Random-walk modeling of free-recall anomalies
Part 1: Contiguity and inter-response intervals

Francesco Fumarola

(Dated: December 13, 2016)

A comparison of three different experimental datasets leads to the identification of an apparently robust feature of free-recall data: a minimum in the rate of contiguous recall near the beginning of the recall process. By simulating verbal retrieval as a homogenous Markov chain with the experimentally observed transition matrix, I show that this behavior cannot be explained in terms of the nonequilibrium initial conditions. I deduce that the process, if markovian, is inhomogenous and I identify a psychological mechanism that would effectively lead to a time-dependent transition matrix. The theory is based on the simultaneous coexistence of competing random walks, all of them homogenous but liable to losing permanently the right to effect retrieval. To test this hypothesis, I construct a many-particle model governed by two free parameters and simulate it with initial conditions taken from the data. Besides reproducing the contiguous-recall curve, the theory leads to a counterintuitive prediction on the behavior of the inter-response intervals as a function of the serial-position lag. Data are found to confirm this prediction. I proceed then to investigate possible consequences of the model, drawing in part on a comparison with classical results from experiments on category recall.

I. REMINISCING-PARTICLE MODELS

Science knows few experiments as simple as free recall. Participants in a free-recall trial are simply presented with a list of words and asked to recall them in any order (Murdock, 1960, 1962; Roberts, 1972; Standing, 1973). Such experiments have been performed systematically for over a century (Binet and Henri, 1894), yet some of their results still defy comprehension (for a review, see Kahana, 2012).

The words are typically random and the input primarily characterized by one parameter: the size $S$ of the lists. As for the subject’s response, a key observable is the “serial-position
lag”, i.e. the difference between the positions of two consecutively recalled words. For example, if the 5th word in the list is remembered right after the 8th, the corresponding lag is $L = 3$. Transitions with lag $L = 1$ are called contiguous. Another instructive observable is the distribution of the inter-response intervals, the time intervals that elapse between two consecutive recalls (Murdock and Okada, 1970; Rohrer and Wixted, 1994).

One standard approach to modeling these experiments is to represent the recall process as the stochastic motion of what we may term a “psychion”, a particle representing the evolution of thought. The space in which this process evolves may be a highly connected graph, as in Romani et al. (2013), or a semantic network (as in Abbot et al, 2015, Kenett et al, 2015), or a lattice (as in Fumarola, 2016), or finally a continuous linear space, as in the subtler model of Howard and Kanaha (2002). The psychion moves continuously, but stochastically, in the region where memories are located, and retrieves all the memories it encounters.

The most compelling support for psychion models comes from a phenomenon known as the lag-recency effect (Howard and Kahana, 1999): words that are close to each other within the list are more likely to be recalled at short distance from each other. This can be taken to signify that the motion of the psychion is continuous. Indeed, suppose that the memory corresponding to each word in the list is created in the neighborhood of the memory created by the previous word; during recall, if the psychion random-walks continuously, contiguous memories will have a higher probability of being encountered contiguously (Howard and Kahana, 2002, 2002b; Polyn, Norman, and Kahana, 2009a, 2009b; Lohnas et al., 2015).

While the continuity of its motion has solid experimental grounding, it is less clear whether the psychion should move in accordance with a time-homogenous process. In this paper, I will point out a robust feature of the data that can in no way be reconciled with the time-homogeneous assumption – namely, the non-monotonous evolution of contiguous recall.

In the next section, I show that simulating the retrieval process through its observed time-averaged transition matrix leads to a systematically incorrect prediction for this feature. In section III, I sketch a possible mechanism that may lie behind the apparent nonhomogeneity. In section IV, I articulate this mechanism into a two-parameter model of free recall and test it numerically. In section V, I report an additional consequence of the model and investigate it through the same sets of data. In section VI, I consider three possible ramifications of the
theory, drawing on a comparison with theoretical and experimental results available in the literature.

II. RATE OF CONTIGUOUS RECALL

I will consider three sets of data. The first dataset comes from experiment I of PEERS (Penn Electrophysiology of Encoding and Retrieval Study). In each experimental trial, 16 words were presented one at a time on a computer screen. Each item was kept on the screen for 3000 ms, followed by an interstimulus interval of 800–1200 ms. Participants were given 75 s to attempt to recall aloud any of the just-presented items. The experiment involved 194 participants and multiple trials were performed on each subject, summing up to 4656 trials. For more details on the experimental procedure, see Healey and Kahana (2016) and Lohnas and Kahana (2013).

The second dataset comes from experiment III of PEERS. It consists of 7360 trials performed on 156 participants in experimental conditions similar to experiment I, and was reported on in Lohnas et al., 2015. The third dataset comes from the experiments described in Polyn et al., 2009a. These experiments involved 45 participants; the lists contained 24 words each, again presented on a computer screen. Each word was shown for 3 seconds; in the retrieval stage, participants were given 90 seconds to recall the words. A total of 1394 trials were performed.

Three differences may be noticed between the experimental conditions of PEERS and those employed by Polyn et al.: (1) the length of the lists, which is 16 in PEERS and 24 in Polyn et al.; (2) the longer time allowed by Polyn et al. for both memorization and recall; (3) the different size of the word-pools, as PEERS experiments used a pool of 1638 words, while the pool of Polyn et al. contained only 1297.

If we start with the assumption that the recall process is a homogenous Markov process, it should be possible to obtain the transition matrix by averaging over all transitions in the recall process. Doing so yields indeed similar matrices in all three experiments. The general properties of this time-averaged transition matrix are shown on the left-hand panel of Fig. 1., with the Polyn et al. matrix as an example. The diagonal is null because immediate repetitions have not been considered. The super- and sub- diagonal elements are the most prominent terms. The transition probability decays quickly as we move toward the off-
FIG. 1: Left-hand panel: probability of first recall as a function of the normalized serial position, for three sets of experiments. Right-hand panel: time-averaged transition matrix for the experiments of Polyn et al., 2009a. The label Ω indicates transitions to the sink state.

diagonal corners. Transitions to the sink state are encoded in a column at the right end of the matrix.

The general distribution of initial conditions for this process (that is, the distribution of the first word to be recalled) has been studied extensively (Murdock 1960, 1962; Bjork and Whitten, 1974). The form it takes in these datasets is shown in the right-hand panel of Fig. 1, plotted for all three sets of experiments as a function of the serial position $x$ normalized by the length $S$ of the lists.

In agreement with the literature, the highest concentration of probability is found at the extremal points – the beginning and the end of the list. In addition, the end-list concentration is more marked for the longer lists of Poly et al. This is easily understood: as the lists of Healey and Kahana are shorter, participants had less time to forget the initial words of the list by the time the recall process began.

Using these three distributions as initial conditions, and the time-averaged transition matrices obtained from each of the datasets, we can fully simulate the resulting Markov process. In particular, we would like to compute the probability of contiguous recall as a function of time. The result is shown in Fig. 2 for all three datasets. The curves resulting from simulations are in red, the experimental curves in blue.

Both simulations and experiments show a large variability in the contiguous-recall probability at the very beginning of the recall process. This is due, in simulations, to the fact that
FIG. 2: Rate of contiguous recall as a function of the order of recall for the three experimental datasets (blue) and for a Markov chain on the list of words, with both the initial conditions and the time-averaged transition matrix taken from the experiments (red).

the initial distribution is distant from the equilibrium state. Unlike the simulated curve, however, the experimental curve experiences a minimum at a recall position $T^*$, such that $T^* \sim 5$ for 16-word lists, and $T^* \sim 7.5$ for 24-word lists. The contiguous recall rate then increases steadily throughout the observed duration of the recall process, while the simulated curve stays close to a stationary value. For very large times, the experimental curves are no longer reliable as too few experimental points are available.

We conclude that the homogeneous Markov theory does not explain the experimental behavior. To understand the possible reasons for this discrepancy, it is useful to derive similar conclusions from a diffusive psychion model, also time-homogeneous, of the class we described above.

For simplicity, let the psychion live in a space of binary arrays, where each site is a string of 0’s and 1’s. The creation of word-memories can then be modeled as the stochastic process that, starting with a given binary array, creates new memories by progressively flipping a random digit. The constraint that no state should correspond to two memories will be enforced, although in sufficiently high dimensions it is mostly unnecessary.

At the beginning of the retrieval trajectory, the psychion is made to lie at the location of one of the memories. We may pick this memory at random according to the distribution of initial words found in the experimental data. The retrieval process evolves by flipping digits in the binary array that describes the psychion’s location. Whenever the location coincide
with a memory, that memory is retrieved.

This is admittedly a crude model (e.g., no mechanism for recall termination has been introduced), yet it predicts the main DC properties of the dynamics. The time-averaged transition matrix obtained numerically from the psychion model is shown, representatively, for the initial conditions of Polyn et al. in the left-hand panel of figure (3). The behavior near the diagonal, as can be seen, is close to the behavior reported in Fig. 1.

We may expect that, if the process is homogenous in time, this microscopic model could also offer an approximation to its relaxation to equilibrium. To ascertain this, we refer again to the time-evolution of the contiguous recall probability. This has been plotted, for all three sets of initial conditions, in the right-hand panel of Figure 3. Note that the variable $T$ in the x-axis represents the order of recall in the verbal process, as in Fig. 2, and not the time variable over which the psychon’s motion develops.

From Fig. 3, we see that the contiguous recall probability behaves similarly in all three cases. It starts by increasing, just as with the shorter lists in the Markov chain model (red curves of Fig. 2), then quickly serves down, and its downward trend slackens only near the end of the recall process. The difference from the experimental behavior (blue curves of Fig. 2) could not be more drastic.

While the basic psychion model provides a fair approximation to time-averaged properties of the recall process, we conclude that it cannot describe the dynamics adequately. In the next section, I will indicate a possible mechanism from which this discrepancy may arise and will propose a modified psychion model to account for the observed behavior.
FIG. 3: Results from simulations of the psychion model. Left-hand panel: time-averaged transition matrix with the initial conditions of Polyn et al., 2009a. Right-hand panel: probability of contiguous recall as a function of the normalized order of recall, for initial conditions taken from the three sets of experiments.

III. A RANDOM WALK ON RANDOM WALKS

We will consider the possibility that the system may not be employing a single psychion, but a number of psychions proportional to the magnitude of the recall task, that is, to the length $S$ of the list. Accordingly, there will be two time scales: a microscopic time scale, over which each of $N$ psychions random-walks independently; and a longer, macroscopic time scale, over which the system random-hops over psychion trajectories.

I will focus here on the simplest possible form of such a “hierarchial” random walk, where only the psychion currently being visited by the system is allowed to retrieve memories (Fig. 4). Psychions that the macroscopic random walk has not yet visited will be called “unexplored”. Let $n_t$ be their number at time $t$; at every time step, the system has a finite probability $q(n_t)$ of hopping onto any of the unexplored psychions, and a probability $1 - n_t q(n_t)$ of staying with the current one. Once a psychion has been abandoned, it is no longer be taken under consideration.

Results from simulations of this model are shown in Fig. 5, for the simple choice $q(n) = \frac{1}{n+1}$. At the beginning of each simulation, each psychion has been placed at the sites of some word-memory, chosen in accordance with the distributions in the left-hand panel of Fig. 1. Notice that, while the model of the previous section had a single free parameter
FIG. 4: Example of dynamics in the many-particle model. Left-hand panel: random-walk on the psychion graph, with corresponding recall events. In the example, only word B is recalled associatively. The rest are located by switching among psychions. Right-hand panel: cartoon depiction of psychionic space. Word-memories are represented as gray cells; stretches of the trajectory during which a psychion is endowed with retrieval privileges are shown as solid lines, the rest as dotted lines.

– the dimensionality $d$ of psychionic space – this many-particle variant includes a second controlling parameter: the psychion-per-word ratio $\kappa$, defined by $N = \kappa S$. The results in Fig. 5 were obtained in seven dimensions for $\kappa = 9$.

The left-hand panel of the figure shows the time-averaged transition matrix obtained from the initial conditions of Polyn et al., 2009a. The process of switching among psychions has visibly smoothened the sharp dominance of contiguous transitions, without altering the overall structure of the DC modes.

In the right-hand panel, the probability of contiguous recall has been plotted as a function of time for the three initial distributions. The same behavior emerges for all sets of experimental conditions: the contiguous recall probability decreases sharply at small times, experiences a minimum shortly after the beginning of the recall process, then increases steadily, all the way to the end of the process. This is precisely what we found in experi-
FIG. 5: Simulation of the memory search through a many-particle model. Left-hand panel: time-averaged transition matrix. Right-hand panel: probability of contiguous recall as function of time. The results displayed correspond to $d = 7, \kappa = 9$.

mental data (see Fig. 2).

We may now explain quite logically the experimental behavior, by reasoning in terms of the many-particle model.

Once the system has located a memory, there are two ways it can locate the next: by hopping onto another psychion that is going to locate it, or by continuing to follow the same psychion that located the first memory. I will refer to the former mechanism as “zapping”, to the latter as free association. At the beginning of the process, zapping dominates: the system is testing the possibilities offered by the many trajectories available. Once most psychions have been explored, the system focuses on the few left, and is forced to make a greater use of associative recall. By construction, zapping is indifferent to serial position. Associative recall, on the contrary, is affected by the lag-recency effect, which favors contiguous memories. Therefore, the probability of contiguous recall will begin to grow.

This interpretation is confirmed by Fig. 6, where the associative component and the zapping component of the contiguous recall probability are shown separately for one of the simulations. The only observable quantity in the figure is the average probability of contiguous recall, shown as a solid curve, which lies between the two dashed curves corresponding to the zapping and associative components. The smaller panel to the right shows the zapping frequency as a function of time; once again, the variable $T$ in the x-axes does not represent
time proper, but the order of recall.

At small times, zapping predominates. The probability of contiguous transitions coincides therefore with the probability of contiguity among zapping transitions. Since the initial condition is concentrated near few definite points (the initial and final memories), zapping transitions possess initially a large degree of contiguity. This ceases to be true as the packet of psychions spreads away from its initial position. Zapping transitions cease to have a considerable chance of being contiguous, and since they are still the dominant type of transitions, the overall probability of contiguous recall goes down.

Gradually, the zapping frequency decreases and the frequency of associative transitions increases. As a consequence, the average probability of contiguous recall coincides less and less with the probability of contiguous transitions in the zapping process, and migrates toward the associative component.

The lag-recency effect enters now into play. Through its action, as can be seen in Fig. 6, the curve of contiguous recall for associative transitions lies always above the curve of contiguous recall for zapping transitions. As associative recall comes to dominate the dynamics, the probability of contiguous transitions therefore increases. By the end of the process it virtually coincides with the probability of contiguous transitions in associative recall.

The hierarchical structure of the search for memories leaves its telltale mark in the presence of two different time scales. The initial drop in contiguous recall occurs over a time scale that depends on the dimensionality of psychionic space and on the distribution of initial conditions. This is a microscopic time scale, independent on the size of the psychion population.

The subsequent increase in contiguous recall, on the other hand, happens on a macroscopic time scale, that could not be obtained from studying individual psychion dynamics. This is the time scale over which the reservoir of unexplored psychions is gradually depleted, and its value is controlled by the parameter $\kappa$. We learn from experiments that $\kappa$ is large enough to allow for the depletion to stretch throughout the recall process, but low enough to let the reservoir be substantially depleted.

Since the macroscopic random-walk is self-avoiding, the many-psychion model is non-markovian. However, it is possible to translate it into a nonhomogenous Markov process as follows. Call $M_0(x, y) = \frac{1}{d} \delta_{|x-y|, 1}$ the elements of the transition matrix for the one-psychion model in the space of binary arrays of length $d$. Since we can account
FIG. 6: Decomposition of the contiguous recall probability in the many-particle model, with the initial conditions of Healey and Kanaha, 2016. The dashed curves are the rate of contiguity among associative and zapping transitions; the solid curve is the rate of contiguous transitions in the full recall process. The right-hand panel shows the time-evolution of the zapping frequency $P_{\text{zap}}$. Results correspond to $d = 7$, $\kappa = 9$.

separately for the randomness of the motion and of the initial conditions, the position of the psychion currently entitled to retrieval evolves according to the transition matrix

$\hat{M}(t) = (1 - \alpha_t) \hat{M}_0 + \alpha_t \hat{M}_{0}^{t+1} e_y n^T$, where $n$ is a constant function equal to unity and $e_y$ is a function worth unity at position $y$ and zero elsewhere; the value of $y_t$ should be picked anew with a probability distribution $\pi_0$ at every time step $t \geq 0$, $e_y_0$ being the initial condition of the process; and the probability distribution of the infinite binary array $\alpha_t$ (defined through the function $q$) constrains it to sum up to $\int_0^\infty dt \alpha_t = \kappa S - 1$.

The functions $\alpha_t$ and $y_t$ are statistically independent, and the only nonmarkovian correlations are those contained in $\alpha_t$. For each given choice of the function $\alpha_t$, and of the initial conditions, the resulting process is a nonhomogeneous Markov chain. Consequently, we can regard it as describing the dynamics of a single collective psychion moving in the random field created by the upper level of the hierarchical search. I will say more on this interpretation in section V.

The rate of contiguous recall has proven to be a highly informative variable. However, it is unlikely to be the only giveaway of many-particle dynamics. From what we have said thus far, we would particularly suppose that some traces of psychionic zapping may be detectable.
in the distribution of the inter-response intervals (IRIs).

The argument is simple: in the many-particle scenario, there are two different ways to perform fast transitions between words during the recall process, by exploiting lag-recency or by exploiting zapping. The same mechanism that creates the lag-recency effect can make associative recall faster, but only if the words involved lie nearby within the list. Hence, fast transitions due to association may only occur for small serial-position lags.

Zapping, on the other hand, can provide fast transition with any lags. Going from a psychion that has found a memory to another psychion that has just found another memory is equally likely regardless of the locations of the memories, and may produce fast transitions between memories located at any lag from each other. Only for very long lags, however, zapping encounters no rival mechanism, because associative recall performs those transitions too rarely to provide any competition.

Therefore, two different sets of fast transitions should exist, one from association, at short lags, and one from zapping, at long lags.

In the next section, I will inspect available data on the IRIs in search of traces of this phenomenon.

IV. TRACES OF ZAPPING IN THE INTER-RESPONSE INTERVALS

Data on the IRIs are not available for the experiments of Lohnas et al. 2105 examined above. I will analyze therefore the two sets of data from Healey and Kahana, 2016 and Polyn et al., 2009a.

Results from the experiment of Healey and Kahana are plotted in Fig. 7. The behavior of the distribution of the inter-response interval $\tau$ is presented in the left-hand panel. Lists are 16 words long, so the maximum lag is 15. For ease of visualization, a few representative values of the lag are shown (see legend).

The presence of a peak at small $\tau$ for transitions with a lag $L = +1$ is no surprise, and agrees with expectations from the lag-recency effect. For longer lags, the peak is suppressed, as seen in the curves corresponding to lags $L = 2, 3, 8$. However, this monotonous suppression has ceased by $L = 12$: the peak has begun to grow again. With the maximal lag $L = 15$, we have recovered a peak as tall as the peak for $L = 3$. This means that the recall process, slow for transitions to intermediate lags, is faster again for transitions to very long lags.
The right-hand panel of Fig. 7 shows the cumulative of the distribution, easier to visualize, for all values of the lag. The growth of the cumulative toward saturation becomes gradually slower as the lag increases from $L = +1$. It reaches its slowest point for $L^* = 10$ (yellow curve); then, it gradually grows faster again. The inversion point is the lag $L^* \sim 0.6S$, where $S$ is the size of the lists. For the maximal lag $L = 15$, the cumulative curve virtually overlaps with the curve for $L = 3$.

Data from the experiments of Polyn et al., 2009a, are similarly displayed in Fig. 8. The left-hand panel shows the distribution of IRIs for a handful of lag values. The behavior of the distribution confirms what we observed in the data of Healey and Kanaha: a suppression of the peak for intermediate lags values, and a new enhancement for near-maximal values. In fact, we find here that the peak for the maximal lag ($L = 23$) is taller than the peak for contiguous transitions ($L = 1$).

These findings are better gauged by plotting again the cumulative distribution, shown in the right-hand panel of Fig. 8 for all lags. It can be seen that the highest suppression of the peak occurs for $L^* = 14$, so we have once again $L^* \sim 0.6S$, where $S$ is the size of the lists. For the maximal lag $L = 23$, the cumulative curve reaches saturation faster than the curve for contiguous transitions.

To better compare these sets of data, I have averaged the IRIs corresponding to all transitions with the same value of the lag. In Fig. 9, the result is plotted as a function of
the ratio between the lag and the size of the lists.

Participants in the experiment of Polyn et al. were accorded a longer time for the memory test. Perhaps, this is why the curve of the mean IRI for Polyn et al. lies above the curve for the data of Healey and Kahana. In spite of this, the two curves follow the same overall pattern.

For short lags, a growth in the mean IRI is observed. This requires no explanation, as we expect the thought process to move with greater speed between memories created within a shorter time from one another. As the lags grow, so does the time it takes to go from one memory to the next. This increasing trend, however, slows down in mid-range, and ceases altogether at values of $L/S$ in the range between 0.5 and 0.7. Then the average IRI begins to decrease as the lag grows.

On the outside of any specific model of retrieval, this decrease is deeply counterintuitive. It seems to suggest that if two events take place within a shorter time interval they will be recalled further apart from each other, for a certain range of time intervals. The Polyn et al. data are particularly surprising, because the curve reaches further down for maximal lags than for minimal lags. A memory situated 23 memories away from the most recently visited memory is recalled faster, on average, than a memory contiguous to the last one we recalled.

To obtain a quantitative understanding of this finding, I employ simulations of the many-
FIG. 9: Mean inter-response interval for transitions with a given lag, computed from the data in Healey and Kahana, 2016 (blue curve) and in Polyn et al., 2009a (red curve) and plotted as a function of the ratio between the lag and the length of the lists.

psychion model with the same parameters as above (seven dimensions, $\kappa = 9$) so as to extract: (1) the mean IRI; (2) the mean IRI among associative transitions; (3) the mean IRI for zapping transitions; (4) the zapping probability. All this is plotted in Fig. 10 as a function of the lag, for both sets of initial conditions. The resulting scenario appears identical with the two initial conditions, and sheds some light on the possible meaning of Fig. 9.

The time analog of the lag-recency effect speeds up associative transitions only for very short lags; in this range, therefore, the mean associative IRI (shown in blue) grows quickly as a function of the lag.

For such short lags, the zapping probability (shown in green) is considerably smaller than unity, and associative transitions play a crucial role. Therefore the average IRI (the only observable quantity, shown in black) reflects strongly the increase of the associative IRI – the more so as the zapping IRI (shown in red) is nearly constant. This results in a growth of the mean IRI, as seen in the experiments.

As the lags increases further, the zapping probability experiences a steep growth and associative transitions become rarer. The mean IRI shifts therefore toward the value of the IRI for zapping transitions. But the zapping IRI lies always beneath the associative IRI (in
FIG. 10: Mean inter-response interval $\tau$ for transitions with lag $L$: restricted to associative transitions ($\tau_{\text{ass}}$, blue), restricted to zapping transitions ($\tau_{\text{zap}}$, red), and unrestricted ($\overline{\tau}$, black). The zapping probability $P_{\text{zap}}$ is shown in green. The simulations were performed in 7 dimensions with $\kappa = 9$. Initial conditions were taken from the experiments mentioned in the headings.

the whole range of lags); therefore for long lags the mean overall IRI decreases.

As an outcome, the behavior of the average IRI as a function of the lag corresponds to the black curves of Fig. 10 – nonmonotonous, increasing for short lags and decreasing for long lags – thus reproducing the experimental results of Fig. 9.

The initial conditions (shown in the left-hand panel of Fig.1) produce a further effect clearly discernible in Fig. 10, which adds slightly to the speeding up of transitions with near-maximal lags. Since the psychion packet is distributed mostly on extremal word memories (near the end and the beginning of the list), zapping transitions with near-maximal lags are “prepared” by the initial condition, and happen without much searching at the very beginning of the recall process. That makes them faster, so the zapping IRI decreases for very long lags.

As we motivated in section II, the concentration of psychions on early word-memories is greater for shorter lists, i.e. in the experiments of Healey and Kahana. Accordingly, the decrease of the zapping IRI at long lags is more sustained for those initial conditions.

A final note: the indentation in the curves of the associative IRI (and, to a lesser extent, in the other curves) does not come from numerical inaccuracy; it is a corollary of the not-so-high dimensionality used for simulations. In the $d$-dimensional space of binary arrays, any two
memories can be at most \( d \) flips apart. If the lag between them is odd, and the dimension not very large, there is a non-negligible chance that their distance will be unity, which makes them easier to recall associatively. This indentation, however, gradually disappears in higher dimensions and has no psychological significance.

V. CONSIDERATIONS ON THE MODEL

A. Functionality of the many-particle search

The qualitative agreement of Fig. 2 with Fig. 5, and of Fig. 9 with Fig 10, is a strong clue that the memory search in free recall is indeed the performance of multiple thought-particles diffusing simultaneously. Nonetheless, the two-parameter model I have employed should not be taken as a realistic portrait of the actual search mechanism. It is only designed to grasp one nontrivial aspect of the psychology involved, the gradual discarding of psychions that causes a dwindling of the zapping phase.

To infer in what ways the real search differs from the basic model I have proposed, a principle of functionality can be invoked. If we accept the above demonstrations that psychionic resources are being deployed, the actual search mechanism must be designed to utilize them efficiently. For instance, psychions may only be discarded if they have proven inefficient for a certain number of steps. Or, when selecting which psychions to visit next, the system may use some approximate knowledge on the current findings of the psychions: e.g., only particles that have just located a memory may compete for retrieval rights.

These more realistic models turn out to require far fewer psychions and far less time to perform retrieval (compare Alon et al., 2008); simulations suggest, however, that the effects I described in this paper are not suppressed.

For these functionalized models, the reformulation of the many-psychion theory as the motion of a single collective psychion, sketched in section III, will no longer be exact. If the macroscopic random-walk uses feedback on the psychion trajectories, the macroscopic and microscopic motions do not decouple. However, the reformulation will still be valid at a mean-field level, with the hopping probability \( \alpha_t \) now solved for self-consistently.

This approach will be further developed elsewhere; let us only notice that, for the class of models we are considering, the mean-field theory holds a profound psychological significance.
It describes the viewpoint of an observer who interacts with the psychion gas from the outside, unawares of its multiplicity.

**B. Corrections to the scaling law**

Let \( R \) be the number of words successfully recalled by the participant in a free-recall experiment with lists of length \( S \). It has long been known that the relationship between \( R \) and \( S \) is a power law, \( R(S) \sim S^\eta \), sometimes referred to as Standing’s power function (Standing, 1973; Murray et al., 1976).

In the data I have used, it appears that repetitions occur mostly toward the end of the recall process. Therefore, a natural candidate for the mechanism of recall termination is the self-intersection of the psychological trajectory. But this assumption leads, in the first approximation, to a scaling exponent \( \eta = 1/2 \), considerably higher than the typical experimental value. In Standing’s original work, it was found that \( \eta_{Standing} = 0.096 \).

Suppose, however, that the recall consists of an initial zapping stage that retrieves \( R_{zap} \) words, followed by an associative stage that retrieves \( R_{ass} \) words, with \( R = R_{zap} + R_{ass} \). The words yet to be retrieved in the associative stage will be \( S - R_{zap} \), so we may write \( R(S) \sim R_{zap}(S) + (S - R_{zap}(S))^{\eta_{ass}} \). By developing this formula, we may obtain in principle some corrections to the scaling law dependent on the asymptotics of \( R_{zap}(S) \).

Given any database on experiments with lists of multiple lengths, extracting \( R_{ass} \) and \( R_{zap} \) as functions of the lists’ length \( S \) would be feasible through an analysis of the inter-response intervals.

Zapping and associative transitions group into different sectors of the IIRIs distribution, as shown above. Counting them and extracting the respective scaling laws may lead to substantial progress in our theoretical understanding of the scaling law.

**C. Gruenewald-Lockhead experiment**

In the above simulations, we have taken the initial distribution of the psychion packet from experimental data. By doing so, we have evaded questions on how such a packet is formed, as our aim was to study the trajectory of psychions and not their triggering mechanism. However, some insight into their formation may be readily obtained through a
comparison with a different type of experiments.

In 1980, Gruenewald and Lockhead performed a systematic measurement of inter-
response intervals in category recall. In this kind of experiments, subjects are requested
to recall the largest number possible of elements of a category, e.g., to name as many ani-
mals as they can in 15 minutes.

Gruenewald and Lockhead established empirically that, during the recall of category
examples, the “reminiscing particle” will always travel faster between two psychological
states if they are semantically closer (Gruenewald and Lockhead, 1980). This finding has
been recently replicated and extended in Vonberg et al. (2014), and Montez et al. (2015).

In our framework, this means that the retrieval of category items is routinely performed
through a single psychion, at least in the absence of cues, reminders, or other alternatives
to presentation. Therefore, the multiplication of psychions in free-recall experiments is due
solely to the recency of the external input.

A simple scenario ensues. Psychions are created by any input at a location that represents
it, and immediately begin to random walk in psychological space. After some time, like all
diffusive particles, they are no longer likely to be found at their original location: this is
how forgetting will occur in this description of episodic memory.

Since the distribution of psychions over newly formed memories is not uniform, an optimi-
ization principle must guide their genesis. In the case of a word list, we find empirically
(Fig. 1) that concentrating psychions at the earliest and latest memories is the optimal
strategy.

It may be conjectured that the concentration of psychions should be proportional to the
information rate of the input, because that will smoothen the difficulty in retrieving high-
information memories. This yields a possible explanation for the “primacy” effect (Murdock,
1960) as due to the sharp transition from zero to nonzero information flow.

The Gruenewald-Lockhead experiment also confirms that, as argued in the Introduction,
the individual motion of a single psychion is continuous. In particular, it will be impossible
to create ex nihilo a psychion with a random position. This still allows for two mechanisms of
psychion evolution to take place: the splitting of an existing psychion, so that two resulting
psychions will continue their random walk independently, and the merging of two psychions
into one.

These two mechanisms, compared to psychion creation through external input, are far
more difficult to measure, at least through recall experiments. Finding viable ways to test their existence would be an important experimental direction.

VI. CONCLUSIONS

This paper is the first installment in a theoretical analysis of certain anomalies discovered in free-recall data. Datasets from Healey and Kahana (2016), Lohnas et al. (2015), and Polyn et al. (2009) have been used to provide evidence that the rate of contiguous recall experiences a minimum near the beginning of the recall process. This minimum appears to be robust, and its position an increasing function of the list’s size.

I have begun by approaching this problem in the most basic way, through time-homogeneous markovian modeling, first on a graph composed by words of the list, then through a one-parameter model describing retrieval as the motion of a random-walking particle (“psychion”) in psychological space. This approach led to an adequate understanding of the time-averaged transition matrix, but not of the nonmonotonous evolution of the contiguous recall probability.

I have then considered a two-parameter model where multiple psychions are allowed to explore psychological space simultaneously. The right to effect retrieval is passed around among psychions and can be retained for any length of time, but no psychion is granted retrieval rights twice. This alters drastically the curve of contiguous recall, producing a behavior compatible with experiments. In the early stages of the process, transitions between different psychions dominate. Later, associative retrieval through single-psychion trajectories is prevalent and contiguous recall intensifies accordingly.

I have then examined the distribution of inter-response intervals (IRIs) with an eye to detecting traces of transitions between different psychions. The relationship between IRIs and lags has proven nonmonotonous, and its main features independent on the size of the lists. A peak at low durations can be found at short lags and a second peak at long lags, while for intermediate lags the peaks tend to be suppressed.

Simulations of the two-parameter model reproduce this curve correctly, and allow to describe its possible meaning. The first peak comes from the enhancement of associative retrieval at low lags; the second peak from the shorter time interval required by transitions between memories retrieved by two different psychions.
I have then examined briefly three general problems arising from the model: its functionality, its impact on the scaling law of free recall, and the behavior of psychions beyond mere free motion. Results from classical experiments on category recall have been used to understand better the triggering mechanism, an optimization principle has been proposed concerning the number of psychions created during presentation, and some potential properties of psychionic motion have been predicted from the principle of continuity.

If the multiplicity of psychions should find any corroboration in further studies, it would be important to understand the connection between neurons and psychions. A single psychion has already been shown (Romani et al., 2013) to represent adequately the dynamics of a simple spiking network during associative retrieval. If psychions are indeed excitations of complex neural systems, it would be useful to understand (1) whether these excitations are local or global, (2) how their time scale emerges from the microscopic time scales of the network, (3) what class of networks possesses a population of modes that cooperate the way psychions do under the appropriate input.

I would like to thank Michael J. Kahana, of the University of Pennsylvania, for encouraging correspondence and for providing the experimental data that motivated this study.

VII. BIBLIOGRAPHY

Alon N., Avin C., Kouchy M., Kozma G., Lotker Z., Tuttle M. (2008). Many random walks are faster than one. In: Proc. of SPAA 2008, 119-128.

Binet A. and Henri. V. (1894). La memoire des mots. L’année psychologique, Bd. I, 1:1-23.

Bjork R. A. and Whitten W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. Cognitive Psychology 6(2):173-189.

Fumarola F. (2016). Verbal perception and the word length effect. Preprint, arXiv:1607.05530.

Gruenewald P.J. and Lockhead G.R. (1980). The free recall of category examples. Journal of Experimental Psychology: Human Learning, 6:225-240.

Healey M. K. and Kahana M. J. (2016). A four-component model of age-related memory change. Psychological Review, 123(1):23-69.

Howard M. W. and Kahana M. J. (2002). A distributed representation of temporal
context. Journal of Mathematical Psychology, 46:269-299.

Howard M. W. and Kahana M. J. (2002b). When does semantic similarity help episodic retrieval? Journal of Memory and Language, 46:85-98.

Howard M. W. and Kahana M. J. (1999). Contextual variability and serial position effects in free recall. Journal of Experimental Psychology: Learning, Memory, and Cognition, 25:923-941.

Howard M. W. and Kahana M. J. (2002). A distributed representation of temporal context. Journal of Mathematical Psychology, 46:269-299

Kahana M. J. (2012). Foundations of Human Memory, Oxford University Press.

Lohnas L. J. and Kahana M. J. (2013). Parametric effects of word frequency effect in memory for mixed frequency lists. Journal of Experimental Psychology: Learning, Memory, and Cognition, 39:1943-1946.

Lohnas L.J., Polyn S.M., and Kahana M.J. (2015) Expanding the scope of memory search: Modeling intralist and intralist effects in free recall, Psychological Revier, 122 (2):337-363.

Montez P., Thompson G., and Kello C.T. (2015). The Role of Semantic Clustering in Optimal Memory Foraging, Cognitive Science: A Multidisciplinary Journal, 39(8):1-14.

Murdock B. B. (1960). The immediate retention of unrelated words. Journal of Experimental Psychology 60:222-234.

Murdock B. B. (1962). The serial position effect of free recall. Journal of Experimental Psychology, 64(5):482-488.

Murdock B., Okada R. (1970). Inter-response times in single-trial free recall. Journal of Experimental Psychology, 86(2):263-267.

Murray D. J., Pye C., and Hockley W. E. (1976). Standing’s power function in long-term memory. Psychological Research, 38(4):319-331

Murray, D. J. (1975). Graphemically cued retrieval of words from long-term memory. Journal of Experimental Psychology: Human Learning and Memory, 1(1):65-70.

Polyn S. M., Norman K. A., and Kahana M. J. (2009a). A context maintenance and retrieval model of organizational processes in free recall. Psychological Review, 116:129-156.

Polyn S. M., Norman K. A., and Kahana M. J. (2009b). Task context and organization in free recall. Neuropsychologia, 47:2158-2163.

Roberts W.A. (1972). Free recall of word lists varying in length and rate of presentation: a test of total-time hypotheses. Journal of Experimental Psychology 92:365-372.
Rohrer D., Wixted J.T. (1994). An analysis of latency and interresponse time in free recall., Mem Cognit.:511-24

Romani S., Pinkoviezky I., Rubin A., Tsodyks M. (2013). Scaling laws of associative memory retrieval. Neural Computation 25:2523-2544.

Sederberg P. B., Howard M. W., and Kahana M. J., (2008). A context-based theory of recency and contiguity in free recall. Psychological Review, 115(4):893-912.

Standing L. (1973). Learning 10.000 pictures. Quarterly Journal of Experimental Psychology 25:207-222.

Vonberg I., Ehlen F., Fromm O., and Klostermann. (2014). The Absoluteness of Semantic Processing: Lessons from the Analysis of Temporal Clusters in Phonemic Verbal Fluency. PLoS One, 9(12).