** Statistical properties of recall in associative retrieval model. (A) Probability to recall an item across many trials versus the relative size of its neuronal representation. (B) Mean position of the item in recall sequence versus its recall probability. (C,D) Fraction of items recalled when underlying representations are random with average sparseness  $f$  (C) and when the fraction of active neurons representing particular word is exactly  $f$  (D).

representations results in lower average performance. The resolution of this paradox lies in the realization that the recall probability of an item as well as the general recall performance critically depends on the statistics of the pool of items in memory. In particular, the reason that the performance is declining with increasing  $f$  is the increasing “variability” in the sizes of neuronal representations, rather than the increase in the “average” size of representations. To further illustrate this point, we simulated the recall performance when all of the lists are composed of randomly chosen items that have representations of the same size. As shown in Figure 2D, for such a composition of the lists, the performance does not depend on the chosen size of neuronal representations.

These counter-intuitive results indicate nontrivial interactions between easy and difficult items in our recall model. To derive quantitative measures characterizing these interactions that can be directly tested with our data set (see Materials and Methods), we again considered lists of randomly chosen items of different sizes, and characterized each recall trial by three variables: the average  $P_{rec}$  of presented items ( $P_{rec-pres}$ ); the number of items recalled ( $N_{rec}$ ) and the average  $P_{rec}$  of recalled items ( $P_{rec-rec}$ ). One can characterize the statistical properties of recall interactions between different items by computing, across trials, the correlation coefficients between  $N_{rec}$  and either  $P_{rec-pres}$  or  $P_{rec-rec}$ . It is instructive to compare our model with the “naïve” one where each item is recalled independently with its corresponding probability  $P_{rec}$ . As shown in Figure 3A and B, these correlations are radically different for the two models. While for the naïve model,  $N_{rec}$  is positively correlated with  $P_{rec-pres}$  and only slightly correlated with  $P_{rec-rec}$ , the associative recall model produces almost no correlation of  $N_{rec}$  to  $P_{rec-pres}$  and negative correlation to  $P_{rec-rec}$ . In other words, lists that are comprised of easier items will on average not be recalled better than the lists of difficult items, and recalling easier items will typically result in fewer items

recalled on this trial. Note that these results are obtained for all values of sparseness  $f$ , i.e., they represent the generic feature of our retrieval model.

### Analysis of the data

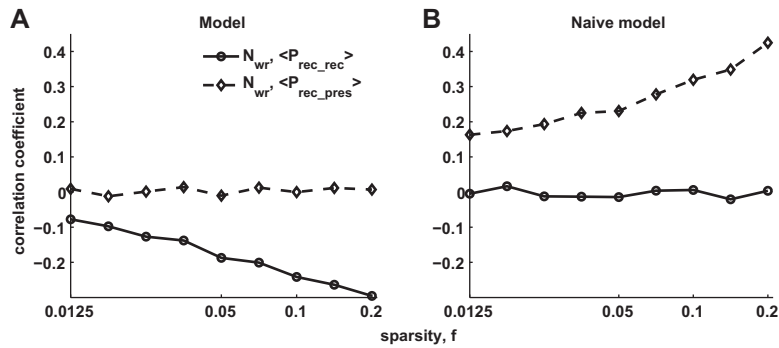
We tested the above predictions by analyzing the large data set of free recall experiments performed by 141 subjects with 112 trials per subject (see Materials and Methods). We first analyzed the data for the existence of easy and difficult words. Figure 4A shows the distribution of recall probabilities,  $P_{rec}$  for all the words in the data set, computed as a fraction of times a given word was recalled when presented. Figure 4B shows the scatter plot of the  $P_{rec}$  obtained from two random disjoint groups of subjects, showing that the words that were easy/difficult for one group tend to be easy/difficult for another group. These results demonstrate that different words exhibit significantly different probabilities to be recalled, and that this distinction is broadly reproducible across subjects. We also found that, as predicted by the model, easy words tend to be recalled earlier than the difficult ones (Fig. 4C).

Having established the distribution of recall probabilities across the word

pool, we computed the correlation coefficients between  $N_{rec}$ ,  $P_{rec-pres}$ , and  $P_{rec-rec}$  across the trials. To test for the reliability of these estimates, we performed nonparametric bootstrapping analysis of correlation values. One thousand bootstrap iterations were performed by randomly dividing the data set in two groups of subjects. Trials of one group were used to estimate  $P_{rec}$  for each word, and another group was used to estimate the correlations.  $N_{rec}$  exhibits negative correlation with  $P_{rec-rec}$  ( $-0.09 \pm 0.02$ ) and small but still positive correlation with  $P_{rec-pres}$  ( $0.02 \pm 0.01$ ). Figure 4D (upper panel) shows the histograms of obtained correlation coefficients. We conclude that the statistics of recall in terms of interactions between easy and difficult words is similar for the model and experiment (see Fig. 3). To stress this point further, we compared our results with a “naïve” model where each word is recalled independently with the corresponding recall probability (Fig. 4D, second panel). As expected,  $N_{wr}$  in this model is positively correlated to  $P_{rec-pres}$  ( $0.13 \pm 0.01$ ) and its correlation to  $P_{rec-rec}$  is very small ( $-0.00 \pm 0.01$ ), i.e., naïve model exhibits very different recall statistics than both the data and associative recall model. The corresponding histograms obtained from simulating the associative retrieval model (with sparseness parameter  $f = 0.02$ ) are shown in Figure 4D, third panel.

### Extension of the associative retrieval model

As opposed to more detailed models of free recall proposed in psychological literature (Raaijmakers and Shiffrin 1980; Howard and Kahana 2002; Laming 2009; Polyn et al. 2009; Lehman and Malmberg 2013), our model cannot account for some of the classical effects on recall resulting from the presentation order, i.e., primacy (better recall of initial words), recency (better recall of last words), and temporal contiguity (tendency to recall neighboring words in temporal proximity, see Murdock and Okada 1970; Howard and Kahana 1999). Moreover, our model is completely

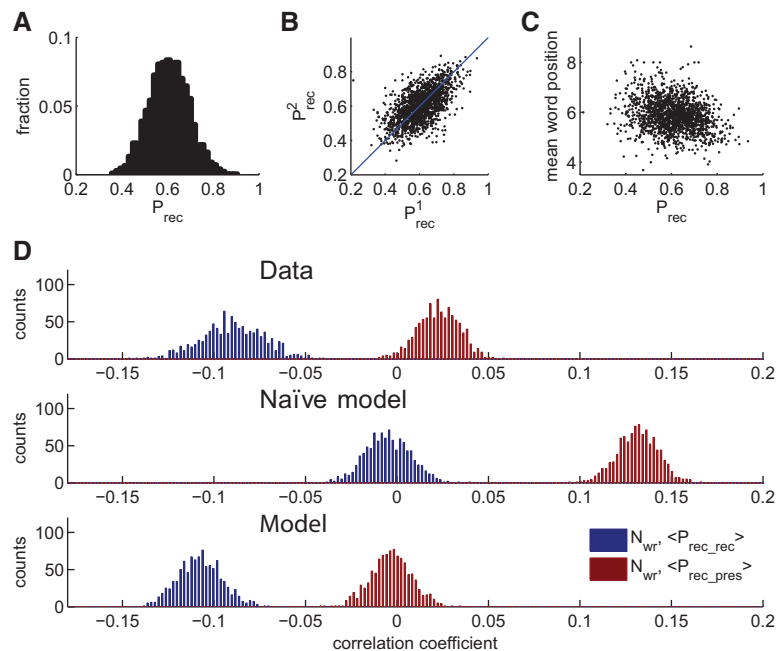


**Figure 3.** Prediction of the model. (A) Correlation coefficient between the number of recalled items and (dashed curve), the average recall probability of presented items; (solid curve), the average recall probability for recalled items for different values of sparseness parameter  $f$ . (B) Same as A for the naïve model (see details in text).

deterministic and hence cannot account for the distribution of interresponse times observed experimentally (Murdock and Okada 1970). In order to verify that the results presented above are not significantly affected by these effects, we extended the model by modifying, for each trial, the similarity measures between presented items in a way that depends on their presentation order and introduced weak stochasticity in the retrieval (see Materials and Methods). The extended retrieval mechanism can be considered as a simple version of the classical Search Associative Memory (SAM) model (Raaijmakers and Shiffrin 1980) with similarities consisting of two components—one based entirely on the internal neuronal representations of memory items and another based on the associations acquired during presentation and rehearsal. More specifically, we assumed that the similarity measures between presented words are modified, temporarily for each trial, in a way that depends on their presentation order (see Materials and Methods). Briefly, the first presented item has a strengthened representation in the network (primacy), thus its similarity to all other items presented on the same trial is enhanced by a certain amount. We also assumed that at the time of each item presentation, one of the previous items could be spontaneously recalled into short-term memory and acquire an additional similarity to a newly presented item, with probability of the spontaneous recall decaying exponentially with the serial distance between the items (contiguity). This mechanism of associating different items in the presented list is similar to the one considered in Raaijmakers and Shiffrin (1980); however, we assume that previous items can reenter the short-term memory even after falling out of it, due to spontaneous recall between the presentations of new items. To account for the asymmetric nature of contiguity, the forward association is stronger than the backward one. Finally, to account for recency, we assumed that after the list presentation, one of the last items remains in the short-term memory and consequently becomes the first item be-

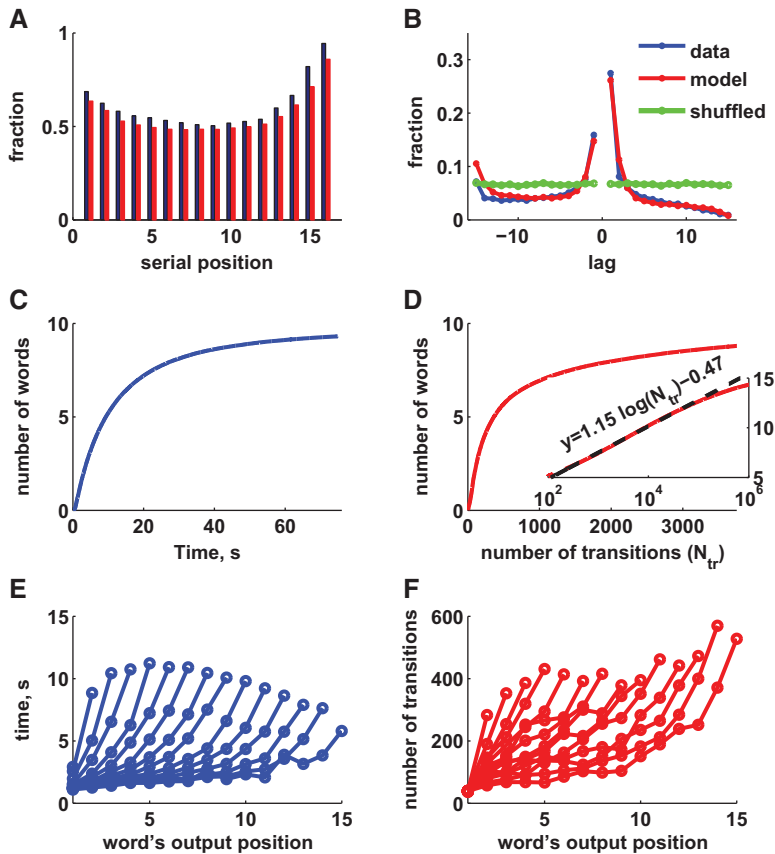
ing recalled. We also introduced stochastic retrieval mechanism by assuming that the transitions occur probabilistically with the corresponding probabilities increasing with the similarities between consecutive words, similar to SAM.

We found that the model is flexible enough to fit the recall data well with respect to serial position curve (recall probability versus position of a word in the list; Fig. 5A) and the shape of the contiguity effect as described in Howard and Kahana (1999) (Fig. 5B). We also obtained similar time course of recall, characterized by the average cumulative number of words recalled as a function of time (Fig. 5C versus D) and average time to recall a new word computed separately for trials with different number of words recalled (Fig. 5E versus D, see also Murdock and Okada 1970). Comparing Figure 5C and D, one can estimate the frequency of oscillations to be  $\sim 50$  Hz, i.e., in the range of a Gamma rhythm, but one should note that this estimate could depend on the particular noise model we chose and could be different in a more realistic neural network model. No precise tuning of parameters is needed to account for the data, illustrating the robustness of the proposed recall mechanism (see Materials and Methods for more details). Since the retrieval process in the extended model is stochastic, theoretically all presented items will be recalled given enough time. We found however that even after  $10^7$  transition cycles the recall performance did not saturate. The long-time behavior of the average number of recalled words is



**Figure 4.** Statistical properties of recall data. (A) The distribution of recall probabilities across the words in the pool. (B) Recall probabilities of the words in the pool, computed separately for two randomly chosen complementary groups of subjects. Correlation coefficient is positive (0.50,  $P < 10^{-3}$ ) indicating similarity of word representation between individuals. (C) Scatter-plot of mean word position in recall sequence versus its recall probability. Correlation coefficient is negative ( $-0.25$ ,  $P < 10^{-3}$ ) (D) Bootstrapped distribution of correlation coefficients between the number of recalled words and: (red) the average recall probability of presented words, (blue) the average recall probability for recalled words. Calculations are based on 9682 trials.





**Figure 5.** Comparison of extended model with data. (A) Serial position curve—fraction of words recalled as a function of position in the study list; (B) Contiguity effect—conditional probability of recalling a word as a function of lag between positions the study list. (C) Average number of words recalled as a function of time to report; (D) Same as C in extended model. Time here is represented by the number of transitions. (E) Average time to recall each subsequent word as a function of its output position. Each curve represents trials with different total number of recalled words (3–16 from left to right); (F) Same as E for extended model.

hard to calculate analytically, but we observed that in the large range of times (between  $10^2$  and  $10^5$  cycles) it can be approximated as a logarithmic function (Fig. 5D, inset). This approximate behavior implies that retrieving each new item requires  $\sim 2.5$  the time already elapsed since the beginning of the trial.

After establishing that the extended model describes classical features of free recall experiments we tested whether the statistical properties of the basic model are retained. To this end, we simulated the recall experiment and compute the bootstrapped correlation coefficients the same way we did with the data and the basic model. Similar to the basic model, the word recall probability is largely defined by the word size (correlation coefficient is 0.62, Fig. 6A). Mean word position is also negatively correlated with word recall probability (correlation coefficient is  $-0.23$ , Fig. 6B). We finally evaluated the statistical interactions between the recall of easy and difficult words in the same way that we did in the basic model above. We again found that the correlation between the number of recalled words and average word recall probability of recalled words was negative and the correlation between the number of recalled words and average word recall probability of presented words was close to zero (Fig. 4, cf. D and C).

According to the extended version of the model, similarities between the presented items are composed of two components. The first component is determined by the intrinsic overlaps between long-term representations of the items and thus stable for

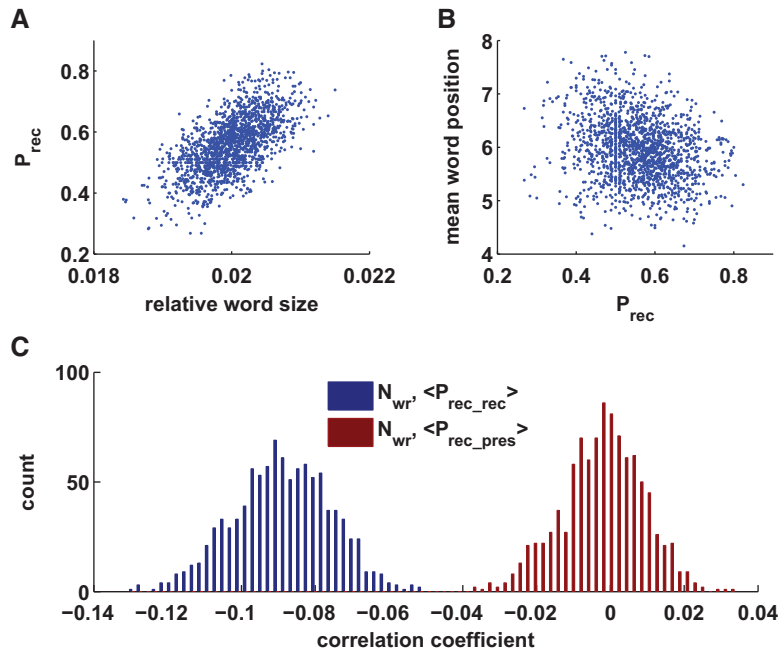
the duration of the experiment, whereas the second component depends on the order in which the items are presented and thus holds for the duration of a single trial. The second component of similarities can be manipulated by the presentation protocol, e.g., presenting some of the words twice, as in the classical list strength effect study (Tulving and Hastie 1972). In accordance with the contiguity mechanism described above, when a word is repeated twice, it acquires additional similarity to two words at the previous positions, as opposed to words that are only presented once. Hence the probability for the repeated word to be recalled is increased even though its long-term neuronal representation is not assumed to change. Following Tulving and Hastie (1972), we simulated the model for lists of 16 words, with half of them repeated twice. We found that the probability to recall a repeated word is increased compared with the control case (lists of 16 words, each presented once), while the probability to recall a word presented once is decreased due to suppression (0.64, 0.56, and 0.39, respectively).

## Discussion

Our retrieval model is based on two assumptions: that items are encoded by random groups of neurons (representations) in dedicated memory networks; and that retrieval of subsequent items from memory is determined by the size of the intersections between the corresponding neuronal representations. Due to its deterministic nature and the randomness of representations, retrieval

process enters into a loop before all of the items in memory are reached. Those assumptions are sufficient to explain the difficulty human subjects encounter when attempting to recall unrelated lists of words. Moreover, the predictions arising from those assumptions are confirmed by the analysis of the experimental data.

Several detailed models of free recall were developed in the psychological literature (Raaijmakers and Shiffrin 1980; Howard and Kahana 2002; Laming 2009; Polyn et al. 2009; Lehman and Malmberg 2013). These models are characterized by precise match to experimental data. Our model shares some of the features with these detailed models, but is much simpler and is explicitly cast in terms of neuronal representations of items in memory. This allowed us to make several predictions regarding the data. More specifically, we introduce the notion of average “recall probability” for a given item determined by the size of its neuronal representation, as measured in free recall experiments with randomly composed lists of unrelated words. Indeed, analysis of the data presented in this contribution showed that different words have different recall probabilities consistently for different groups of subjects (easy versus difficult words). The model also robustly predicts the pattern of recall statistics characterized by specific correlations between the recall of easy and difficult words. Surprisingly, due to the interaction between easy and difficult words, there is no significant correlation between the average recall probability of presented words and how many of them will be recalled.



**Figure 6.** Statistical properties of extended model. (A) Probability of correct responses as a function of word size; (B) mean word position in recalled sequence as function of word-recall probability. (C) Bootstrapped distribution of correlation coefficients between the number of recalled words and (red), the average recall probability of presented words; (blue), the average recall probability for recalled words.

Moreover, there is a negative correlation between the average recall probability of “recalled” words and their number. These predictions are supported by the analysis of the data. A more direct test of the model would be to specifically assemble lists of either most easy or most difficult words—the average number of them recalled should not be significantly different.

One interesting issue not addressed by our model is how subjects avoid reporting words not presented for recall. Analysis showed that the rate of such mistakes (“intrusions”) is low,  $\sim 0.3$  intrusions per trial. Moreover, most of the intrusions involve words that were presented to the subject in one of the previous trials, i.e., may constitute a delayed recall. One possible mechanism for restricting the recall to the current list is tying the neuronal representations of the words together via the representation of a list (“context”), as in the influential Temporal Context Model of Kahana (Howard and Kahana 1999).

Our basic model can be extended to account for classical features of free recall, namely primacy, recency, and contiguity. In the extended model, similarities between the presented items are determined by two components, the first one is constant and determined by the intrinsic overlaps between long-term representations of the items, and the second one is formed at each trial and depends on the order in which the items are presented for recall. The first component is a random one and is chiefly responsible for the fact that only a diminishing fraction of words can be recalled for longer lists, while the second component can potentially improve the recall by overcoming the randomness of the first one. We also showed that a weakly stochastic version of the model results in temporal dynamics of recall that is similar to experimentally measured one.

Taken together, the results presented in this study strongly suggest that associative search for information stored in memory, based on internal representations of memory items, underlies retrieval process in free recall and similar experimental paradigms.

We note that similar process, termed “latching dynamics” was suggested as a mechanism of free association transitions in the context of attractor neural network models of cerebral cortex (Russo et al. 2008). Some of the characteristic features of this mechanism can be captured in simplified recall models that provide a parsimonious explanation for many of the classically observed characteristics of recall. The model establishes the size of long-term neuronal representation of a word as an important feature that controls its recall probability. The number of neurons that encode a word in memory can in principle be measured in neurophysiological experiments (or possibly inferred from EEG/MEG imaging, see e.g., Burke et al. 2013), which would constitute the direct test of model predictions. Another future research question is whether the word’s recall probability can be “predicted” by such observable attributes as its usage frequency, emotional valence, orthographic, and syllabic length etc. In the interesting recent study (Lohnas and Kahana 2013), it was reported that words of high and low usage frequency tend to have higher recall probability compared with words with intermediate frequency. This tendency, even though statistically

significant, is however rather weak compared with the variability in  $P_{rec}$  of words with similar frequency, i.e., the latter is not enough to predict the former on a word by word basis. We found that word length is also significantly correlated with recall probability (Katkov et al. 2014; see also Shulman 1967 for earlier study of recognition memory), but also this effect is weak. It will be interesting to find how well a combination of measurable word attributes can constrain its recall probability, and how it correlates to the size of the word’s long-term neuronal representation.

## Materials and Methods

### Basic model

We assume that each word is represented by a randomly chosen group of neurons in the dedicated memory network of  $N$  neurons (Romani et al. 2013). To mimic the experimental protocol (see below), we generated a large pool of such groups specified by binary patterns  $\{\xi_i^w = 0;1\}$  where  $w = 1, \dots, W$  indicates different words in the pool of size  $W = 1638$  and  $i = 1, \dots, N$  indicates the neurons in the network, such that  $\xi_i^w = 1$  if neuron  $i$  is participating in the encoding of the word  $w$ . The components  $\xi_i^w$  were drawn independently with probability of “1” given by the sparseness parameter  $f$ , and were then fixed throughout the simulated experiment. We then simulated numerous recall trials (matching experiment). At each simulated trial, a set of  $L = 16$  items  $\{w_1, \dots, w_L\}$  was randomly selected for presentation and the set of retrieved items with serial positions  $\{k_1, k_2, \dots, k_r\}$  was determined as follows: the serial position of the first item to be recalled ( $k_1$ ) was chosen randomly, and subsequent transitions between states were determined by the similarity matrix  $S$  between the items, each element of which computed as the number of neurons in the intersection between the corresponding representations:

$$S^{km} = \sum_{i=1}^N \xi_i^{w_k} \xi_i^{w_m}.$$

More specifically, the next retrieved item ( $k_{n+1}$ ) is the one that has the maximal similarity to the currently retrieved item ( $k_n$ ), excluding the item that was retrieved just before the current one ( $k_{n-1}$ ).

### Extended model

$W = 1638$  items were represented by random groups of neurons in the network of size  $N = 10^5$ , specified by binary patterns  $\xi_w (w = 1, \dots, W)$  of length  $N$ . These representations were fixed throughout the simulated experiment. The components of  $\xi_w$  were drawn independently with probability of "1" given by the sparseness parameter  $f$ . At each trial, a set of  $L = 16$  items  $\{w_1, \dots, w_L\}$  was randomly selected for presentation. Retrieval was modeled as a series of states "visited" by the network. Following Romani et al. (2013), transitions between states were determined by the similarity matrix  $S$  between the items (see below). More specifically, if the system visited states corresponding to items  $i_1, i_2, \dots, i_k$  then the next item is  $i_{k+1} \neq i_k$  is chosen with probability  $P(i_{k+1}|i_k) \propto e^{-\delta S_{k+1, i_k}}$ , where parameter  $\delta$  determines the degree of stochasticity. The system was allowed to evolve for predefined time  $T = 3750$  transitions, and a new item was retrieved when the state corresponding to this item appeared in the sequence of states for the first time. At that moment, random uniformly distributed time (from 0 to 75 transitions) was added to the system to model the report of the recalled item.

At each trial, the similarity between the presented items  $w_k$  and  $w_m$  was defined as the sum of three terms,  $O_{km} + P_{km} + C_{km}$ , standing for "overlap," "primacy," and "contiguity" correspondingly, each term defined as following.

$$O_{km} = \sum_{i=1}^N \xi_{w_k}^i \xi_{w_m}^i.$$

The primacy was modeled by increased overlap between first and subsequent items -  $P_{km} = \{\varepsilon \langle \xi_{w_k} \rangle \langle \xi_{w_m} \rangle\} N$  when  $k$  or  $m = 1$ ; 0 otherwise. Contiguity was modeled as increased overlap between an item,  $m$ , and one of the previously presented items  $k$ :  $C_{km} = \varepsilon \langle \xi_{w_k} \rangle \langle \xi_{w_m} \rangle N$ .  $C_{mk} = \gamma C_{km}$ . Here  $k < m$  was drawn from distribution  $P(k) \sim e^{-\gamma \beta (m-k)}$ ,  $\gamma \leq 1$  represents asymmetry of contiguity, and  $1/\beta$  is the average number of previous items available for association. Recency was modeled by selecting initial item,  $w_k$ , in the recall sequence with distribution  $P(k) \sim e^{-\beta(L-k)}$ . Overall, the model can be considered as stripped of SAM model (Raaijmakers and Shiffrin 1980).

### Model parameters

Parameters of the model were chosen the following way. Some parameters after initial exploration were fixed at plausible values that produce good fits to some of the features of experimental data. Parameter  $f = 0.02$  was chosen as it produced the scaling dependence between the average number of recalled words and number of presented words similar to experimental observations (Romani et al. 2013). Parameter  $\gamma = 0.8$  was fixed to produce the ratio between probabilities of recalling following and previous words similar to those obtained from experimental data. The rest of the parameters ( $\beta = 1.06$ ,  $\delta = 0.098$ ,  $\varepsilon = 0.33$ ,  $\vartheta = 0.9$ ) were rounded after fitting to experimental data using simulated annealing. The qualitative behavior of the model does not depend on exact values of the parameters and the annealing procedure was used to find the range of parameters leading to model behavior similar to that observed in the data. For the optimization we used simulated annealing procedure, where each iteration we modeled free recall experiments with the same number of trials, and the cost function consisted of weighted sum of square differences between simulated and observed measures: the overall distribution of  $N_{wr}$  across trials, serial position curve (recall probability versus position of a word in the list) and distribution of recall chains of consecutively presented words.

### Chain length

For each trial, recall was broken into sequences of items either in the presented (positive length) or reverse (negative length) order. Isolated items were assigned chain length 0. For example, a report  $w_1, w_2, w_7, w_5, w_4$ , and  $w_3$  consists of three chains:  $w_1, w_2$ , (length 2);  $w_7$  (length 0);  $w_5, w_4, w_3$  (length -3).

### Experimental methods

The data reported in this manuscript were collected in the laboratory of M. Kahana as part of the Penn Electrophysiology of Encoding and Retrieval Study (see Miller et al. 2012 for details of the experiments). Here we analyzed the results from the 141 participants (age 17–30) who completed the first phase of the experiment, consisting of seven experimental sessions. Participants were consented according to the University of Pennsylvania's IRB protocol and were compensated for their participation. Each session consisted of 16 lists of 16 words presented one at a time on a computer screen and lasted ~1.5 h. Each study list was followed by an immediate free recall test. Words were drawn from a pool of 1638 words. For each list, there was a 1500 msec delay before the first word appeared on the screen. Each item was on the screen for 3000 msec, followed by jittered 800–1200-msec interstimulus interval (uniform distribution). After the last item in the list, there was a 1200–1400-msec jittered delay, after which the participant was given 75 sec to attempt to recall any of the just-presented items. Only trials without errors (no intrusions and no repeated recalls of the same words) were used in the analysis.

### Acknowledgments

We are grateful to M. Kahana for generously sharing the data obtained in his laboratory with us. The laboratory of Kahana is supported by NIH grant MH55687. M.T. is supported by EU FP7 (Grant agreement 604102), Israeli Science Foundation and Foundation Adelis. S.R. is supported by Human Frontier Science Program long-term fellowship.

### References

- Amit DJ. 1989. *Modeling brain function: the world of attractor neural networks*. Cambridge University Press, Cambridge.
- Binet A, Henri V. 1894. La memoire des mots. *L'annee psycholog., Bd. I* **1**: 1–23.
- Burke JF, Zaghoul KA, Jacobs J, Williams RB, Sperling MR, Sharan AD, Kahana MJ. 2013. Synchronous and asynchronous theta and gamma activity during episodic memory formation. *J Neurosci* **33**: 292–304.
- Gelbard-Sagiv H, Mukamel R, Harel M, Malach R, Fried I. 2008. Internally generated reactivation of single neurons in human hippocampus during free recall. *Science* **322**: 96–101.
- Hasselmo ME, Wyble BP. 1997. Free recall and recognition in a network model of the hippocampus: simulating effects of scopolamine on human memory function. *Behav Brain Res* **89**: 1–34.
- Hills TT, Jones MN, Todd PM. 2012. Optimal foraging in semantic memory. *Psychol Rev* **119**: 431–440.
- Hopfield JJ. 1982. Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci* **79**: 2554–2558.
- Howard MW, Kahana MJ. 1999. Contextual variability and serial position effects in free recall. *J Exp Psychol Learn Mem Cogn* **25**: 923–941.
- Howard MW, Kahana MJ. 2002. A distributed representation of temporal context. *J Math Psychol* **463**: 269–299.
- Katkov M, Romani S, Tsodyks M. 2014. Word length effect in free recall of randomly assembled word lists. *Front Comput Neurosci* **8**: 129.
- Laming D. 2009. Failure to recall. *Psychol Rev* **116**: 157–186.
- Lehman M, Malmberg KJ. 2013. A buffer model of memory encoding and temporal correlations in retrieval. *Psychol Rev* **120**: 155–189.
- Lohnas LJ, Kahana MJ. 2013. Parametric effects of word frequency in memory for mixed frequency lists. *J Exp Psychol Learn Mem Cogn* **39**: 1943–1946.
- Miller JF, Weidemann CT, Kahana MJ. 2012. Recall termination in free recall. *Mem Cognit* doi: 10.3758/s13421-011-0178-9.
- Murdock BB. 1960. The immediate retention of unrelated words. *J Exp Psychol* **60**: 222–234.
- Murdock BB. 1962. The serial position effect of free recall. *J Exp Psychol* **64**: 482–488.

- Murdock BB, Okada R. 1970. Interresponse times in single-trial free recall. *J Exp Psychol* **86**: 263–267.
- Murray DJ. 1975. Graphemically cued retrieval of words from long-term memory. *J Exp Psychol Hum Learn* **1**: 65.
- Murray DJ, Pye C, Hockley WE. 1976. Standing's power function in long-term memory. *Psychol Res* **38**: 319–331.
- Polyn SM, Norman KA, Kahana MJ. 2009. A context maintenance and retrieval model of organizational processes in free recall. *Psychol Rev* **116**: 129–156.
- Raaijmakers JGW, Shiffrin RM. 1980. SAM: a theory of probabilistic search of associative memory. *Psychol Learn Motiv Adv Res Theory* **14**: 207–262.
- Roberts WA. 1972. Free recall of word lists varying in length and rate of presentation: a test of total-time hypotheses. *J Exp Psychol* **92**: 365–372.
- Romani S, Pinkoviezky I, Rubin A, Tsodyks M. 2013. Scaling laws of associative memory retrieval. *Neural comput* **25**: 2523–2544.
- Russo E, Namboodiri VM, Treves A, Kropff E. 2008. Free association transitions in models of cortical latching dynamics. *New J Phys* **10**: 015008.
- Shulman AI. 1967. Word length and rarity in recognition memory. *Psychon Sci* **9**: 211–212.
- Standing L. 1973. Learning 10,000 pictures. *Q J Exp Psychol* **25**: 207–222.
- Tulving E, Hastie R. 1972. Inhibition effects of intralist repetition in free recall. *J Exp Psychol* **93**: 297–304.
- Tulving E, Pearlstone Z. 1966. Availability versus accessibility of information in memory for words. *J Verb Learn Verb Behav* **5**: 381–391.

*Received April 1, 2014; accepted in revised form November 11, 2014.*