



Chunking dynamics: heteroclinics in mind

Mikhail I. Rabinovich¹, Pablo Varona^{2*}, Irma Tristan¹ and Valentin S. Afraimovich³

¹ BioCircuits Institute, University of California, San Diego, La Jolla, CA, USA

² Grupo de Neurocomputación Biológica, Departamento de Ingeniería Informática, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Madrid, Spain

³ Instituto de Investigación en Comunicación Óptica, Universidad Autónoma de San Luis Potosí, San Luis Potosí, México

Edited by:

Tobias Alecio Mattei, Ohio State University, USA

Reviewed by:

Maurizio Mattia, Istituto Superiore di Sanità, Italy

Hiroshi Okamoto, RIKEN Brain Science Institute, Japan

*Correspondence:

Pablo Varona, Grupo de Neurocomputación Biológica, Departamento de Ingeniería Informática, Escuela Politécnica Superior, Universidad Autónoma de Madrid, C/Francisco Tomás y Valiente, 11, 28049 Madrid, Spain
e-mail: pablo.varona@uam.es

Recent results of imaging technologies and non-linear dynamics make possible to relate the structure and dynamics of functional brain networks to different mental tasks and to build theoretical models for the description and prediction of cognitive activity. Such models are non-linear dynamical descriptions of the interaction of the core components—brain modes—participating in a specific mental function. The dynamical images of different mental processes depend on their temporal features. The dynamics of many cognitive functions are transient. They are often observed as a chain of sequentially changing metastable states. A stable heteroclinic channel (SHC) consisting of a chain of saddles—metastable states—connected by unstable separatrices is a mathematical image for robust transients. In this paper we focus on hierarchical chunking dynamics that can represent several forms of transient cognitive activity. Chunking is a dynamical phenomenon that nature uses to perform information processing of long sequences by dividing them in shorter information items. Chunking, for example, makes more efficient the use of short-term memory by breaking up long strings of information (like in language where one can see the separation of a novel on chapters, paragraphs, sentences, and finally words). Chunking is important in many processes of perception, learning, and cognition in humans and animals. Based on anatomical information about the hierarchical organization of functional brain networks, we propose a cognitive network architecture that hierarchically chunks and super-chunks switching sequences of metastable states produced by winnerless competitive heteroclinic dynamics.

Keywords: cognitive dynamics, stable heteroclinic channel, transient dynamics, low dimensionality of brain activity, hierarchical sequences, chunking and superchunking, cognition modeling principles

INTRODUCTION

Chunking is a dynamical phenomenon that the brain uses for processing long informational sequences. The concept of chunk was introduced by Miller (1956). His key notion is that short-term storage is not rigid but amenable to strategies such as chunking that can expand its capacity. Miller's work drew plenty of attention to the concept of short-term memory and its functional characteristics. Chunking involves two processes: concatenation of units in a block and segmentation of the blocks. In general, chunking is related to the hierarchical organization of perceptual, cognitive, or behavioral sequential activity. In particular, in motor control (see Rosenbaum et al., 1983) sequences can consist of sub-sequences and these can in turn consist of sub-sub-sequences, etc. The natural hierarchical organization of long sequences is a result of the activity of specific brain functional networks. Such networks include many different brain areas and some of them are also organized in a hierarchical manner. A well-known example is Broca's area that has been suggested to act as a "supramodal syntactic processor," able to process any type of hierarchically organized sequences (Grossman, 1980; Tettamanti and Weniger, 2006), a hypothesis based on the findings that this region is not only involved in processing language syntax (Musso et al., 2003), but also in syntax like aspects of non-linguistic tasks, for example, the performance of specific movements and music (Fadiga

et al., 2009) as several fMRI studies (Bahlmann et al., 2008, 2009) seem to confirm. Clerget et al. hypothesize that motor behavior shares some similarities with language (Clerget et al., 2013), namely that a complex action can be viewed as a chain of subordinate movements, which need to be combined according to certain rules in order to reach a given goal (Dehaene and Changeux, 1997; Dominey et al., 2003; Botvinick, 2008).

What are the mechanisms that transform the extremely complex, noisy, and many-dimensional brain activity into a rather regular, low-dimensional, and even predictable cognitive behavior, e.g., what are the mechanisms underlying the dynamics of the mind, including chunking? This is one of the most challenging questions in today's neuro- and cognitive science. Recent continuous advances in non-invasive brain imaging allow assessing the structural connectivity of the brain and the corresponding evolution of the spatio-temporal activity in detail.

In our view, metastability is a key element of transient cognitive dynamics participating in chunking processes. The idea of the spatiotemporal organization of brain dynamic activity through transient, metastable states emerged more than 15 years ago (Kelso, 1995; Friston, 1997). According to this scenario, such dynamics can be represented as a sequential switching between different metastable states (for a description of the mathematical basis of this scenario see Rabinovich et al., 2008a,b). Metastable

transient dynamics represent a balance between the segregation of focused cognitive processing and the flexible integration of distributed brain areas. Such integration is necessary for the performance of a specific cognitive function (Bressler and Kelso, 2001; Meehan and Bressler, 2012). The existence of connections that are prevalent over long periods of time supports the well-regarded concept of a hierarchical organization of neural processing (Engel et al., 2001), which is the basis for the understanding of the origin of the chunking dynamics. Because the dimensionality of cognition depends on the number of activated (in contrast to the potentially observable) metastable states, it is important to remember that the brain chooses the necessary metastable states and suppresses those which are irrelevant to the goal of the cognitive process, resulting in a reduced dimensionality. The low-dimensionality of brain cognitive dynamics is based on two important issues: first, the manner of the cognitive task encoding—an external or internal stimulus determining a specific cognitive task excites a set of elements of the community networks which are responsible for the performance of such cognitive activities; and second, the existence of a specific hierarchical organization of the global brain networks that operate for the performance of a specific cognitive task by a moderate number of brain modes.

Based on experimental data suggesting that the processing of sequential cognitive activity on computational grounds is implemented in the brain by spatiotemporally pattern dynamics (see also Sahin et al., 2009), we build here a general dynamical model that produces hierarchical chunking of sequences, which suggests a plausible neural mechanism of chunking dynamics in the brain. This model is reasonably low-dimensional, which allows a detailed dynamical analysis.

MATERIALS AND METHODS

A top-down approach to model transient cognitive dynamics taking into account the experimental observations described in the introduction is to use kinetic equations for the description of spatiotemporal mental modes that contain the discussed metastable states as equilibrium points. The set of brain patterns that sequentially change in the process of the cognitive task performance determine the spatial structure of the modes and the associated connection matrix among them. Using such type of models we can integrate our knowledge about the description of brain activity based on these new ideas related to heteroclinic sequences and their interactions, i.e., heteroclinic networks.

As a top-down departing point, we need a mathematical object that can describe robust transient dynamics and their associated information processing. Once we have this object, we can implement it through a set of canonic equations that can be used to study transient activity at different brain description levels, and in particular to address chunking dynamics. A mathematical image of robust transient sequential dynamics must have two principal features. First, it must be resistant to noise and reliable even in the context of small variations in initial conditions, so that the succession of states visited by the system (its trajectory, or transient) is stable. Second, the transients must be input-specific to contain information about what caused them. These are two fundamental contradictions regarding the use of transient dynamics

for the description of brain activity. Transient dynamics are inherently unstable: any transient depends on initial conditions and cannot be reproduced from arbitrary initial conditions. On the other hand, dynamical robustness in principle prevents sensitivity to informative perturbations. These contradictions can be solved through the concept of metastability, which was introduced to cognitive science at the end of the last century (Kelso, 1995; Friston, 1997, 2000; Fingelkurts and Fingelkurts, 2006; Oullier and Kelso, 2006; Gros, 2007; Ito et al., 2007).

A stable heteroclinic channel (SHC) is a mathematical object that meets the above discussed requirements, which can implement such stable transients. A SHC is defined by a sequence of successive metastable “saddle” states that are connected by separatrices. Under proper conditions, all the trajectories in the neighborhood of these saddle metastable states that form the chain remain in the channel, ensuring robustness and reproducibility over a wide range of control parameters (Rabinovich et al., 2008b). The stability of a channel means that trajectories in the channel do not leave it until the end of the channel is reached.

A simple model to implement SHCs is a generalized Lotka–Volterra equation with N interactive elements:

$$\frac{dA_i(t)}{dt} = A_i(t)F\left(\sigma_i(S_k) - \sum_{j=1}^N \rho_{ij}A_j(t)\right) + A_i(t)\eta_i(t)$$

$$i = 1, \dots, N \quad (1)$$

where $A_i(t) \geq 0$ is the activity rate of element i , σ_i is the gain function that controls the impact of the stimulus, S_k is an environmental stimulus, ρ_{ij} determines the interaction between the variables, η_i represents the noise level, and F is a function, in the simplest case a linear function. The state portrait of the system often contains a heteroclinic sequence linking saddle points. These saddles can be interpreted as successive and temporary winners in a never-ending competitive game, i.e., winnerless competition (WLC) dynamics (Rabinovich et al., 2001, 2006). In neural systems, because a representative model must produce sequences of connected neuronal population states (the saddle points), the neural connectivity ρ_{ij} must be asymmetric, as determined by the theoretical examination of this model (Huerta and Rabinovich, 2004). Although many connection statistics probably work for stable heteroclinic-type dynamics, it is likely that connectivity within biological networks is, to some extent at least, the result of optimization by evolution and synaptic plasticity. It is important to emphasize that Equation (1) is just an elementary building block for different levels of the chunking hierarchy that we will describe below.

Models like the generalized Lotka–Volterra equations allow establishing the conditions necessary for transient stability, and display stable, sequential, and cyclic activation of its components, the simplest variant of WLC. A network with several degrees of freedom and asymmetric connections can generate structurally stable sequences—transients, each shaped by one input. Asymmetric inhibitory connectivity helps to solve the apparent paradox that sensitivity and reliability can coexist in a network (Huerta and Rabinovich, 2004; Nowotny and Rabinovich, 2007; Rabinovich et al., 2008b; Rabinovich and Varona, 2011). The

neurons or modes participating in a SHC are assigned by the stimulus, by virtue of their direct and/or indirect input from the neurons activated by that stimulus. The joint action of the external input and a stimulus-dependent connectivity matrix defines the stimulus-specific heteroclinic channel. In addition, asymmetric inhibition coordinates the sequential activity and keeps a heteroclinic channel stable.

The WLC concept is directly related to the sequential dynamics of metastable states that are activated by inputs that do not destroy the origin of a competitive process. This paradigm can explain and predict many dynamical phenomena in neural networks with excitatory and inhibitory synaptic connections. Based on the requirement of the stability, this formalism has been used (i) to assess the dynamical origin of finite working memory (WM) capacity based upon WLC amongst available informational items (Bick and Rabinovich, 2009; Rabinovich et al., 2012); (ii) to build a dynamical model of information binding for transients that can describe the interaction of different sensory information flows that are generated concurrently (Rabinovich et al., 2010a); (iii) to model the sequential interaction between emotion and cognition (Rabinovich et al., 2010b); (iv) to represent attention dynamics (Rabinovich et al., 2013); and (v) to assess the dynamics of pathological states in mental disorders (Bystritsky et al., 2012; Rabinovich et al., 2013). Here we focus on a model of hierarchical chunking dynamics that can represent several forms of cognitive activity such as WM and speech construction.

As we discussed in the Introduction, chunking is grouping or categorizing related issues or information into smaller, most meaningful and compact units. Think about how hard it would be to read a long review paper without chapters, subchapters, paragraphs, and separated sentences. Chunking is a naturally occurring process that can be actively used to break down problems in order to think, understand, and make improvisation more efficiently. This is because it is easier to process chunked tasks or perceptual data. In particular, it is much easier to learn and recall such data. Mathematically, the “chunking principle” can be viewed as the transformation of a chain of metastable states along a transient process to the chain of groups of such states. It is a key dynamical idea that nature may use to make cognitive information processing more effective in the context of a complex environment.

Chunking processes in human perception, learning, and performance of a cognitive task can be both automatic and directly linked to the environmental stimuli, and controllable by a goal-oriented intrinsic signal (Gobet et al., 2001). It is important to note that chunking is a strategy that supports increasing speed and accuracy through the formation of hierarchical memory structures and complex task-dependent behavioral sequences. Two competitive processes form temporal chunking sequences—one separates long sequences into shorter groups of information items to be easily performed, and the second connects them to express a long sequence as a unified thought or behavioral action (Friederici et al., 2011; Chekaf and Matha, 2012).

Hierarchical chunking dynamics can be implemented in a model of cognitive networks whose information processing relies on SHCs. **Figure 1** illustrates a chunking heteroclinic cognitive network for two hierarchical informational groups—elementary

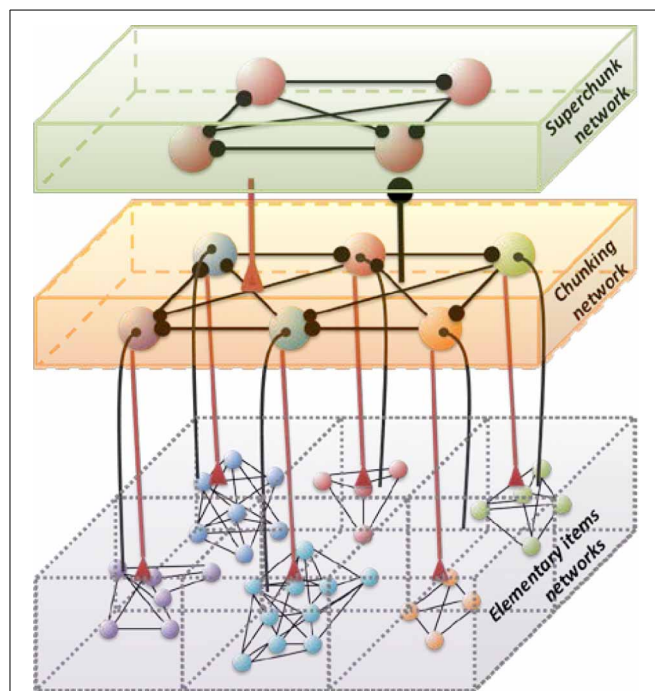


FIGURE 1 | Architecture of the three level cognitive network

responsible for the grouping of informational items. Each level of hierarchy is described by its own Lotka–Volterra type Equations (see 2–6) with connection matrices ρ , ξ and ζ . Black circles represent inhibitory connections; triangles represent excitatory connections responsible for the choosing of the informational items or units (metastable states). Spheres represent the informational items or units (metastable states). Different colors indicate different chunks. All connections inside the elementary items are inhibitory.

items and chunking (integrated) informational items including many elementary units interacting through dynamical connections. It is reasonable to hypothesize that functionally there are two different cognitive networks from at least two different hierarchical levels that are responsible for the: (i) organization of the sequence of items inside chunks, and (ii) the formation of the chunk sequence. In particular, this hypothesis is supported by an experiment with chunking during visuomotor sequence learning (Sakai et al., 2003). It has been shown that each motor cluster is processed as a single memory unit—a chunk. A learned visuomotor sequence is a sequence of chunks that contains several elementary movements. The authors of this work have shown that a key role in the process of chunking formation is played by a brain network including the dominant parietal area, the basal ganglia, and the presupplementary motor area (see also Ribas-Fernandes et al., 2011 and Bor and Seth, 2012, where authors discuss the chunking structure of conscious processes).

Below we suggest a three level hierarchical model for the description of the chunking dynamics. Inhibition plays a key role in this model as is responsible for the execution of three functions: (i) competition between elementary informational items in order to produce stable sequences of metastable states, (ii) generation of the chunking sequence, and (iii) control of the performance of the sequential task. In recent years, the investigation of the hierarchical control between different levels of representation and

information processing has become one of the hot subjects in cognitive science. This issue is important for understanding how the mind controls behavior and itself. In particular, the relationship between chunking (a sequence-level process) and task-set inhibition (a task-level process) in the performance of task sequences was investigated in (Koch et al., 2006; Schneider, 2007; Li et al., 2010), for a description of “chunks of chunks”—“superchunks” see Rosenberg and Feigensohn (2013).

To understand the emergence of hierarchical chunking dynamics in a model we need to depart from Equation (1) in the following direction, c.f. **Figure 1**):

$$\dot{X}_i^{lk} = X_i^{lk} \left(\sigma_i^{lk}(S, C) \cdot Y^{lk} - \sum_j \rho_{ij}^{lk}(S, C) X_j^{lk} \right) \quad (2)$$

$$\tau \dot{Y}^{lk} = Y^{lk} \left(\left(V^l - \beta(C) \sum_i X_i^{lk} \right) - Z^{lk} \right) \quad (3)$$

$$\theta(C) \dot{Z}^{lk} = \sum_m \xi_i^{km}(S, C) Y^{lm} - Z^{lk} \quad (4)$$

$$T \dot{V}^l = V^l \left(\left(1 - \delta(C) \sum_j Y^{lj} \right) - W^l \right) \quad (5)$$

$$\Theta(C) \dot{W}^l = \sum_q \zeta^{lq}(S, C) V^q - W^l \quad (6)$$

Here X_i^{lk} characterizes the i -th informational item associated with the k -th chunk and l -th superchunk, $\sigma_i^{lk}(S, C)$ is the growth rate for each informational item determined by the stimulus S and the cognitive task C , and $\rho_{ij}^{lk}(S, C)$ is the matrix of inhibitory connections among basic informational items. In this model Y^{lk} characterizes the k -th chunk associated to the l -th superchunk V^l , with corresponding characteristic times τ and T , respectively, and $\beta(C)$ represents the strength of the inhibition between the informational items and the chunk, and $\delta(C)$ between the chunks and the superchunk. Also, Z^{lk} describes the synaptic dynamics for the k -th chunk associated to the l -th superchunk with $\xi_i^{km}(S, C)$, the matrix of inhibitory connections between chunks (black circles in **Figure 1**); and W^l describes the synaptic dynamics for the l -th superchunk with $\zeta^{lq}(S, C)$, the matrix of inhibitory connections between superchunks, the corresponding characteristic times are $\theta(C)$ and $\Theta(C)$. In this model, $\beta(C)$ and $\delta(C)$ are adaptation parameters that determine the timing relationship between a basic informational chain and the chunking and superchunking modulation. The chunking variables also satisfy the generalized Lotka–Volterra—canonic equations which allows them to form a stable sequence. Because of this, in fact, chunking variables play the role of cognitive controllers. The parameters for Equations (3)–(5) in the simulations below were chosen with this scope. Since chunking dynamics has to take into account of the characteristic time of the chunk formation, the competition between different chunks has to be delayed—we used for this an inhibition described by a first order kinetic model. At the same time,

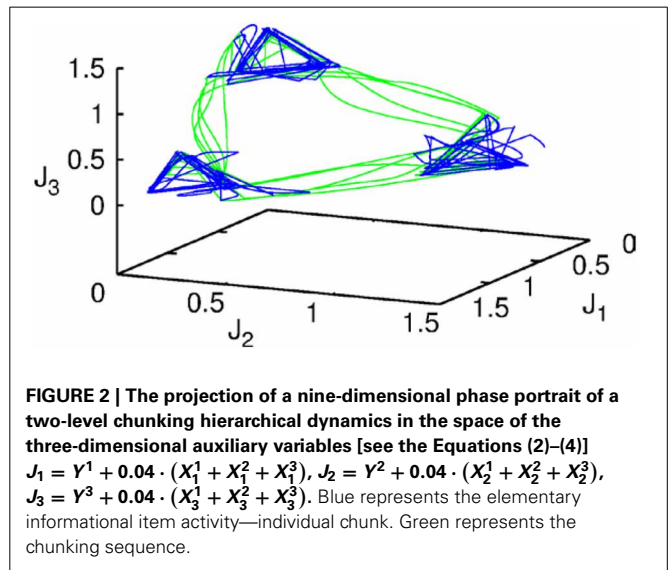
the competition among elementary informational items is implemented by fixed weight ρ_{ij} instantaneous synapses. The same logic has been applied for the description of the highest level of the hierarchy—the superchunks.

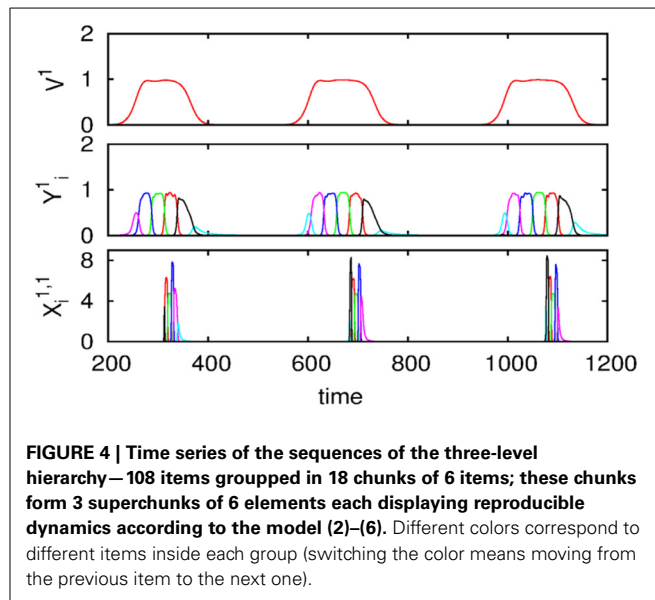
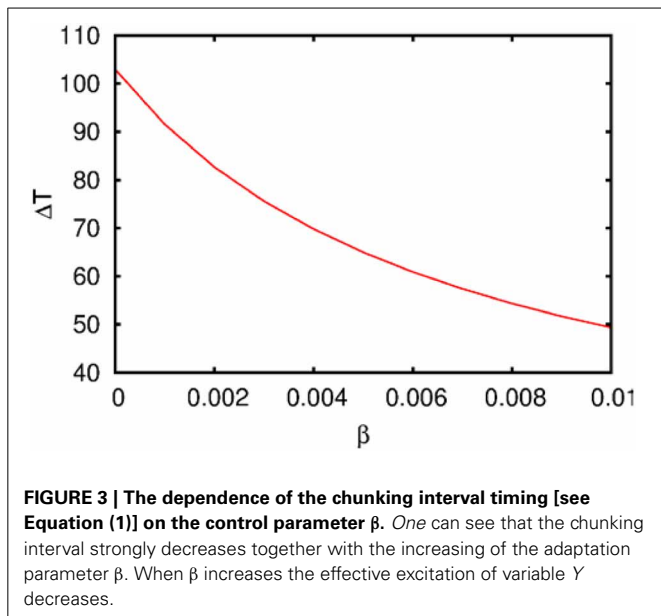
RESULTS: HIERARCHICAL SEQUENCES—CHUNKING AND SUPER-CHUNKING

Let us first represent the phase portrait of a simple two-level chunking dynamics. We carried out numerical simulations of the model for the dynamics within chunks of informational items for the following parameters $N^k = 3$, $M = 3$ (number of “chunks” or “episodes”), $\sigma^1 = [7.24, 5.85, 8.30]$, $\sigma^2 = [9.93, 6.00, 5.18]$, $\sigma^3 = [8.29, 7.86, 9.16]$, and given these values, $\rho_{ii}^k = 1.0$, $\rho_{i_n-i_n}^k = \frac{\sigma_{i_n}^k}{\sigma_{i_n}^k} + 0.51$, and $\rho_{i_n+i_n}^k = \frac{\sigma_{i_n+1}^k}{\sigma_{i_n}^k} - 0.5$, $i = 1, \dots, N^k$, $k = 1, \dots, M$ as well as the parameters considered for the synaptic dynamics described by Equations (3) and (4): $\tau = 0.7$, $\theta = 2.0$, $\xi^{kk} = 1.0$, $\xi^{k_n k_{n+1}} = 1.4$ and $\xi^{k_n k_{n-1}} = 0.5$, $k = 1, \dots, M$ and $\beta = 0.01$. The results of these simulations are shown in **Figures 2, 3**.

Figure 2 shows the phase portrait of the chunking dynamics when the superchunk formation is absent: the system is described by Equations (2)–(4), $V = 1$. This example illustrates a closed chunking sequence (green) that consists of several heteroclinic cycles that represent the elementary chunks (blue). In general, the number of elementary items in each chunk are different and the chunking sequence can be open.

Figure 3 illustrates the timing between chunks along the sequence. The emergence of the chunking sequence shown in **Figure 2** is the result of a modulational instability in the two-level hierarchical network whose dynamics is described by Equations (2)–(4). This instability is oscillatory. The characteristic period of the oscillation is ΔT . The analytical investigation of the dependence of ΔT on the control parameters τ , θ , β and connection matrices ρ , ξ is a non-realistic problem because of the non-linear feedback between the dynamical variables X and Y . However, it is reasonable to think that the key parameter in this problem is





β which determines the level of excitability of variable Y and, according to the feedback, also controls the excitability of X (term $\sigma_i^{lk}(S, C) \cdot X_i^{lk} \cdot Y^{lk}$) in the right hand side of Equation (2). In **Figure 3** we represent the numerical analysis of the dependence of ΔT on the parameter β —increasing β , i.e., decreasing the excitability leads to the decreasing of the timing interval ΔT .

We also carried out numerical simulations of a high-dimensional model that describes the dynamics of chunk and super-chunk formation with the following parameters: $N^{lk} = 6$, $M^l = 6$ (number of chunks), $P = 3$ (number of superchunks), $\sigma^{l1} = [6.94, 5.11, 8.94, 5.86, 8.33, 9.62]$, $\sigma^{l2} = [5.48, 5.66, 5.39, 9.89, 9.99, 5.82]$, $\sigma^{l3} = [7.65, 8.98, 9.21, 6.02, 5.71, 5.12]$, $\sigma^{l4} = [7.61, 7.73, 5.62, 7.93, 5.80, 5.39]$, $\sigma^{l5} = [5.11, 9.99, 5.52, 5.66, 5.50, 8.21]$, $\sigma^{l6} = [5.84, 9.39, 7.08, 5.16, 8.37, 6.87]$, and given these values, $\rho_{ii}^{lk} = 1.0$, $\rho_{i_{n-1}i_n}^{lk} = \frac{\sigma_{i_{n-1}}^{lk}}{\sigma_{i_n}^{lk}} + 0.5$

$$1, \rho_{i_{n+1}i_n}^{lk} = \frac{\sigma_{i_{n+1}}^{lk}}{\sigma_{i_n}^{lk}} - 0.5, i = 1, \dots, N^{lk}, k = 1, \dots, M^l, l = 1, \dots,$$

$$P, \text{ and } \rho_{i_{n-1}i_n}^{lk} = \rho_{i_{n-1}i_n}^{lk} + \frac{\sigma_i^{lk} - \sigma_{i_{n-1}}^{lk}}{\sigma_{i_n}^{lk}} + 2, i \neq \{i_{n-1}, i_n, i_{n+1}\}, \text{ as}$$

well as the parameters considered for the synaptic dynamics between chunks described by the equations $\tau = 0.8$, $\theta = 2.0$, $\xi_l^{kk} = 1.0$, $\xi_l^{k_n k_{n-1}} = 0.5$, $\xi_1^{k_n k_{n+1}} = 1.4$, $\xi_2^{k_n k_{n+1}} = 1.3$, $\xi_3^{k_n k_{n+1}} = 1.5$, $k = 1, \dots, M^l$, $l = 1, \dots, P$, $\xi_l^{k k_n} = \xi_l^{k_{n-1} k_n} + 2$, $k \neq \{k_{n-1}, k_n, k_{n+1}\}$, and $\beta = 0.01$. Finally, the parameters for the synaptic dynamics between superchunks were $T = 5$, $\Theta = 10$, $\zeta^{ll} = 1.0$, $\zeta^{l_{n-1} l_n} = 0.5$, $\zeta^{l_{n-1} l_{n+1}} = 1.4$, $l = 1, \dots, P$, and $\delta = 0.01$. The result of these simulations are displayed in **Figure 4**, which shows three levels of information hierarchy: original informational chain (lower panel), chunked chain (middle panel), and superchunking chain (upper panel).

As illustrated in **Figure 2**, the sequence of chunks can be considered as a heteroclinic cycle of metastable states where each metastable state itself is a heteroclinic cycle of elementary informational items. Based on this self-similarity, we can expect that

the chunking chain as a result of a second heteroclinic instability generates the next level of modulation—the superchunk sequence. Our expectation is confirmed in **Figure 4** that shows the time series of the three level network (2)–(6) (c.f. **Figure 1**) dynamics. In this figure, one can see the generation of sequences of superchunks. All together, the sequences informational items, chunks and superchunks can be interpreted as “words,” “sentences,” and “paragraphs.”

For the sake of simplicity we have illustrated here the phenomenon of stability just for a closed-loop clustered chunking-superchunking sequence. In the general case of open sequence, it is possible to formulate the sufficient conditions for the existence and stability of the non-closed channel based on the estimation of the saddle values of the metastable states (elementary items)—the channel is stable in the case that all of them are larger than one in absolute value (Afraimovich et al., 2004; Bick and Rabinovich, 2010). The formulation of the necessary conditions is a more complex problem and is still under consideration. The imposed stability conditions determine the behavior of the trajectories inside the neighborhood of the heteroclinic network independently of the initial conditions as computer experiments have confirmed (Afraimovich et al., 2004; Bick and Rabinovich, 2010).

The above described numerical results can be justified by an analytical study of the system

$$\begin{cases} \dot{X}_i^k = X_i^k \left(\sigma_i^k \cdot Y^k - \sum_{j=1}^{N^k} \rho_{ij}^k X_j^k \right), \\ \tau \dot{Y}^k = Y^k \left(1 - \beta \sum_{i=1}^{N^k} X_i^k - Z^k \right), \\ \theta \dot{Z}^k = \sum_{m=1}^M \xi^{km} Y^m - Z^k \end{cases} \quad (7)$$

$i = 1, \dots, N^k$, $k = 1, \dots, M$. For the sake of simplicity, let us assume that $\tau = \theta \ll 1$, so one can apply geometric singular perturbation theory (see, for instance, Jones, 1995; Hek, 2010 and references therein). In order to avoid confusion, it is important to say that the assumption $\tau = \theta \ll 1$ implies that, in contrast to the dynamics of X , the chunking dynamics is a composition of fast and slow motions. The fast motions lead variables Y -th and Z -th to a neighborhood of the slow manifold in the phase space. The evolution of the chunk variables on this manifold in the vicinity of the metastable states is much slower than the X variables. This corresponds to the intuitively clear fact that the “enveloping” variables mimic the averaging

dynamics of X . Computer experiments confirm this explanation (see Figure 4).

The limit slow manifold has the equations $Y^k \left(1 - \beta \sum_{i=1}^{N^k} X_i^k - Z^k\right) = 0$, $\sum_{m=1}^M \xi^{km} Y^m - Z^k = 0$, thus, $\sum_{m=1}^M \xi^{km} Y^m = 1 - \beta \sum_{i=1}^{N^k} X_i^k$. Denote by ξ the $m \times m$ -matrix ξ^{km} . If $\det \xi \neq 0$, we find

$$Y^k = \frac{1}{\det \xi} \left(\sum_{m=1}^M \eta^{mk} - \beta \sum_{m=1}^M \eta^{mk} \sum_{i=1}^{N^m} X_i^m \right) \quad (8)$$

Table 1 | Sequential dynamics in neural and cognitive systems.

Phenomenon/image	Model	References	Comments
Voting paradox / Structurally stable heteroclinic cycle	Kinetic (rate) equation, Lotka–Volterra model	Krupa, 1997; Stone and Armbruster, 1999; Ashwin et al., 2003; Postlethwaite and Dawes, 2005	J. C. Borda and the Marquis de Condorcet (De Borda, 1781; Saari, 1995) analyzed the process of plurality elections at the French Royal Academy of Sciences. They predicted the absence of a winner in a 3 step voting process (Condorcet’s triangle)
Learning sequences	Hopfield type non-symmetric networks with time delay including spiking neuron models	Amari, 1972; Kleinfeld, 1986; Sompolinsky and Kanter, 1986; Minai and Levy, 1993; Deco and Rolls, 2005	Networks proposed to explain the generation of rhythmic motor patterns and the recognition and recall of sequences
Latching dynamics	Potts network is able to hop from one discrete attractor to another under random perturbation to make a sequence	Treves, 2005; Russo et al., 2008; Russo and Treves, 2011; Linkerhand and Gros, 2013	The dynamics can involve sequences of continuously latching transient states
Sequential memory with synaptic dynamics / Chaotic itinerancy sequences of Milnor attractors or attractor ruins	Spike-frequency-adaptation mechanism Noisy dynamical systems. Cantor coding	Tsuda, 2009	Proposed to be involved in episodic memory and itinerant process of cognition
Winnerless sequential switchings along metastable states/Stable heteroclinic channel	Generalized coupled Lotka–Volterra equations	Afraimovich et al., 2004; Rabinovich et al., 2008a,b	Information processing with transient dynamics at many different description levels from simple networks to cognitive processes
Winnerless competitive dynamics in spiking brain networks	Random inhibitory networks of spiking neurons in the striatum	Ponzi and Wickens, 2010	Neurons form assemblies that fire in sequential coherent episodes and display complex identity–temporal spiking patterns even when cortical excitation is constant or fluctuating noisily
Sequences of sequences / Hierarchical transient sequences	Recognition of sequence of sequences based on a continuous dynamical model	Kiebel et al., 2009	Speech can be considered as a sequence of sequences and can be implemented robustly by a dynamical model based on Bayesian inference. recognition dynamics disclose inference at multiple time scales

where η^{km} is the cofactor of the entry ξ^{mk} of the matrix ξ . Substituting this expression into the first equation of the system (7) we obtain the system

$$\dot{X}_i^k = X_i^k \left(\sigma_i^k \frac{1}{\det \xi} \sum_{m=1}^M \eta^{mk} - \sum_{j=1}^{N^k} \rho_{ij}^k X_j^k - \frac{\beta}{\det \xi} \sum_{m=1}^M \eta^{mk} \sum_{i=1}^{N^k} X_i^m \right) \quad (9)$$

$i = 1, \dots, N^k$, $k = 1, \dots, M$, which is similar to the binding model described in Rabinovich et al. (2010a). In particular, the “in-chunk” dynamics in (9) corresponds to the dynamics in the modality subspace in Rabinovich et al. (2010a). The main peculiarity of the system (9) is that the rates of coupling coefficients between different chunks have the common factor β , so if $\beta = 0$ then the interaction between different chunks is absent. Similarly to the study in Rabinovich et al. (2010a), one can impose conditions under which there exists a heteroclinic cycle for each chunk and successive heteroclinic connections between saddle points in different cycles. The last claim has the form $\beta > \beta_{cr}$ where β_{cr} depends on the parameters of the system (9). If τ is small then because of the geometric singular perturbation theory, the imposed conditions shall guarantee the existence of a heteroclinic network in the original system (7) corresponding to the “in-chunk” and “inter-chunk” dynamics.

Observations on the temporal chunk signal have focused on the use of pauses in behavior to probe chunk structures in WM. On the basis of some of these studies, a hierarchical process model has been proposed, which consists of four hierarchical levels describing different kind of pauses. The lowest level consists of pauses between strokes within letters. On higher levels, there are pauses between letters, words, and phrases. Each level is associated with a larger amount of processing when retrieving these chunks from memory (Cheng and Rojas-Anaya, 2006). Writing may be an effective approach to the study of cognitive phenomena that involves the processing of chunks. In Cheng and Rojas-Anaya (2003), it was demonstrated that in the writing of simple number sequences the duration of pauses between written elements (digits) that are within a chunk are shorter than the pauses between elements across the boundary of chunks. This temporal signal is apparent in un-aggregated data for individual participants in single trials. Mathematically the time intervals between chunks and super-chunks are controlled by parameter β (see Equation 3).

DISCUSSION

In this paper we have shown how the architecture of hierarchical mental model networks affected their associated functions. The discussed examples illustrate that networks with metastable states having several unstable separatrices exhibit very diverse cognitive functions (behavior). Complex heteroclinic networks allow completely new dynamical phenomena, and one of the primary challenges is the assessment of the existence and stability of hierarchical—chunking processes that can represent cognitive activity.

It is important to remind that the modeling of cycling and sequential dynamics in behavior and cognition has a long history (see several representative efforts in Table 1). Most of these

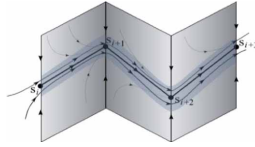
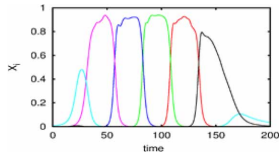
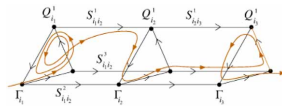
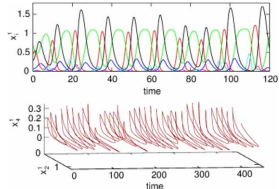
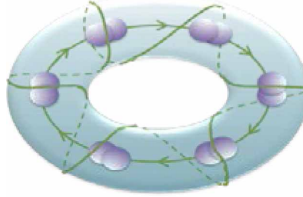
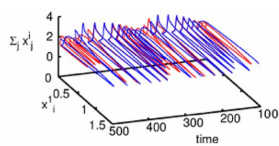
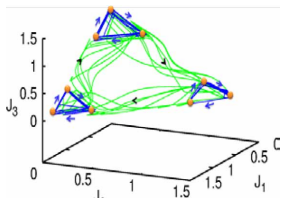
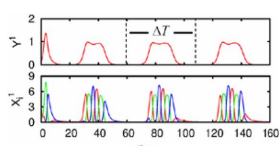
models are based on Hopfield type networks. The main problem there is to keep the stability of the recall sequences against noise.

The results of chunking dynamics reported in this paper can be viewed as relevant in the description of different cognitive tasks. For example, in WM, humans encode items and synthesize them. With that, we give meaning to ideas and find a relevant place for them in our cognitive world. In these actions the interaction between WM and chunking are reciprocal—first of all WM is the “engine” of chunking, and on the other hand, the chunking makes WM capacity higher.

The model of chunking dynamics discussed in this paper relies on heteroclinic dynamics. It is important to emphasize that the main features of the SHC do not depend on the specific model used. The conditions of existence and the dynamical features of SHCs can be implemented in a wide variety of models: from simple Lotka–Volterra descriptions to complex Hodgkin–Huxley models, and from small networks to large ensembles of many elements (Varona et al., 2002; Venaille et al., 2005; Nowotny and Rabinovich, 2007; Rabinovich et al., 2012). The intrinsic hierarchical nature of the SHC at different temporal and spatial scales allows implementing many types of cognitive dynamics. Within this framework, brain networks can be viewed as non-equilibrium systems and their associated computations as unique patterns of transient activity, controlled by incoming input. The results of these computations can be reproducible, robust against noise, and easily decoded. Using asymmetric inhibition appropriately, the space of possible states of large neural systems can be restricted to connected saddle points, forming SHCs. These channels can be thought of as underlying reliable transient brain dynamics. Table 2 summarizes four types of heteroclinic networks that can describe different aspects of sequential dynamics in cognitive processes: (i) A canonic heteroclinic network that produces reproducible sequential switching from one metastable state to another inside one modality (like in a simple WM task); (ii) A network displaying inhibitory-based heteroclinic binding dynamics that is responsible for the stable perception of a subject based on three different modalities; (iii) Two different modalities dynamically coordinated by excitatory connections; (iv) A chunking heteroclinic network that controls the grouping of elements of sequential behavior.

Mathy and Feldman have recently suggested to use the Kolmogorov complexity and compressibility (Mathy and Feldman, 2012) for the definition of a “chunk”: a chunk is a unit in a maximally compressed code. The authors presented a series of experiments in which they manipulated the compressibility of stimulus sequences by introducing sequential patterns of variable length. To explore the influence of chunking on the capacity limits of WM, and departing from Bick and Rabinovich (2009), authors in Li et al. (2013) have suggested a model for chunking in sequential WM. This model also uses hierarchical bidirectional inhibition-connected neural networks with WLC. Assuming no interaction between a basic sequence and a chunked sequence, and the existence of an upper bound to the inhibitory weights the network, authors show that chunking increases the number of memorized items in WM from the “magical number” 7–16 items. The optimal number of chunks and the number of the memorized items in each chunk correspond to the “magical number 4.”

Table 2 | Heteroclinics in mind.

Phenomenon	Network formalism*	Phase portrait	Time series
Sequential heteroclinic switching	$\dot{X}_i = X_i \left(\sigma_i - \sum_{j=1}^N \rho_{ij} X_j \right)$		
Sequential heteroclinic binding and information flow	$\dot{X}_i^l = X_i^l \left(\sigma_i^l - \sum_{j=1}^N \rho_{ij}^l X_j^l - \sum_{m=1}^L \sum_{j=1}^N \xi_{ij}^{lm} X_j^m \right)$		
Heteroclinic cooperation	$\tau_i^m \dot{X}_i^m = X_i^m \cdot \left[\sigma_i^m - \sum_{j=1}^{K^m} \rho_{ij}^m X_j^m + \sum_{k=1}^M \sum_{j=1}^{K^m} \xi_{ij}^{mk} X_j^k \right]$		
Hierarchical chunking memory and learning	$\begin{aligned} \dot{X}_i^k &= X_i^k \left(\sigma_i^k \cdot \gamma^k - \sum_j^N \rho_{ij}^k X_j^k \right) \\ \tau \dot{\gamma}^k &= \gamma^k \left(\left(1 - \beta \sum_i^N X_i^k \right) - Z^k \right) \\ \theta \dot{Z}^k &= \sum_{m=1}^M \xi^{km} \gamma^m - Z^k \end{aligned}$		

*See the definition of the variables and parameters in the text.

Recent experiments have confirmed the existence of three levels of cognitive hierarchy—see Rosenberg and Feigenson (2013). In this paper authors reported that infants can unify the representation of chunks into “super-chunks.”

The chunking models discussed above can be generalized on more complex cases. In particular, by adding attention control in the network hierarchy, it is possible to analyze the binding of sequences of chunks. The brain could use such binding to perform many cognitive functions like the coordination of visual perception with speech comprehension, or the coordination of music chunks and word chunks in singing processes. It is well-known that viewing a speaker’s articulatory movements substantially improves a listener’s ability to understand spoken words, especially under noisy environmental conditions like in a crowded cocktail party. Ross and coauthors claimed that this effect is most pronounced when the auditory input is weakest. As a result of attentional binding—multisensory integration—, substantial gain in multisensory speech enhancement is achieved at even the lowest signal-to noise ratios (Ross et al., 2007).

The dynamics of hierarchical heteroclinic networks is also able to explain and predict the coordination of behavioral elements with different time scales (for a study about the coordination of sensorimotor dynamics see Jantzen and Kelso, 2007). Functionally, such kind of synchronization can be the result of

learning—the changing of the strength of inhibitory connections between agents at the different levels of the hierarchy in order to coordinate the dynamics with different time scales (see Figure 3). Additionally, it is important to note that the winnerless competitive learning process itself can be chaotic (Komarov et al., 2010), which provides wider possibilities for adaptability.

ACKNOWLEDGMENTS

Mikhail I. Rabinovich acknowledges support from ONR grant N00014310205. Pablo Varona was supported by MINECO TIN2012-30883. Irma Tristan acknowledges support from the UC-MEXUS-CONACYT Fellowship. Valentin S. Afraimovich was partially supported by Ohio University Glidden Professorship program.

REFERENCES

Afraimovich, V. S., Zhigulin, V. P., and Rabinovich, M. I. (2004). On the origin of reproducible sequential activity in neural circuits. *Chaos* 14, 1123. doi: 10.1063/1.1819625
 Amari, S.-I. (1972). Learning patterns and pattern sequences by self-organizing nets of threshold elements. *IEEE Trans. Comput.* C-21, 1197–1206. doi: 10.1109/T-C.1972.223477
 Ashwin, P., Field, M., Rucklidge, A. M., and Sturman, R. (2003). Phase resetting effects for robust cycles between chaotic sets. *Chaos* 13, 973–981. doi: 10.1063/1.1586531

- Bahlmann, J., Schubotz, R. I., and Friederici, A. D. (2008). Hierarchical artificial grammar processing engages Broca's area. *Neuroimage* 42, 525–534. doi: 10.1016/j.neuroimage.2008.04.249
- Bahlmann, J., Schubotz, R. I., Mueller, J. L., Koester, D., and Friederici, A. D. (2009). Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Res.* 1298, 161–170. doi: 10.1016/j.brainres.2009.08.017
- Bick, C., and Rabinovich, M. I. (2009). Dynamical origin of the effective storage capacity in the brain's working memory. *Phys. Rev. Lett.* 103:218101. doi: 10.1103/PhysRevLett.103.218101
- Bick, C., and Rabinovich, M. I. (2010). On the occurrence of stable heteroclinic channels in Lotka–Volterra models. *Dyn. Syst.* 25, 1–14. doi: 10.1080/14689360903322227
- Bor, D., and Seth, A. K. (2012). Consciousness and the prefrontal parietal network: insights from attention, working memory, and chunking. *Front. Psychol.* 3:1–14. doi: 10.3389/fpsyg.2012.00063
- Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends Cogn. Sci.* 12, 201–208. doi: 10.1016/j.tics.2008.02.009
- Bressler, S. L., and Kelso, J. A. S. (2001). Cortical coordination dynamics and cognition. *Trends Cogn. Sci.* 5, 26–36. doi: 10.1016/S1364-6613(00)01564-3
- Bystritsky, A., Nierenberg, A. A., Feusner, J. D., and Rabinovich, M. (2012). Computational non-linear dynamical psychiatry: a new methodological paradigm for diagnosis and course of illness. *J. Psychiatr. Res.* 46, 428–435. doi: 10.1016/j.jpsychires.2011.10.013
- Chekaif, M., and Matha, F. (2012). “Chunking memory span of categorizable objects,” in *53rd Annual Meeting of the Psychonomic Society* (Minneapolis, MN: Psychonomic Society Publication).
- Cheng, P. C. H., and Rojas-Anaya, H. (2003). “Writing out a temporal signal of chunks: patterns of pauses reflect the induced structure of written number sequences,” in *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (Mahwah, NJ: Lawrence Erlbaum), 424–429.
- Cheng, P. C.-H., and Rojas-Anaya, H. (2006). “A temporal signal reveals chunk structure in the writing of word phrases,” in *Proceedings of the Twenty-Eighth Annual Conference of the Cognitive Science Society* (Mahwah, NJ: Lawrence Erlbaum).
- Clerget, E., Andres, M., and Olivier, E. (2013). Deficit in complex sequence processing after a virtual lesion of left BA45. *PLoS ONE* 8:e63722. doi: 10.1371/journal.pone.0063722
- De Borda, J. C. (1781). *Memoire sur les Elections au Scrutin*. Paris: Academie Royale des Sciences.
- Deco, G., and Rolls, E. T. (2005). Sequential memory: a putative neural and synaptic dynamical mechanism. *J. Cogn. Neurosci.* 17, 294–307. doi: 10.1162/0898929053124875
- Dehaene, S., and Changeux, J. P. (1997). A hierarchical neuronal network for planning behavior. *Proc. Natl. Acad. Sci. U.S.A.* 94, 13293–13298. doi: 10.1073/pnas.94.24.13293
- Dominey, P. F., Hoen, M., Blanc, J.-M., and Lelekov-Boissard, T. (2003). Neurological basis of language and sequential cognition: evidence from simulation, aphasia, and ERP studies. *Brain Lang.* 86, 207–225. doi: 10.1016/S0093-934X(02)00529-1
- Engel, A. K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* 2, 704–716. doi: 10.1038/35094565
- Fadiga, L., Craighero, L., and D'Ausilio, A. (2009). Broca's area in language, action, and music. *Ann. N.Y. Acad. Sci.* 1169, 448–458. doi: 10.1111/j.1749-6632.2009.04582.x
- Fingelkurts, A. A., and Fingelkurts, A. A. (2006). Timing in cognition and EEG brain dynamics: discreteness versus continuity. *Cogn. Process.* 7, 135–162. doi: 10.1007/s10339-006-0035-0
- Friederici, A. D., Bahlmann, J., Friedrich, R., and Makuuchi, M. (2011). The neural basis of recursion and complex syntactic hierarchy. *Biolinguistics* 5, 87–104. Available online at: <http://www.biolinguistics.eu/index.php/biolinguistics/article/view/170>
- Friston, K. J. (1997). Transients, metastability, and neuronal dynamics. *Neuroimage* 5, 164–171. doi: 10.1006/nimg.1997.0259
- Friston, K. J. (2000). The labile brain. II. Transients, complexity and selection. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 355, 237–252. doi: 10.1098/rstb.2000.0561
- Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C.-H., Jones, G., Oliver, I., et al. (2001). Chunking mechanisms in human learning. *Trends Cogn. Sci.* 5, 236–243. doi: 10.1016/S1364-6613(00)01662-4
- Gros, C. (2007). Neural networks with transient state dynamics. *New J. Phys.* 9, 109. doi: 10.1088/1367-2630/9/4/109
- Grossman, M. (1980). A central processor for hierarchically-structured material: evidence from Broca's aphasia. *Neuropsychologia* 18, 299–308. doi: 10.1016/0028-3932(80)90125-6
- Hek, G. (2010). Geometric singular perturbation theory in biological practice. *J. Math. Biol.* 60, 347–386. doi: 10.1007/s00285-009-0266-7
- Huerta, R., and Rabinovich, M. (2004). Reproducible sequence generation in random neural ensembles. *Phys. Rev. Lett.* 93:238104. doi: 10.1103/PhysRevLett.93.238104
- Ito, J., Nikolaev, A. R., and van Leeuwen, C. (2007). Dynamics of spontaneous transitions between global brain states. *Hum. Brain Mapp.* 28, 904–913. doi: 10.1002/hbm.20316
- Jantzen, K. J., and Kelso, J. S. (2007). “Neural coordination dynamics of human sensorimotor behavior: a review,” in *Handbook of Brain Connectivity*, eds V. K. Jirsa and A. McIntosh (Berlin; Heidelberg: Springer), 421–461.
- Jones, C. K. R. T. (1995). Geometric singular perturbation theory. *Lect. Notes Math.* 1609, 44–118. doi: 10.1007/BFb0095239
- Kelso, J. A. S. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior*. Cambridge, MA: The MIT Press.
- Kiebel, S. J., von Kriegstein, K., Daunizeau, J., and Friston, K. J. (2009). Recognizing sequences of sequences. *PLoS Comput. Biol.* 5:e1000464. doi: 10.1371/journal.pcbi.1000464
- Kleinfeld, D. (1986). Sequential state generation by model neural networks. *Proc. Natl. Acad. Sci. U.S.A.* 83, 9469–9473. doi: 10.1073/pnas.83.24.9469
- Koch, I., Philipp, A. M., and Gade, M. (2006). Chunking in task sequences modulates task inhibition. *Psychol. Sci.* 17, 346–350. doi: 10.1111/j.1467-9280.2006.01709.x
- Komarov, M. A., Osipov, G. V., and Burtsev, M. S. (2010). Adaptive functional systems: learning with chaos. *Chaos* 20, 045119. doi: 10.1063/1.3521250
- Krupa, M. (1997). Robust heteroclinic cycles. *J. Nonlinear Sci.* 7, 129–176. doi: 10.1007/BF02677976
- Li, G., Ning, N., Ramanathan, K., He, W., Pan, L., and Shi, L. (2013). Behind the magical numbers: hierarchical chunking and the human working memory capacity. *Int. J. Neural Syst.* 23:1350019. doi: 10.1142/S0129065713500196
- Li, K. Z. H., Blair, M., and Chow, V. S. M. (2010). Sequential performance in young and older adults: evidence of chunking and inhibition. *Neuropsychol. Dev. Cogn. Sect. B Aging Neuropsychol. Cogn.* 17, 270–295. doi: 10.1080/13825580903165428
- Linkerhand, M., and Gros, C. (2013). Generating functionals for autonomous latching dynamics in attractor relict networks. *Sci. Rep.* 3:2042. doi: 10.1038/srep02042
- Mathy, F., and Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition* 122, 346–362. doi: 10.1016/j.cognition.2011.11.003
- Meehan, T. P., and Bressler, S. L. (2012). Neurocognitive networks: findings, models, and theory. *Neurosci. Biobehav. Rev.* 36, 2232–2247. doi: 10.1016/j.neubiorev.2012.08.002
- Miller, G. A. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97. doi: 10.1037/h0043158
- Minai, A. A., Levy, W. B. (1993). “Sequence learning in a single trial,” in *INNS World Congress on Neural Networks* (Hillsdale, NJ: Erlbaum), 505–508.
- Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Büchel, C., et al. (2003). Broca's area and the language instinct. *Nat. Neurosci.* 6, 774–781. doi: 10.1038/nn1077
- Nowotny, T., and Rabinovich, M. I. (2007). Dynamical origin of independent spiking and bursting activity in neural microcircuits. *Phys. Rev. Lett.* 98:128106. doi: 10.1103/PhysRevLett.98.128106
- Oullier, O., and Kelso, J. A. S. (2006). Neuroeconomics and the metastable brain. *Trends Cogn. Sci.* 10, 353–354. doi: 10.1016/j.tics.2006.06.009
- Ponzi, A., and Wickens, J. (2010). Sequentially switching cell assemblies in random inhibitory networks of spiking neurons in the striatum. *J. Neurosci.* 30, 5894–5911. doi: 10.1523/JNEUROSCI.5540-09.2010
- Postlethwaite, C. M., and Dawes, J. H. P. (2005). Regular and irregular cycling near a heteroclinic network. *Nonlinearity* 18, 1477–1509. doi: 10.1088/0951-7715/18/4/004

- Rabinovich, M., Huerta, R., and Laurent, G. (2008a). Neuroscience. Transient dynamics for neural processing. *Science* 321, 48–50. doi: 10.1126/science.1155564
- Rabinovich, M., Tristan, I., and Varona, P. (2013). Neural dynamics of attentional cross-modality control. *PLoS ONE* 8:e64406. doi: 10.1371/journal.pone.0064406
- Rabinovich, M., Volkovskii, A., Lecanda, P., Huerta, R., Abarbanel, H. D., and Laurent, G. (2001). Dynamical encoding by networks of competing neuron groups: winnerless competition. *Phys. Rev. Lett.* 87:68102. doi: 10.1103/PhysRevLett.87.068102
- Rabinovich, M. I., Afraimovich, V. S., Bick, C., and Varona, P. (2012). Information flow dynamics in the brain. *Phys. Life Rev.* 9, 51–73. doi: 10.1016/j.plrev.2011.11.002
- Rabinovich, M. I., Afraimovich, V. S., and Varona, P. (2010a). Heteroclinic binding. *Dyn. Syst. An Int. J.* 25, 433–442. doi: 10.1080/14689367.2010.515396
- Rabinovich, M. I., Huerta, R., Varona, P., and Afraimovich, V. S. (2006). Generation and reshaping of sequences in neural systems. *Biol. Cybern.* 95, 519–536. doi: 10.1007/s00422-006-0121-5
- Rabinovich, M. I., Huerta, R., Varona, P., and Afraimovich, V. S. (2008b). Transient cognitive dynamics, metastability, and decision making. *PLoS Comput Biol* 4:e1000072. doi: 10.1371/journal.pcbi.1000072
- Rabinovich, M. I., Muezzinoglu, M. K., Strigo, I., and Bystritsky, A. (2010b). Dynamical principles of emotion-cognition interaction: mathematical images of mental disorders. *PLoS ONE* 5:e12547. doi: 10.1371/journal.pone.0012547
- Rabinovich, M. I., and Varona, P. (2011). Robust transient dynamics and brain functions. *Front. Comput. Neurosci.* 5:24. doi: 10.3389/fncom.2011.00024
- Ribas-Fernandes, J. J. E., Solway, A., Diuk, C., McGuire, J. T., Barto, A. G., Niv, Y., et al. (2011). A neural signature of hierarchical reinforcement learning. *Neuron* 71, 370–379. doi: 10.1016/j.neuron.2011.05.042
- Rosenbaum, D. A., Kenny, S. B., and Derr, M. A. (1983). Hierarchical control of rapid movement sequences. *J. Exp. Psychol. Hum. Percept. Perform.* 9, 86–102. doi: 10.1037/0096-1523.9.1.86
- Rosenberg, R. D., and Feigenson, L. (2013). Infants hierarchically organize memory representations. *Dev. Sci.* 16, 610–621. doi: 10.1111/desc.12055
- Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., and Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cereb. Cortex* 17, 1147–1153. doi: 10.1093/cercor/bhl024
- Russo, E., Nambodiri, V. M. K., Treves, A., and Kropff, E. (2008). Free association transitions in models of cortical latching dynamics. *New J. Phys.* 10:015008. doi: 10.1088/1367-2630/10/1/015008
- Russo, E., and Treves, A. (2011). An uncouth approach to language recursivity. *Biolinguistics* 5, 133–150. Available online at: <http://www.biolinguistics.eu/index.php/biolinguistics/article/view/186>
- Saari, D. G. (1995). *Basic Geometry of Voting*. Berlin: Springer. doi: 10.1007/978-3-642-57748-2
- Sahin, N. T., Pinker, S., Cash, S. S., Schomer, D., and Halgren, E. (2009). Sequential processing of lexical, grammatical, and phonological information within Broca's area. *Science* 326, 445–449. doi: 10.1126/science.1174481
- Sakai, K., Kitaguchi, K., and Hikosaka, O. (2003). Chunking during human visuo-motor sequence learning. *Exp. Brain Res.* 152, 229–242. doi: 10.1007/s00221-003-1548-8
- Schneider, D. W. (2007). Task-set inhibition in chunked task sequences. *Psychon. Bull. Rev.* 14, 970–976. doi: 10.3758/BF03194130
- Sompolinsky, H., and Kanter, I. I. (1986). Temporal association in asymmetric neural networks. *Phys. Rev. Lett.* 57, 2861–2864. doi: 10.1103/PhysRevLett.57.2861
- Stone, E., and Armbruster, D. (1999). Noise and O(1) amplitude effects on heteroclinic cycles. *Chaos* 9, 499–506. doi: 10.1063/1.166423
- Tettamanti, M., and Weniger, D. (2006). Broca's area: a supramodal hierarchical processor? *Cortex* 42, 491–494. doi: 10.1016/S0010-9452(08)70384-8
- Treves, A. (2005). Frontal latching networks: a possible neural basis for infinite recursion. *Cogn. Neuropsychol.* 22, 276–291. doi: 10.1080/0264329042000329
- Tsuda, I. (2009). Hypotheses on the functional roles of chaotic transitory dynamics. *Chaos* 19, 015113. doi: 10.1063/1.3076393
- Varona, P., Rabinovich, M. I., Selverston, A. I., and Arshavsky, Y. I. (2002). Winnerless competition between sensory neurons generates chaos: a possible mechanism for molluscan hunting behavior. *Chaos* 12, 672–677. doi: 10.1063/1.1498155
- Venaille, A., Varona, P., and Rabinovich, M. I. (2005). Synchronization and coordination of sequences in two neural ensembles. *Phys. Rev. E Stat. Nonlin. Soft. Matter Phys.* 71:61909. doi: 10.1103/PhysRevE.71.061909

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 18 October 2013; accepted: 10 February 2014; published online: 14 March 2014.

Citation: Rabinovich MI, Varona P, Tristan I and Afraimovich VS (2014) Chunking dynamics: heteroclinics in mind. *Front. Comput. Neurosci.* 8:22. doi: 10.3389/fncom.2014.00022

This article was submitted to the journal *Frontiers in Computational Neuroscience*. Copyright © 2014 Rabinovich, Varona, Tristan and Afraimovich. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.