# Research Summary (2000 ~ 20010)


## Jun Tani

Lab. for Behavior and Dynamic Cognition, RIKEN BSI


## 1. Introduction

The main objective of our research has been to understand the mechanisms of higher-order cognitive brain functions with particular focus on the problems of compositionality. Compositionality means that the whole can be derived by combining reusable parts; for example, complex goal-directed actions can be generated by combining reusable behavior primitives into specific sequences according to Arbib's motor schemata theory. Although it is considered in conventional cognitive science that compositionality can be achieved by means of symbol representation and manipulation of the representations at a higher cognitive level, our motivation has been to investigate an alternative process whereby compositionality could be developed in the neural dynamics of distributed activity through accumulative learning of sensory-motor experiences.

With this motivation, we have conducted a series of neural modeling studies at the connectionist level, inspired by related evidence obtained from cognitive neuroscience studies. The models have been examined in robotics experiments for the purpose of exploring novel phenomena appearing in the interaction between neuro-dynamics and physical actions which could provide us new insights to understanding nontrivial brain mechanisms.

We first studied how a set of sensory-motor patterns can be learned as behavior primitives with distributed representation by proposing a model recurrent neural network with parametric biases (RNNPB). Our robotics experiments clarified its capability for *generalization* in learning and also showed how learned behavior primitives can be adaptively shifted in accordance with dynamic environmental changes by means of "regression" of the current sensory inputs.

As a next step, we investigated how a set of acquired behavior primitives can be manipulated in a goal-directed manner with utilizing hierarchy. Our robotics experiments showed that a *functional hierarchy* can be self-organized by utilizing multiple timescale dynamics in a proposed network model such that the fast dynamics part acquires primitive patterns and the slow dynamics part learns a sequential combination of them for each action goal. Its developmental processes were detailed.

Moreover, we studied possible mechanisms of the further higher level cognitive functions assumed for the prefrontal cortex (PFC), which include the formation of abstract concepts and executive control of acquired rules. We examined how linguistically represented concepts can be learned through associative learning of a set of sentences and corresponding action patterns by proposing an extension of the RNNPB model. The experimental results obtained with a minimal setup showed that a set of concepts composed of verbs and object nouns can be developed with generalization even when there is a "poverty of exemplars". Finally, our simulation experiment suggested that mechanisms of the executive control for rule switching and self-monitoring of confidence can be developed by self-organizing the necessary functional hierarchy presumed between the PFC and the posterior cortices. The next section describes details of the research achievements which elucidate that the higher cognitive mechanisms which are developed in the dynamics of distributed neural activities can afford *generalization* and *fluidity* as well as compositionality.

## 2. Research achievements

**Learning behavior primitives** First, we focused on the problem of how a set of behavior primitives which frequently appear in relation to a particular sensory-motor sequence pattern can be learned for generation as well as recognition. To address the problem, we [Tani99] initially proposed a model of gated modular networks focusing on its local representation scheme (Wolpert and Kawato independently proposed a similar idea in their MOSAIC model in the same period.) The goal was to learn to store each behavior primitive in a specific local modular network as segmented from experiences of continuous sensory-motor flow by means of associated gate operations. The modular networks compete to predict the current sensory-motor sequence pattern and the gate of the winner—determined by minimum prediction error—opens exclusively. The winner is then entitled to learn to become an expert for the pattern.

Although the scheme worked successfully in a simple simulation experiment of navigation learning, it could not be scaled easily to more complex problems dealing with robots with larger degrees of freedom (DOF). There is potential instability in the gate opening as the result of miss-categorization of similar patterns due to competition from multiple modules. Also, the scheme cannot provide sufficient generalization among learned patterns because the patterns are learned separately in local modules.

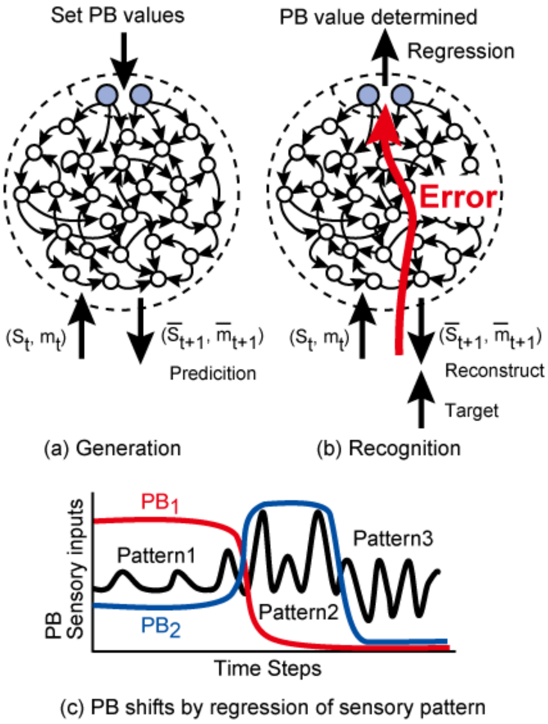

Fig1. RNNPB

Therefore, as a next step, we examined an alternative scheme based on distributed representation and proposed the RNNPB model [Tani03b]. RNNPB is a discrete time RNN associated with PB units in which a set of sensory-motor sequence patterns $(s_t, m_t)$ can be learned in a distributed way with self-determining a specific PB vector value for each of them. The learning is conducted by means of back-propagation through time (BPTT) algorithm in order to determine the optimal synaptic weight matrix as well as a specific PB value corresponding to each of learned patterns. The learned sensory-motor sequence patterns can be regenerated by setting corresponding PB values (Fig.1a). Also, given sensory sequence patterns can be recognized by means of the PB regression scheme; the optimal PB value for reconstructing the target pattern is searched by back-propagating the error generated between the target sequence and the reconstructed one to the PB for its update (Fig.1b). Thus, the RNNPB can be regarded as a generative model which can both generate and recognize learned sensory-motor sequence patterns. Our experiment [Tani03b] showed that a set of discrete movements and periodic movements can be learned simultaneously as fixed point attractors and limit cycling ones, respectively. PB can be regarded as a bifurcation parameter shifting attractor structure from one to another. It was also found that similar patterns are learned with similar PB values. This characteristics can provide generalization capability to the system. [Tani04b] showed that principal feature dimensions can be extracted in PB vector dimensions. These results suggest that the system has some generalization capability in learning.

By utilizing the PB regression mechanism in an on-line way, robots can adapt to dynamically changing situations both in social and environmental contexts while adequately altering their own behavior patterns. As an example in social contexts, a humanoid robot with 8 DOF arms was tutored to imitate multiple movement patterns demonstrated by a human experimenter [Ito04]. More specifically, the robot, facing with the experimenter, learned to predict the experimenter's movement patterns for both hands perceived through vision and also to generate its own corresponding arm movement patterns in terms of motor sequences. In the imitation game after the learning, the robot could successfully follow sudden shifts in patterns from currently engaged ones to others, as demonstrated by the experimenter, by adequately shifting the PB values (Fig1.c). In this model, the function role of PB corresponds to that of mirror neurons found in the ventral premotor cortex of monkeys by Rizzolatti's group because the same PB activation accounts for both recognizing others' particular movement patterns and generating own corresponding movement patterns. The autonomous shifting of the PB by the error regression observed in this experiment may account for neuronal mechanisms for reading of other's intentions.

Another illustrative example is about adaptation to situational changes in environment. The robot was tutored for two different ball-play patterns, one for laterally rolling a ball and the other for moving the ball up and down with visual perception of the ball position [Ito06]. As shown in video-1[1], the ball-play patterns can switch from one pattern to another intermittently, triggered by irregular ball

[1] All videos can be found at http://bdc.brain.riken.jp/~tani/STL09 with *ID* "brain" and *password* "tani"

movements. The prediction error generated by these irregular situational changes causes modulation of the PB values by the regression, which results in the autonomous shifting of the currently engaged play patterns from one to another.

**Self-organization of functional hierarchy** We next investigated how a set of acquired behavior primitives can be combined in sequences to achieve various goal-directed actions. For this purpose, we proposed a continuous time RNN model, the so-called multiple timescale RNN (MTRNN) [Yama08b], which consists of a fast dynamics part with a smaller time constant $\tau$ in the lower level and a slow dynamics part with a larger $\tau$ in the higher level (Fig.2a), where the activation dynamics of each neural unit $a_i$ can be described as:

$$\tau \cdot \dot{u}_i = -u_i + \sum w_{ij} a_j + I_i \,, \quad a_i = sigmoid(u_i) \quad (1) \quad \text{where } u_i \text{ is the potential.}$$

We consider that the slow dynamics part may correspond to the PFC and the premotor as it is often observed that the preparatory period required to build up neural activation in these areas can take a relatively long time, of the order of seconds. The fast dynamics part, on the other hand, may correspond to the posterior cortices, especially the inferior parietal cortex (IPL). Usually, the build-up in the posterior cortical neurons is much faster, within much less than a second. The lower level network receives current visuo-proprioceptive (VP) state (vt, pt) as inputs and generates its prediction outputs (vt+1, pt+1) for the next time step. Here, proprioceptive state essentially means the current body posture in terms of joint angle positions for robots. The network can generates a variety of VP sequences depending on the initial states set in the higher level network by utilizing initial sensitivity characteristics. The network is trained to regenerate a set of VP sequences by determining optimal synaptic weights of the whole network and specific initial states of some of the higher level units corresponding to each of the sequences. The initial states of other neural units are set as neutral. Also, the network can generate motor imagery of the learned sequences without receiving real sensation but by feeding its own prediction to the next sensory inputs by closing the loop. A particular assumption in this model is that the IPL might generate prediction for coming multimodal sensations by receiving the top-down intention for a particular action program from the frontal cortices [Nishi09].

The MTRNN was tested in an experiment involving developmental tutoring of multiple tasks that dealt with object manipulation by a humanoid robot. Three different task sequences were considered, each of which was composed of sequences of various behavior primitives including reaching for the object, lifting it up and down (UD), moving it to the left and right (LR), moving it forward and backward (FB), and rotating it (RO)[Nishi09]. Fig.2b shows successful regenerations of the three tasks achieved at the end of the iterative tutoring where the VP sequence and activation dynamics of the slow units and the fast units are plotted for each task sequence (also see video-2). It is observed that the slow dynamics generates different profiles started from each distinct initial state which result in different sequential combinations of the primitives. Therefore, it is considered that the initial state may represent an abstract action program. Also, note that the fast dynamics correlates closely with the VP

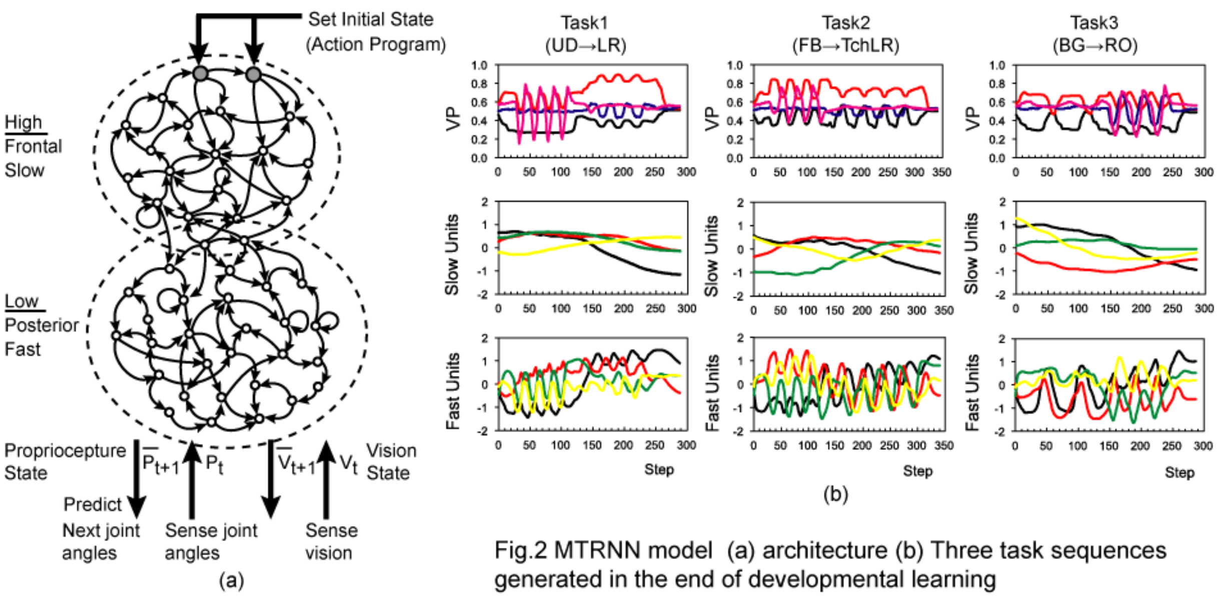


Fig.2 MTRNN model (a) architecture (b) Three task sequences generated in the end of developmental learning

sequences. These observations as well as other analytical results [Yama08b] confirmed that a set of behavior primitives are acquired embedded in the fast dynamics and that the slow dynamics in the higher level interacts differently with the fast dynamics in the lower level depending on its initial state, which results in the generation of various sequential combinations of the primitives. The role of slow dynamics might be analogous to the shifting of the PB in RNNPB. Our analysis confirmed that the difference in the time constants between the two parts is essential for the self-organization of this type of functional hierarchy. This self-organization mechanism might be general, and regardless of particular learning schemes such as BPTT, because we showed that adaptation by genetic algorithm can also develop similar functional hierarchy by utilizing timescale differences [Paine05].

The observation of the development processes during iterative tutoring revealed some interesting results. It was shown that motor imagery develops faster than physical action and also that primitives in the lower level develop faster than the sequencing of them in the higher level. These observations correspond to some empirical studies of infant development as described in [Nishi09].

Furthermore, it was shown that MTRNN can generate various action plans by combining behavior primitives to achieve desired goal states [Arie09]. Action programs in terms of the initial states are searched such that distal states in imagery VP sequences generated from the initial states can match with the desired goal states. It was shown that such planning can generate even novel combinations of learned episodic sequences. This experiment is shown in video-3.

Finally, it should be mentioned that the model achieved more than just compositionality of combining primitives. Elaboration between the slow and the fast dynamics during the learning process enabled quite smooth transitions between one primitive to another, rather than just connecting discrete objects of primitives. Moreover, the slow dynamics carries contextual information like current counts of cycle times for periodic movements as well as goals of heading. Although this counting of cycle times was sometimes imprecise (±1) in our experiment, the transitions never occur in the middle of performing primitives. When combinatorial action sequences can be generated to carry context, produce smooth transitions of the primitives and achieve generalization, the appeared structures are not just compositional but also as fluid and “organic” [Tani08b, Nishi09].

**Associative learning of language and action** This study focused on the problem of how linguistically represented concepts can be acquired via behavioral experiences. For this purpose, we examined how a set of simple sentences consisting of verbs and object nouns can be understood and the corresponding actions be produced by robots by extending the RNNPB scheme [Sugi05]. The new model consists of a linguistic module and an action module which are bound by PB units (Fig.3a). For the association of sentences and corresponding actions, the network is trained under the constraint that the PB values for generating the word sequences and those for generating the corresponding actions should become identical. This can be conducted by back-propagating the prediction errors from both modules to the PB units shared by both. After the learning converges for all the pairs, the capability of understanding sentences can be

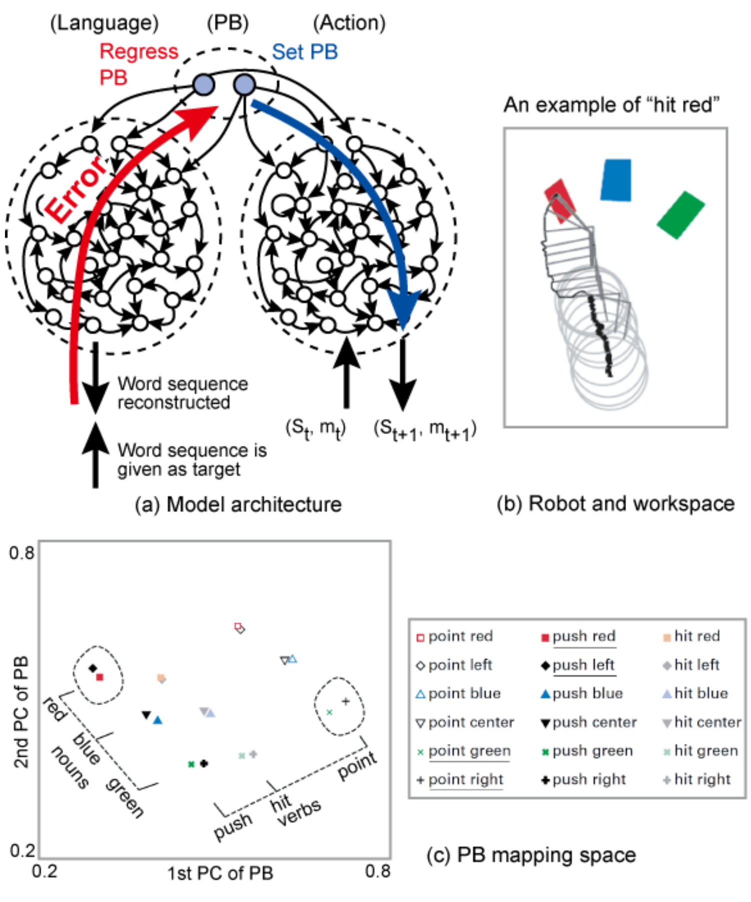

Fig3. Language action association learning

tested as follows. A particular word sequence is shown to the linguistic module as a target to be reconstructed by the PB regression scheme. Then, the PB value obtained as the result of recognizing the word sequence is used to activate the action module in order to generate prediction of the corresponding sensory-motor sequence, and the resultant robot behavior is observed. The scheme was evaluated through the following robotics experiment.

A physical mobile robot equipped with a one DOF arm and vision was placed in a workspace where red, blue and green objects were always located left, center, and right of the robot, respectively (Fig.3b). A set of sentences consisting of 3 verbs (point, push, hit) and 6 object nouns (left, center, right, red, blue, green) were considered. For example, "push red" means that the robot is to move to the red object and push it with its body and "hit left" means that the robot is to move to the object to its left and hit with its arm. Note that "red" and "left" are synonymous in the workspace setting, as are "blue" and "center" and "green" and "right". For given combinations of verbs and nouns, corresponding actions in terms of sensory-motor sequences of more than 100 steps are tutored by guiding the robot while introducing slight variance of object positions at each trial. In order to investigate generalization capability in the learning, only 14 out of 18 possible sentences were trained.

As a result of the association learning, it was found that the robot could generate correct actions for all 18 sentences. We examined how those sentences are mapped to the PB vector space. Fig.3c shows this sentence mapping to the PB space with its two principal components. Observe that the mapping appears with a 2-dimensional grid structure with one dimension for verbs and the other for nouns, where all sentences with the same verbs followed by synonymous nouns appear close in the space. It is noted that even some sentences with unlearned combinations such as "push red|left" and "point green|right" are mapped to adequate positions in the grid (indicated by dotted circles). This result implies that the concepts are acquired as generalized with their compositional structures self-organized in the distributed neural representation via the associative learning even when there is a "poverty of exemplars". The result also suggests importance of the hub-like connectivity by the PB units which may correspond to the Broca's area that may bind language related information from the middle temporal cortex and action related information from the premotor and motor cortices. The concepts or meaning may emerge in interactions among multiple modalities through such hub connectivity. (See video-4 for the related robotics experiment.)

**Executive controls for rule switching and meta-cognition for confidence** We are interested in how the PFC can handle executive controls of rule switching while monitoring confidence for it because this topic should involve problems that challenge higher cognitive brain mechanisms. To this end, we explored the possible neuronal mechanisms for a rule switching task similar to the Wisconsin card sorting test (WCST) with betting options for own successful action outcomes using neuro-robot simulations [Mani09a,b,c]. A robot acquires three different behavioral response rules to a stimulus (Fig.4a). The same rule repeats for successive trials, but it can be changed unpredictably, similarly to an ordinal WCST. The robot is rewarded in the successful trials and punished in the failure trials, and it has to recognize the rule shifts from such feedback. Additionally, the robot can bet on its success at each onset of a trial by outputting an adequate betting rate. The higher the betting rate, the greater the gain (loss) of fitness in the case of success

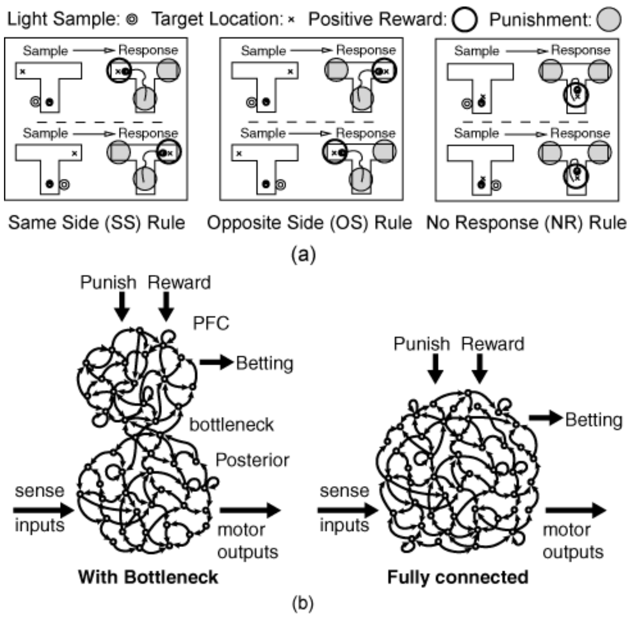


Fig.4 Simulation set up. (a) Three rules were used: the robot should enter the branch on the same side as the light turning on under the "same side" (SS) rule; it should move in the opposite direction under the "opposite side" (OS) rule; and it should remain near the start position independent of the side of the light source in the "no response" (NR) rule. (b) CTRNNs with bottleneck and fully-connected architectures. In the network with a bottleneck, the lower part receives light and range sensor inputs and it outputs motor commands, whereas the higher level receives reward and punishment signals in their corresponding zones and it outputs betting rate. There are no input/output related segregations in the fully-connected case.

(failure). We employed a CTRNN model with bottleneck connectivity added between the PFC and the posterior parts (Fig.4b). In this model, the time constant of all units are set with the same value. We then compared task performance in cases with and without a bottleneck while keeping the number of neural units the same. The synaptic weights of the networks are modulated in the direction of the larger fitness in the repeated trials during the task acquisition phase by using a genetic algorithm.

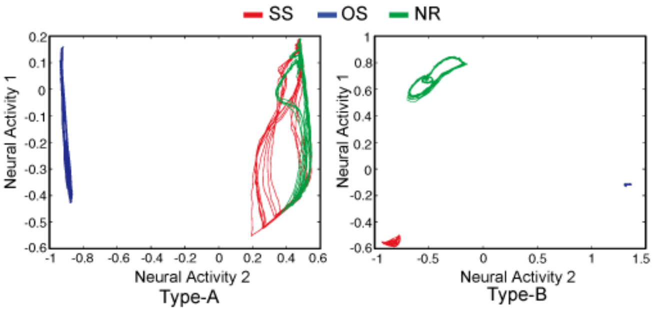

Fig.5 Phase plots of the neural activity generated in Type-A and Type-B cases. The axes are the two principal components of populations in the higher level. The attractor corresponding to each rule is drawn with a different color. In Type-A, some overlaps appear between SS and OS plots because of the similarity in their rules. On the contrary three attractors appear as distinct ones in Type-B.

The simulation results showed that cases with a bottleneck significantly outperformed those without. This suggests that a certain amount of information segregation between the PFC and the posterior cortex may enhance the task performance. It was observed that three attractors representing three different rules appear in the collective neural activities in the PFC part and that rule switching is achieved by state transition from the current attractor to another as triggered by the punishment signal (Fig.5). It was also observed that the betting is generated by achieving adequate mapping from this neural state. If the neural state is within one of the attractors which is following the current rule correctly, the betting rate is mapped to the higher values, and otherwise to the lower values, which indicates that each basin of the attractor becomes the region of the high confidence generated with correct executions of the rule. This could be a general mechanism for realizing executive control for rule switching since only this type of dynamic mechanism was found to be self-organized in multiple runs of simulations with a simple, general network model described. Furthermore, a nontrivial finding of this study was that attractor-based encoding by evolution produced two types of rule representation: an analogical representation (Type-A) that incorporates similarity between rules, and a distinct representation (Type-B) for the same set of rules (Fig.5). The qualitative difference in the geometrical arrangements of the attractors in these two types of rule representation causes different levels of stability in rule execution, which results in different meta-cognitive characteristics of self-confidence about the rules. This may account for the everyday psychological observation that we sometimes acquire rules as tacit knowledge (Type-A) and at other times as explicit, symbolic knowledge (Type-B).

**Neuro-phenomenology** Finally, we also studied the problems of consciousness and "self" at the intersection between neuroscience, nonlinear dynamics and phenomenology. In particular, our studies focused on the interactions between the top-down pathway of predicting future which originates from the frontal cortices and the bottom-up pathway of regressing past experience which proceeds through the posterior cortices. Our robotics experiments suggested that consciousness may arise due to conflicts generated between these two pathways where "authentic self" in terms of Heidegger's notion might appear as dynamical states of *self-organized criticality* as discussed in [Tani09].

**Summary** The abovementioned research results reveal that the higher cognitive brain functions accounting for compositional action generation, compositional concept formation and executive control for rule switching can be developed by means of self-organizing necessary dynamic structures in distributed neural activity rather than by employing a symbolic representation framework. In the case of long-term consolidation learning of sensory-motor experiences, the structures developed are not just compositional but are also generalized and fluid. Although any link to the neurophysiology of such a higher-order cognitive mechanism is necessarily speculative, we consider that this type of abstract computational modeling is very important for outlining the possible underlying brain mechanisms and furthermore, for extracting their organizing principles.

## 3. References


[Tani09] Tani J.: "Autonomy of 'Self' at criticality: The perspective from synthetic neuro-robotics.

*Adaptive Behavior*, Vol.17, No.5, pp.421-443, 2009.
[Nishi09] Nishimoto R., Tani J.: “Development of hierarchical structures for actions and motor imagery: a constructivist view from synthetic neuro-robotics study”, *Psychological Research*, Vol.73, pp.545-558, 2009.
[Arie09] Arie H., Endo T., Arakaki T., Sugano S., Tani J.: “Creating novel goal-directed actions at criticality: a neuro-robotic experiment”, *New Mathematics and Natural Computation*, Vol.5, No.1, pp.307-334, 2009.
[Mani09c] Maniadakis M., Trahanias P. and Tani J.: “Self-organizing neural mechanisms for higher-order cognitive functions in human prefrontal cortex”, Submitted.
[Mani09b] Maniadakis M., Trahanias P., Tani J.: “Explorations on artificial time perception”, *Neural Networks*, Vol.22, pp.509-517, 2009.
[Mani09a] Maniadakis M., Tani J.: “Acquiring rules for rules: neuro-dynamical systems account for meta-cognition”, *Adaptive Behavior*, Vol.17, No.1, pp.58-80, 2009.
[Yama08b] Yamashita Y., Tani J.: “Emergence of functional hierarchy in a multiple timescale neural network model: a humanoid robot experiment”, *PLoS Computational Biology*, Vol.4, Issue.11, e1000220, 2008
[Tani08b] Tani J., Nishimoto R., Paine R.: “Achieving ‘organic compositionality’ through self-organization: Reviews on brain-inspired robotics experiments”, *Neural Networks*, Vol.21, pp.584-603, 2008.
[Ito06] Ito M., Noda K., Hoshino Y., Tani J.: “Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model”, *Neural Networks*, Vol.19, pp.323-337, 2006.
[Paine05] Paine R.W. and Tani J.: “How hierarchical control self-organizes in artificial adaptive systems”, *Adaptive Behavior*, Vol.13, No.3, pp.211-225, 2005.
[Sugi05] Sugita Y., Tani J.: “Learning semantic combinatoriality from the interaction between linguistic and behavioral processes”, *Adaptive Behavior*, Vol.13, No.1, pp.33-52, 2005
[Ito04] Ito M., Tani J.: “On-line imitative interaction with a humanoid robot using a dynamic neural network model of a mirror system”, *Adaptive Behavior*, Vol.12, No.2, pp.93-115, 2004.
[Tani04b] Tani J., Ito M., Sugita Y.: “Self-organization of distributedly represented multiple behavior schemata in a mirror system: Reviews of robot experiments using RNNPB”, *Neural Networks*, Vol.17, pp.1273-1289, 2004.
[Tani03a] Tani J.: “Learning to generate articulated behavior through the bottom-up and the top-down interaction processes”, *Neural Networks*, Vol.16, No.1, pp.11-23, 2003
[Tani03b] Tani J., Ito M.: “Self-organization of behavioral primitives as multiple attractor dynamics: A robot experiment”, *IEEE Trans. on SMC. Part A*, Vol.33, No.4, pp.481-488, 2003.
[Tani99] Tani J., Nolfi S.: “Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems”, *Neural Networks*, Vol.12, pp.1131-1141, 1999.