Complementary Working Memories using Free-Energy Optimization for Learning Features and Structure in Sequences

Alexandre Pitti[1*], Mathias Quoy[1], Catherine Lavandier[1] and Sofiane Boucenna[1]

Abstract— **We propose a global framework for modeling the cortico-basal system (CX-BG) and the fronto-striatal system (PFC-BG) for the generation and recall of audio memory sequences; ie, sound perception and speech production. Our genuine model is based on the neural architecture called INFERNO standing for Iterative Free-Energy Optimization of Recurrent Neural Networks. Free-energy (FE) corresponds to the prediction error on internal or external noise. FE minimization is used for exploring, selecting and learning in PFC the optimal choices of actions to perform in the BG network (eg sound production) in order to reproduce and control the most accurately possible the spike trains representing sounds in CX. The difference between the two working memories relies in the neural coding itself, which is based on temporal ordering in the CX-BG networks (Spike Timing-Dependent Plasticity) and on the rank ordering in the sequence in the PFC-BG networks (gating or gain-modulation). We detail in this short paper the CX-BG system responsible to encode the audio primitives at few milliseconds order, and the PFC-BG system responsible for the learning of temporal structure in sequences. Two experiments done with a small and a big audio database show the capabilities of exploration, generalization and robustness to noise of the neural architecture to retrieve audio primitives as well as long-range sequences based on structure detection. Although both learning mechanisms are implemented with the same algorithm of rank-order coding, the CX-BG system realizes a model-free recurrent neural network (INFERNO) and the PFC-BG system implements a gated recurrent neural network (INFERNO GATE).**

I. Introduction

In different brain areas, working memories (WMs) are hypothesized to embed neural processes with forward and inverse models that can encode, anticipate and eventually control incoming signals to be more robust and to overcome their variability [1]. Two brain areas namely the Basal Ganglia (BG) that selects actions with respect to current states [2] and the Prefrontal Cortex (PFC) that represents forthcoming actions with respect to current contexts [3], are important for embedding these WMs. Being part of two different loops but connected at the BG level, they realize reactive (BG) and proactive (PFC) control, processing information differently and at different speed, see Fig. 1.

II. Related works

On the one hand, some evidences indicate that the striatum in BG has a principal function in learning-related plasticity associated with selecting one set of actions from many, resulting in the acquisition of habitual behavior [4]. On the other hand, PFC achieves behavioral planning in terms of the end result, rather than in terms of the movement required to perform the task [5].

Graybiel and Grafton suggest in [6] that proactive control is associated with sustained and/or anticipatory activation of lateral PFC, which reflects the active maintenance of task goals. By contrast, reactive control should be reflected in transient activation, along with a wider network of additional brain regions such as the BG. Therefore, these two control mechanisms differ in terms of the involvement during learning and retrieving tasks or sequences, with the BG dynamics working at a faster pace than the PFC.

In Machine Learning, reactive and proactive control relate to what is called model-free and model-based systems in Reinforcement Learning (RL) [7], [8], having one system for stimulus-response tasks doing greedy-like optimization and the other learning distinct policies for prediction, which serves for planning goal-directed behaviors.

These features are linked therefore to what is called now the Bayesian theory of the brain [9] and to the paradigm of predictive coding for cognition [10]. These general theories describe how our expectations (as well as our errors) on sensory inputs are used as attention signals to adjust the prior expectations for the next events. Brain areas are hypothesized to use error prediction as a core information to *control* their dynamics from each others, not just for binding them mutually.

Under this framework, two or more brain networks can interact dynamically (eg Cortex CX with Basal Ganglia BG) so that we have always one network (eg the controller) that infers the reliability of another (eg the observer) with respect to a specific context. Along with Bayes theory, predictive coding has also its link with optimal control theory [11], which we think interesting in terms of perspective for modeling the corticostriatal system as it moves the problem of learning and retrieving memory sequences into a control problem.

Problem-solving tasks are good examples for understanding the involvement of the BG-PFC loops in goal-directed behaviors under uncertainty especially during infancy. These goal-directed behaviors are also called *task-sets* in cognitive and developmental sciences [12]. Task sets relate to the novel capabilities acquired by twelve-months-old babies such as tool-use, sustained attention, spatial memory, asymmetric imitation and rule-based learning and are argued to be linked to the maturation of the PFC [13], [14]. Other crucial examples during infancy are speech production and the sequential organization of actions [15], [16], [17]. These two important cognitive tasks presumably involve the BG and PFC loops

[1] University Paris-Seine, University of Cergy-Pontoise, ENSEA, CNRS, France `surname.name@u-cergy.fr`

to adjust timely and orderly motor primitives [18] and [19].

This neural process has been particularly studied for speech and language sequences because auditory modality is the sense especially sensitive to temporal structure. In the case of speech production, Romanski and colleagues propose that the phonotopical level requires the implementation of high-order models for encoding words or sentences as articulatory vocal tracks [20]. In other experiments with 3 months-old children [15], [16], [17], the stronger activation of the PFC has been observed for detecting temporal dissonance in regular temporal structures of spoken sequences of totally random syllabes such as the ABA structure in “tomato” or “mifumi”, independent of the syllabes pronounced for the A or B items [21], [22].

A. Proposal framework for feature extraction and structure learning in sequences

In line with these finding, we propose a neural architecture to model the CX, BG and PFC systems that combines model-free and model-based learning for retrieving and controlling long-range memory sequences hierarchically at the signal level and at the abstract level, see Figs. 1. The two working memories are developped within the same framework of predictive coding and reinforcement learning [23] but each system is working differently to code information and to minimize online error and external noise. The models use spiking neural networks (SNN) in order to learn temporal delays between pre- and post-synaptic firing neurons with the mechanism of Spike Timing-Dependent Plasticity (STDP) [24]. In line with the framework of free-energy minimization [10], we exploit also intrinsic noise within the system in order to realize stochastic descent gradient and novelty detection.

We propose that these neural mechanisms can serve for the learning of temporal delays between neurons in a self-organizing manner and makes possible the discovery of causes and effects necessary for active inference and predictive coding. This extends previous researches in which we developped several models of WMs using SNNs corresponding to different brain areas. These models exploited noise or novelty to iteratively infer a model and minimize error prediction either to control one system’s dynamics (eg the hippocampus or BG-like model-free networks [25], [26]) or to select dynamically a better controller (eg a PFC-like model-based network [27]).

In our BG-like network modeled in [26] and [28], we showed that it is possible to control long-range memory sequences of spikes –, above 1000 iterations without loss,– and to solve the so-called credit assignment problem by infering causes and effects, even with long-range delays. Because of its property to optimize and control dynamics iteratively using noise or free-energy, we named our network INFERNO, which stands for Iterative Free-Energy Optimization for Recurrent Neural Networks [26].

Our framework will be applied to speech learning (perception and production). The global framework combines the corticostriatal and prefrontal systems for the recognition and generation of audio memory sequences, see Fig. 1. In the first part, the WM of the cortico-striatal (CX-BG) system is developped in order to better describe the process for retrieving audio primitives for a short time scale. The model named INFERNO network is then used to solve the credit assignment problem for retrieving the motor primitives that cause specific sound signal (vocal tracks). In principle, the motor primitives should be the vocal articulatory motions of the mouth and of the vocal cords, or the sound generated by a vocoder but in our case it is simply the sound vector reconstruction. In the second part, the combination with the prefrontal system (PFC) will be presented, with the use of Rank-Order neurons for learning the temporal organization within memory sequences and for predicting the next ones; the sensitivity of Rank-Order or gated neurons to the items’ order within a sequence will serve for finding structure within signals. We named this gated version of the INFERNO network, INFERNO GATE and we have showed that it is possible to retrieve long range sequences through iterative optimization for long time scale. More detailed explanations about these two models are presented in extented papers [28], [29].

B. Neural model for the cortico-striatal system and the fronto-basal system

In our comprehension of the free-energy optimization strategy proposed by Friston [30], [31], [23], it is similar to a reinforcement learning process without the need of value-functions to minimize online error prediction. To us, it conveys the learning problem into the ones of optimal control and predictive coding. We can apprehend the cortico-striatal system and the fronto-basal system as two learning systems that attempt to perform an optimal control and resolve error prediction among their dynamics. In Fig. 2, we display the first network, INFERNO, with the Primary Auditory Cortex (PAC) system and the Superior Temporal Gyrus (STG) layer modeled with SNNs to encode incoming inputs, the Striatum layer that categorizes the state of the STG dynamics and the Globus Pallidus that attempt to control back the input dynamics of the PAC and STG with a reentrant loop. The error prediction is evaluated and minimized over time by supervision of the STR units (critic) and by noise generation and stochastic search done on the GP output layer (actor).

This paper is organised as follows. Two experimental setups for sound sequences are presented in section IV, respectively for a limited learning database (only one speaker, 3 minutes length) and for a larger database (several speakers of different genders, 30 minutes length). The results of these two experiments are developped and discussed in section V.

III. Methods

The neural architecture INFERNO [26] consists of two coupled learning systems arranged. The first network corresponds to one recurrent neural network of spiking neurons (SNNs) and the second network consists on one associative map. The SNN implements a forward model of the incoming signals whereas the associative map implements an inverse

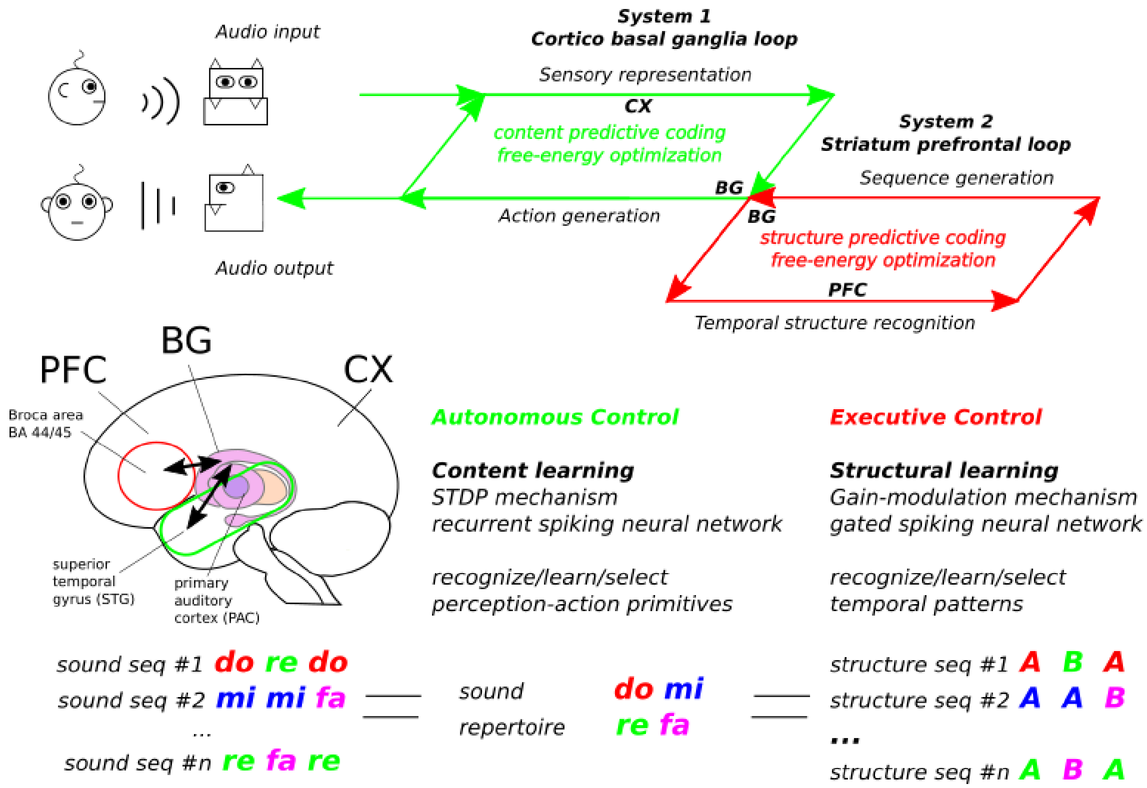

Fig. 1. Framework for sequence learning based on iterative optimization. Cortico-basal (CX-BG) and Fronto-striatal (PFC-BG) loops.

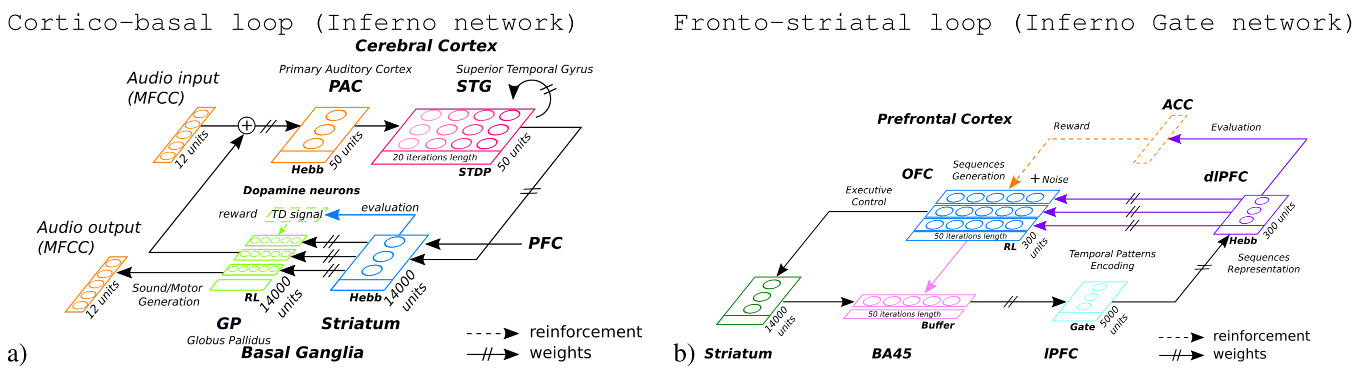

Fig. 2. In a), framework of the INFERNO architecture for audio primitive retrieving based on iterative optimization through cortico-basal ganglia loop (CX-BG) [26], [28]. The Primary Auditory Cortex (PAC) receives and categorizes the audio vectors as a first stage, the Superior Temporal Gyrus (STG) integrates over time its output that are categorized at the end by the Striatum (STR) in the basal ganglia. The Globus Pallidus (GP) searches and retrieves the audio vectors that best matches the STG dynamics recognized by the striatal units. The iterative optimization process is done by minimizing noise with a temporal difference reinforcement signal. In b), framework of the Inferno Gate architecture for structure learning based on iterative optimization through prefrontal-basal ganglia loop (PFC-BG) [29]. The Broca area (BA45) receives sequences, the lateral Prefrontal Cortex (lPFC) learns and detects temporal patterns in sequence, the dorsal Prefrontal Cortex (dlPFC) categorizes and represents sequences as a distribution of temporal primitives. The Orbito-Frontal Cortex (OFC) searches and retrieves the sequences that best match the BA45 sequences recognized by the lPFC units. The iterative optimization process is done by minimizing noise with a temporal difference reinforcement signal in ACC.

model aimed at retrieving and controlling those signals. The inverse-forward controller can be modeled with the function $Y_{out} = f(I)$ for the SNN and with the function $I = g(Y_{out})$ for the associative map, in which I is the input vector and Y_{out} are the output dynamics.

In order to minimize error, the second network generates intrinsic noise I_{noise} to control the dynamics of the first one following a RL mechanism. The activity of the SNN Y_{out} is compared to one desired goal vector Y_{des} to compute the error E between Y_{des} and Y_{out} and the current input is kept for the next step $I(t+1) = I(t) + I_{noise}$, if and only if it diminishes the gradient ΔE. Over time, I converges to I_{opt} its optimum value, and Y_{out} converges to Y_{des} the desired vector. This scheme is in line with actor-critic algorithms and predictive coding. Its organization is similar to novel architectures combining two or more competitive neural networks such as auto-encoders or the generative adversarial networks.

We showed in [26] that this variational process is similar to a stochastic descent gradient algorithm performed iteratively and can solve the temporal credit assignment problem for delays above dizains of iterations. For instance, the convergence to the desired goal after a certain delay can be viewed as the retrieval of a memory sequence for such duration. Furthermore, the free-energy minimization is generative in the sense that it can retrieve novel solutions I for the same output Y. This can be viewed as a synchronization process toward attractor memories [32].

IV. Results

A. Experiment 1 – self-supervised vocal babbling

In this experiment, we make to learn the Primary Auditory Cortex (PAC), STG and Striatum layers in an unsupervised manner so that the three structures self-organize to sparse distributions using Hebb law for the PAC and the Striatum whereas the STL learns the temporal dependencies across time using the STDP learning mechanism; the direction of the information flow is PAC→STG→STR.

The experimental setup for Experiment 1 in section IV-A consists on a small audio dataset of 2 minutes length of a native french woman speaker repeating three times five sentences. The audio .wav file is translated into MFCC vectors (dimension 12) sampled at 25ms each and tested either with a stride of 10ms and no stride. The whole sequence represents 14.000 MFCC vectors for the case with strides and 10.000 MFCC vectors for the case with no strides. The number of Striatal and GP units are chosen so that they correspond to the number of MFCC vectors, which means 14000 units (10.000 units) for each layer. We do so in order to test the reliability of our architecture to retrieve input data with an orthogonal representation. The compression rate is however low (1:1).

The learning stage of the MFCC in the different layers is as follows. While the PAC first receives at each iteration the MFCC vectors, the STG integrates with a temporal horizon of 20 iterations the different dynamics. Then, the third layer, the Striatum, categorizes the current state of the STG network in a higher dimension. In so far, the learning stage is feed-forward from PAC→STG→STR and the categorization is done in an unsupervised manner.

Once several passes are done over the complete audio sequence, the neurons stabilize to certain representations. It is possible then to perform an active exploration stage in the other direction – which means STR→GP→PAC→STG→STR for retrieving the corresponding audio entries in GP through reinforcement learning.

This stage corresponds to a motor babbling in which the audio inputs are generated in GP and evaluated after a delay in STR. The prediction error in STR is used to drive the dynamics in GP using free-energy and to control the PAC layer and STG dynamics via an iterative optimization process. Over time, each audio vector is reinforced for each GP-Striatal pair whenever the GP auditory pattern makes to fire its corresponding Striatal unit. The audio pattern converges to an optimal MFCC vector for which the Striatal unit was the most active. As proposed by several neuroscientists, the GP layer may control indirectly the Striatal layer through the cortical dynamics [4], [33], [23]. The prediction error may drive the amount of noise within the system and the ratio between exploration and exploitation. This scheme corresponds to a predictive coding mechanism, which can solve the temporal credit assignment problem between causes (in GP) and delayed effects (in STG).

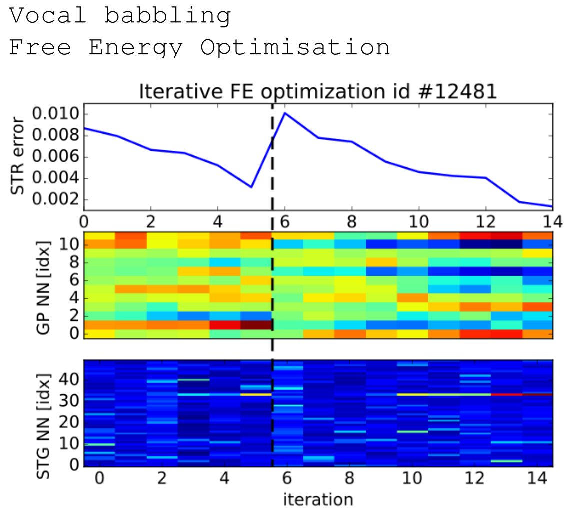

Fig. 3. Free-energy optimization. error minimization of one Striatal unit (top chart) using noise to retrieve GP vectors for which the Striatal units fire maximally (middle chart). The STG unit displays different spike trains for which a solution is found (bottom charts). The dashed lines correspond to a reset of the GP dynamics in order to show that the minimization process is always present and that different solutions can be retrieved dynamically.

We display in Fig. 3 one example of retrieved GP dynamics (middle chart) for which the prediction error in Striatum is diminished over time (top chart) with respect to the spatio-temporal patterns of the STG layer (bottom chart). The dashed line corresponds to a reset performed on the GP dynamics in order to observe dynamically the error minimization mechanism at work. The sample corresponds to the optimization process for one Striatal unit and for one GP vector. During the free-energy descent gradient, the GP vector converges to one audio pattern for which the STG activity is the most recognized by the corresponding Striatal unit. As showed in the graphs, the optimization process does not necessarily converge to the same minima after the reset done on the GP vector but can be stacked to another one. This means that different patterns of activity in the GP layer can influence in a similar way the activity in the STG layer. Therefore, the categorization done in STR is not perfectly orthogonal (sparse) and different solutions coexist to retrieve the STG spatio-temporal dynamics.

When reconstituting the .wav file in Fig. 4 from the retrieved MFCC vectors, we can observe a gradual refining of the audio waveform from the four periods with respect to the ground truth displayed at the bottom chart. The sequence is showed for 11 seconds although the global test was performed over two minutes length of the audio database.

After four exposures of the neural architecture to the audio

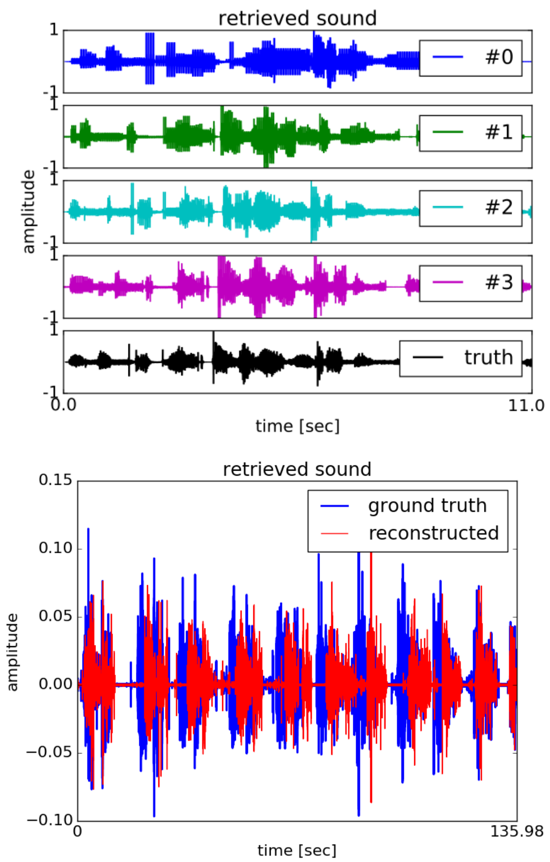

Fig. 4. Waveform reconstruction for the four learning periods.

sequence, the retrieved signals are gradually converging to the correct waveform. At period #0, the waveform is very discrete with square-like pattern and the amplitude and the wavelength are not respected. Gradually from period #1 to #3, we can observe a refinement of the waveform matching the ground truth curve[1].

B. Experiment 2 – structure learning and serial recall of audio sequences

This section explains the processing done in the network INFERNO Gate in order to learn the temporal structure in sequences and to retrieve back original sequences from incomplete information.

The iterative optimization process is done at the dlPFC level for retrieving memory sequences at the OFC level with the error rate computed at the ACC. The information flow corresponds to dlPFC→OFC→Broca Area (B45)→lPFC→dlPFC, see Fig. 2 b).

In order to understand better the global process, we display in Fig. 5, the final retrieved sequence in OFC with respect to the one represented by the dlPFC units is displayed in a), the raster plot of the iterative search of exact sequences in OFC and Broca Area (B45) is showed resp. b).

[1]We provide the link of the different .wav files at `https://promethe.u-cergy.fr/alexpitt/inferno`.

The desired sequences we want to reconstruct in OFC are presented at Fig 5 a) in the top chart. The raster plot of the reconstructed OFC/B45 dynamics are plotted in Fig 5 b).

The exploration search is performed after the learning stage. Over time, a sequence in OFC is explored iteratively using noise so that the dlPFC activity is maximal activity level and that ACC reaches a minimal value.

We display in Fig 5 a) in red, the retrieved OFC sequence in the top chart and the serial order for which the two different dlPFC neurons are the most sensitive to in the bottom chart.

In the top chart, we can observe that the reconstructed OFC sequences in red follow a similar pattern to the ones in blue although the identity of the neurons is not completely preserved. Nonetheless, we can see that the ordinal information in the bottom chart is matched, which means that the proposed sequence in the top chart follows the temporal pattern encoded in the lPFC and in the dlPFC layers.

Hence, despite the indices in the sequence have been lost in the encoding process, the system is capable to retrieve the memory sequences from incomplete information (due to compressive rank) with small error.

In order to analyze the accuracy of the Inferno Gate network, we plot in Fig 6 the euclidean error normalized between $[0, 1]$ made by the network during recall with respect to the number of items given as input vector, resp. b). The exploration stage was limited to 10.000 iterations for each experiment and we plot the retrieved sequences for 20 items given out of 50 items to retrieve, resp. in Fig. 6 b). The grey areas indicate the part of sequence given to the system to restitute the missing part.

In Fig. 6 a), the error rate computed from goal sequence encoded in lPFC/dlPFC and retrieved sequences in the OFC layer with respect to the amount of items given from 0 to 80% of the sequence given at the B45 level. The more the number of units to search are few, the more accurate is the recall. If we provide 40% of the items of the sequence we want to retrieve, the error on the neurons id is particularly small and almost error free if 80% of the neurons are given. In b), serial recall in OFC layer from incomplete information. Retrieved goal sequence when 40%, information are furnished to the system. In the top charts, the generated sequences in OFC layer with identity fo the STR neurons are displayed in red with the goal sequences to retrieve in blue. The more information is given to the system, the easier is the explorative search to retrieve the missing units identity. In the bottom chart, although the rank order in the temporal patterns of the units in lPFC is respected, this does not warranty that the units identity is retrieved correctly in the OFC sequences.

V. Discussion

We have presented the neural architecture INFERNO based on free-energy minimization using recurrent spiking neural networks for modeling the CX-BG loop [26]. This neural architecture is used for learning temporal sequences and for retrieving vocal 'motor' primitives by evaluating sensory feedback [28], although in our case, we did not

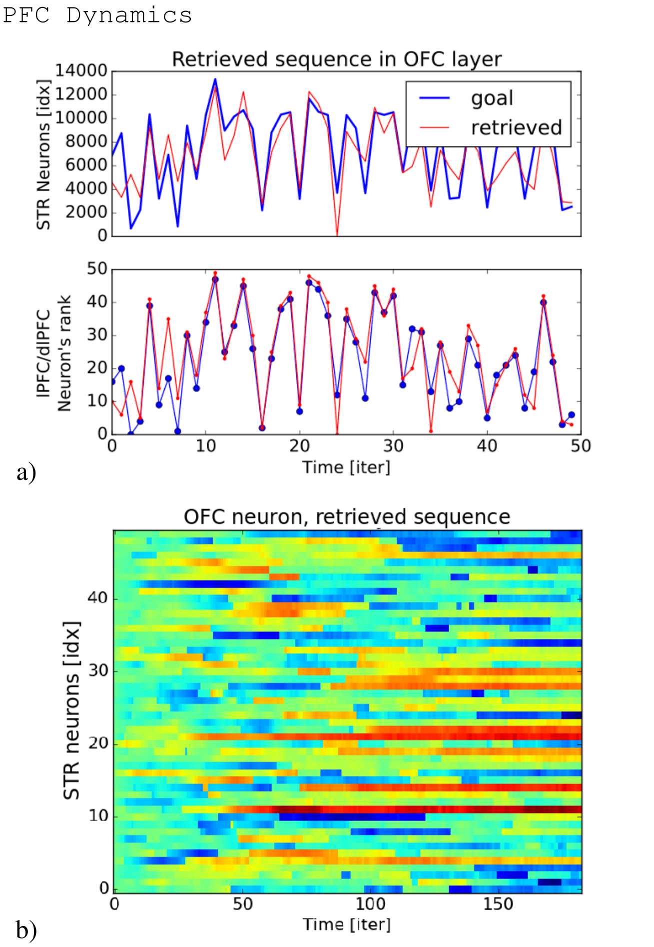

Fig. 5. Free-energy optimization for retrieving sequence in OFC layer. In a), the temporal pattern retrieved in lPFC/dlPFC layers and the temporal pattern of the goal sequence in blue are displayed in the bottom chart. In c), the OFC layer achieves to retrieve with some uncertainty a temporal sequence but the global structure of the sequence and the rank orders are mostly respected.

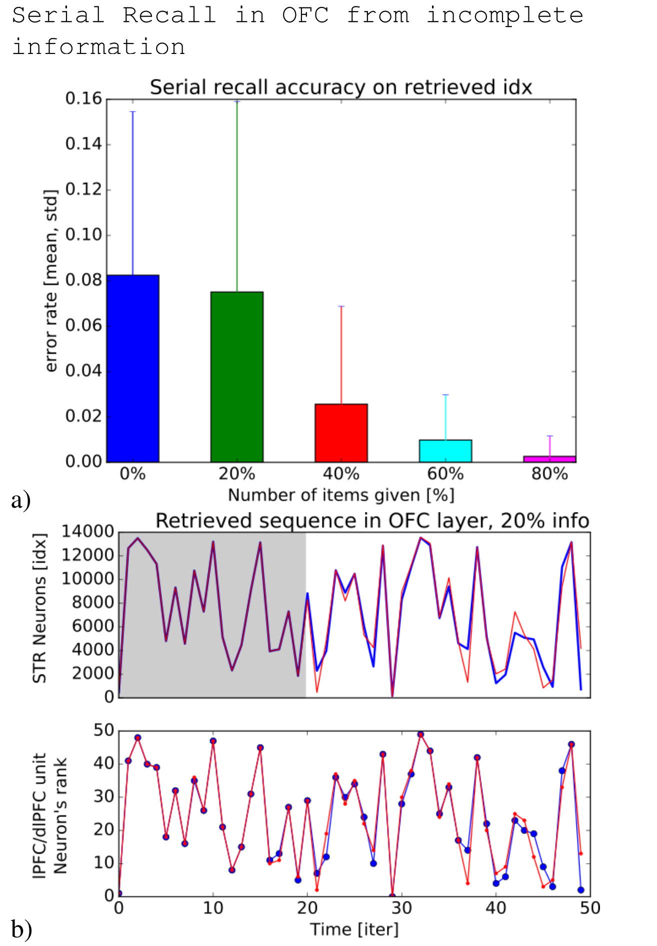

Fig. 6. Performance analysis of the Inferno gate architecture for retrieving sequences with respect to the amount of information given.

model the vocal articulatory system for sound generation but regenerated the sound primitives instead. We will model this part in further works.

The INFERNO network has two features, namely generalization and robustness to temporal delays. On the one hand, the number of units in the Striatum layer imposes a dimensionality reduction depending on the number of sound primitives to be learned (eg the number of MFCC vectors). On the other hand, the temporal chains in the CX layer permits to solve the temporal credit assignment problem and to link causes and effects thanks to STDP.

In the first experiments we have designed the network with the same number of STR units as MFCCs to retrieve (14.000 units) in order to have an orthogonal representation with few overlapping. The exploration of the audio primitives in a self-organized manner is similar to a vocal babbling, testing different sounds till convergence to the correct ones.

In second part, we have presented a novel neural architecture for modeling the fronto-striatal (PFC-BG) loop and learning temporal sequences. This network, named Inferno Gate [29], extends our original neural architecture Inferno presented earlier. Here, we have showed its effectiveness in the more challenging tasks of speech recognition and production based on structure learning. Although the two networks are similar in their functional organization, the encoding type is different. The first network uses the STDP mechanism for learning temporal correlations between spiking events whereas the second one uses a gating mechanism for binding the item's rank and their position within a sequence.

By discriminating content (which sound) and contextual information (when to play it in the sequence), we have showed that the two networks are capable to robustly learn the temporal structure within sequences and to retrieve the items identity in the correct order. In the first experiment, we did not compare the results of the network Inferno with standard RNN methods such as LSTM, as with 5% error reconstruction it performed advantageously. We will do the comparison in future research. In the second experiment, we compare the performances of the LSTM network with Inferno Gate in [29] and showed that our network outperforms the LSTM. The rationale is that by strictly separating the structure to learn and items to locate, the problem becomes easier even with a high number of items and with longer sequences.

In future researches, we envision to extend our framework to social and developmental robotics with speech generation and with visual and motor integration [34], [35], [36].

ACKNOWLEDGEMENTS

We would like to thank Mathieu Lagrange (LS2N, UMR 6004 CNS, Ecole Centrale de Nantes) for providing the audio database. This work was partially funded by EQUIPEX-ROBOTEX (CNRS), chaire dexcellence CNRS-UCP and project Labex MME-DII (ANR11-LBX-0023-01).

REFERENCES

[1] D. Wolpert, K. Doya, and M. Kawato, "A unifying computational framework for motor control and social interaction," *Philosophical Transactions of the Royal Society*, vol. 358, pp. 593–602, 2003.

[2] K. Doya, "Metalearning and neuromodulation," *Neural Networks*, vol. 15, p. 495506, 2002.

[3] T. Buschman and E. Miller, "Goal-direction and top-down control," *Phil. Trans. R. Soc. B*, vol. 369, p. 20130471, 2014.

[4] A. Graybiel, "The basal ganglia and chunking of action repertoires," *Neurobiol Learn Mem*, vol. 70, p. 119136, 1998.

[5] J. Tanji, K. Shima, and H. Mushiake, "Concept-based behavioral planning and the lateral prefrontal cortex," *Trends in Cognitive Sciences*, vol. 11, no. 12, pp. 528–534, 2007.

[6] A. Graybiel and S. Grafton, "The Striatum: Where Skills and Habits Meet," *Cold Spring Harb Perspect Biol*, vol. 7, p. a021691, 2015.

[7] A. Barto, "Adaptive critics and the basal ganglia," *In J. Houk, J. Davis, D. Beiser (Eds.), Models of information processing in the basal ganglia. Cambridge, MA MIT Press.*, pp. 215–232, 1995.

[8] W. Schultz, P. Dayan, and P. Montague, "A neural substrate of prediction and reward," *Annu. Rev. Neurosci.*, vol. 275, pp. 1593–1599, 1997.

[9] J. Tenenbaum, C. Kemp, T. Griffiths, and N. Goodman, "How to grow a mind statistics, structure, and abstraction," *Science*, vol. 331, no. 6022, pp. 1279–1285, 2011.

[10] K. Friston, J. Kilner, and L. Harrison, "A free energy principle for the brain," *Journal of Physiology-Paris*, vol. 100, no. 1-3, p. 7087, 2006.

[11] T. E and M. Jordan, "Optimal feedback control as a theory of motor coordination," *Nat Neurosci*, vol. 5, p. 12261235, 2002.

[12] H. Harlow, "The formation of learning sets," *Psychological Review*, vol. 56, no. 1, pp. 51–65, 1949.

[13] A. Diamond, "Development of the ability to use recall to guide action, as indicated by infants' performance on a-not-b," *Child Development*, vol. 74, pp. 24–40, 1985.

[14] Y. Munakata, "Computational cognitive neuroscience of early memory development," *Developmental Review*, vol. 24, pp. 133–153, 2004.

[15] J. Saffran, R. Aslin, and E. Newport, "Statistical learning by 8-month-old infants," *Science*, vol. 274, p. 19261928, 1996.

[16] T. Nazzi, J. Bertoncini, and J. Mehler, "Language discrimination by newborns: toward an understanding of the role of rhythm," *J. Exp. Psychol. Hum. Percept. Perform*, vol. 24, p. 756766, 1998.

[17] G. Marcus, S. Vijayan, S. Bandi Rao, and P. Vishton, "Rule learning by seven-month-old infants," *Science*, vol. 283, p. 7780, 1999.

[18] P. Barone and J. J.P., "Prefrontal cortex and spatial sequencing in macaque monkey," *Exp Brain Res*, vol. 78, p. 447 464, 2018.

[19] M. Botvinick, Y. Niv, and A. Barto, "Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective," *Cognition*, vol. 113, no. 3, p. 262280, 2009.

[20] L. Romanski, B. Averbeck, and D. M., "Neural representation of vocalizations in the primate ventrolateral prefrontal cortex," *J Neurophysiol*, vol. 93, p. 734747, 2005.

[21] C. Wacongne, E. Labyt, V. van Wassenhove, T. Bekinschtein, L. Naccache, , and S. Dehaene, "Evidence for a hierarchy of predictions and prediction errors in human cortex," *Proc. Natl. Acad. Sci. USA*, vol. 108, p. 2075420759, 2011.

[22] S. Dehaene, F. Meyniel, C. Wacongne, L. Wang, and C. Pallier, "The neural representation of sequences from transition probabilities to algebraic patterns and linguistic trees," *Neuron*, vol. 88, pp. 2–19, 2015.

[23] K. Friston and S. Kiebel, "Predictive coding under the free-energy principle," *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 364, pp. 1211–21, 2009.

[24] G. Bi and M. Poo, "Activity-induced synaptic modifications in hippocampal culture, dependence of spike timing, synaptic strength and cell type," *J. Neurscience*, vol. 18, pp. 10 464–10 472, 1998.

[25] A. Pitti and Y. Kuniyoshi, "Modeling the cholinergic innervation in the infant cortico-hippocampal system and its contribution to early memory development and attention," *Proc. of the International Joint Conference on Neural Networks (IJCNN11)*, pp. 1 – 8, 2011.

[26] A. Pitti, P. Gaussier, and M. Quoy, "Iterative free-energy optimization for recurrent neural networks (inferno)," *PLoS ONE*, vol. 12, no. 3, p. e0173684, 2017.

[27] A. Pitti, R. Braud, S. Mah, M. Quoy, and P. Gaussier, "Neural model for learning-to-learn of novel task sets in the motor domain," *Frontiers in Psychology*, vol. 4, no. 771, 2013.

[28] A. Pitti, , M. Quoy, C. Lavandier, and S. Boucenna, "Feature and structural learning of memory sequences with recurrent and gated spiking neural networks using free-energy: application to speech perception and production i," *unpublished*, 2019.

[29] A. Pitti, M. Quoy, C. Lavandier, and S. Boucenna, "Gated spiking neural network using iterative free-energy optimization and rank-order coding for structure learning in memory sequences (inferno gate)," *Neural Networks*, vol. 121, pp. 242–258, 2020.

[30] K. Friston, J. Daunizeau, and S. Kiebel, "Reinforcement learning or active inference?" *PLoS ONE*, vol. 4, no. 7, p. e6421, 2009.

[31] K. Friston, "The free-energy principle a rough guide to the brain?" *Trends in Cognitive Science*, vol. 4, no. 7, pp. 293–301, 2009.

[32] I. Tsuda, "Chaotic itinerancy and its roles in cognitive neurodynamics," *Current Opinion in Neurobiology*, vol. 31, p. 6771, 2015.

[33] K. Doya, "Metalearning, neuromodulation, and emotion," *G. Hatano, N. Okada, H. Tanabe (Eds.), Affective Minds*, pp. 101–104, 2000.

[34] A. Pitti, A. Blanchard, M. Cardinaux, and P. Gaussier, "Gain-field modulation mechanism in multimodal networks for spatial perception," *12th IEEE-RAS International Conference on Humanoid Robots Nov.29-Dec.1, 2012. Business Innovation Center Osaka, Japan*, pp. 297–302, 2012.

[35] O. Dermy, S. Boucenna, A. Pitti, and A. Blanchard, "Developmental learning of audio-visual integration from facial gestures of a social robot," *preprint*, 2016. [Online]. Available: https://hal.archives-ouvertes.fr/hal-02185423v1

[36] P. Valentin, S. Boucenna, P. Gaussier, and A. Pitti, "Robot recognizing vowels in a multimodal way," *IEEE ICDL-EPIROB, 2019*, 2019. [Online]. Available: https://hal.archives-ouvertes.fr/hal-02185423v1