Self-organization based on auditory feedback promotes acquisition of babbling

Tomohiro Takimoto, Yuji Kawai, Jihoon Park, and Minoru Asada
Graduate School of Engineering, Osaka University
2-1 Yamada-oka, Suita, Osaka, 565-0871 Japan
Email:{tomohiro.takimoto, kawai, jihoon.park, asada}@ams.eng.osaka-u.ac.jp

***Abstract*—Interaction between the body and the brain network is important for cognitive and behavioral development. Sensory feedback to the brain, originating from the body, accelerates self-organization of the brain network. This self-organization may lead to the acquisition of new behaviors. However, how self-organization promotes behavior acquisition and how this brain-behavior interaction develops remain unclear. We propose a recurrent spiking neural network (RSN) model of the acquisition of canonical babbling, and show that self-organization of the RSN based on auditory feedback can promote such acquisition. In this model, the output of the RSN is converted to vocalization, and its sound spectrum is fed back to the RSN. Synaptic weights in the RSN are updated via spike-timing-dependent plasticity (STDP). The output weights of the RSN are modulated by the dopamine STDP, i.e., reward learning to acquire the babbling. The study demonstrated that in the model incorporating STDP under auditory feedback, babbling was acquired faster than it was in the model without STDP. Our analysis indicated that self-organization enhanced the complexity of dynamics of the RSN, resulting in faster reward learning. We also found that there was an optimal balance between STDP and dopamine STDP, which implies that self-organization that is too fast or too slow may be disadvantageous with regard to behavior acquisition.**

I. Introduction

Developing brains are undergoing rapid organization through interaction between neural activity and bodily behaviors. The brain network is organized via neural plasticity based on sensory feedback derived from bodily behaviors. The self-organized brain produces more mature behavior, and is fed back the resulting sensory information. This brain-behavior interaction is responsible for brain development [1]. However, sensory disorders that impede sensory feedback may lead to atypical brain development. An understanding of the mechanisms involved in brain-behavior interaction in infancy is important with regard to the formulation of a method of rehabilitation for such atypically developing children.

In the case of vocal development, canonical babbling evidently facilitates the self-organization of infants'brains, especially with regard to associations between their articulatory control input and the resulting sounds [2]. Canonical babbling contains canonical syllables that can be represented as a rhythmic series of consonants and vowels, for example "CVCVCV." It develops from single syllables, e.g., "da" or "pa" to reduplicated syllables, e.g., "dadada" over several months. Infants with severe or profound hearing impairment reportedly exhibit delayed development of canonical babbling [3], [4], and produce fewer consonants per their utterance

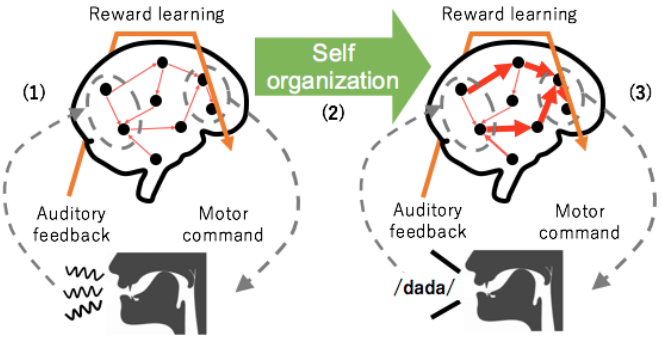


Fig. 1. Self-organization of the brain network promotes the acquisition of canonical babbling. (1) The immature brain network produces an awkward vocalization. (2) The network is organized through the audiomotor loop. At the same time, the agent learns more sophisticated babbling based on reward. (3) The self-organization of the network and the reward learning are mutually enhancing, resulting in the rapid acquisition of babbling.

[5]. These observations suggest the possibility that canonical babbling is acquired through a combination of vocalization and auditory feedback.

A study investigating adult brains reported that several neural activities in the sensorimotor cortex dynamically represent movements of speech-articulators (e.g., tongue, jaw, lips, and larynx) and vocalization of phonemes (consonants and vowels) [6]. That study suggested that the dynamic organization of the speech sensorimotor cortex during the generation of multi-articulator movements is one of the foundations of speaking ability. Such body representation could be reorganized through motor experiences, which are sustained by brain plasticity (e.g., see [7]). In line with these facts, infant brains may be organized through canonical babbling, which represents the speech-articulator relationship. This self-organization may contribute to their speech development. However, the mechanism of this mutual development of the brain and behavior remains unclear.

Computational models of development through interaction between the body and the brain network have been described [8], [9]. Recently, Yamada et al. [8] constructed a very realistic fetus simulation, and investigated the self-organization of a recurrent spiking neural network (RSN) as a result of intrauterine brain-body-environment coupling. That model incorporated the RSN as the somatosensory cortex, which received tactile

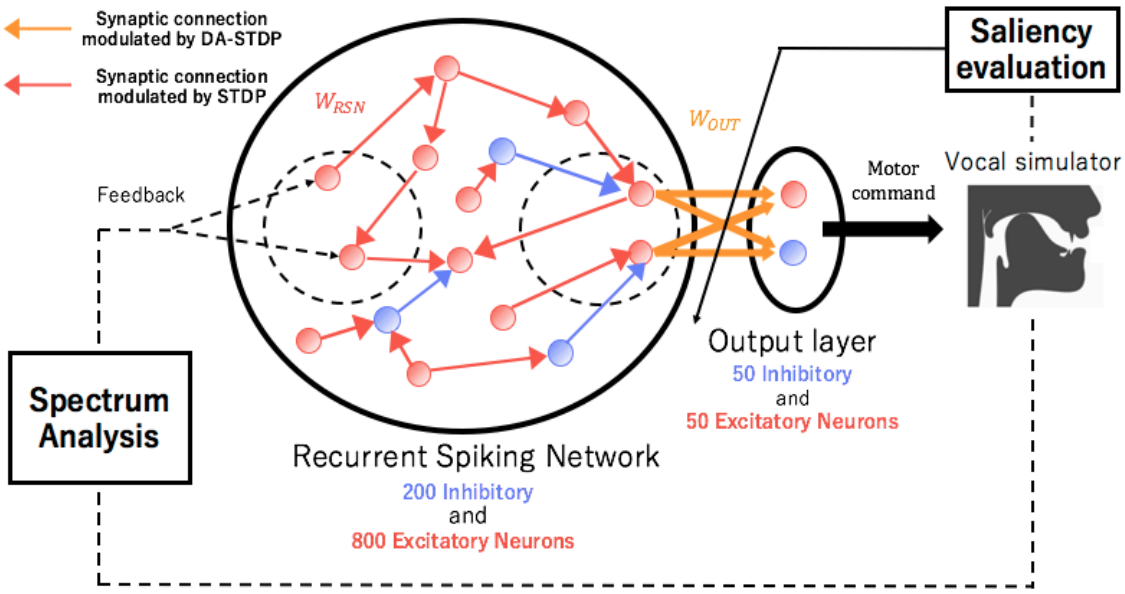


Fig. 2. The model of babbling acquisition through self-organization of the recurrent spiking neural network (RSN). The red and blue circles indicate excitatory and inhibitory spiking neurons, respectively. The arrows represent connections between neurons. Each neuron is randomly connected to one hundred other neurons in the RSN. One hundred excitatory neurons project their spikes to the output layer, represented by the red circles enclosed within the right broken circle in the RSN. One hundred excitatory neurons receive the feedback, represented by the red circles enclosed within the left broken circle in the RSN.

information from the body. The RSN was organized via spike-timing-dependent plasticity (STDP) while it received sensory feedback. They found that self-organized RSN represented body-parts of the fetus. However, the brain-body interaction was one way with regard to simulation, that is, from body to brain (sensory feedback), while bodily behavior is generated via constant input. Therefore, they did not address the issue of how the self-organized brain affects behavior.

Warlaumont et al. [10] proposed a model of canonical babbling learning using an RSN and dopamine-modulated STDP (DA-STDP) [11]. The RSN was only activated by random external input, and its synaptic weights remained constant. Some neurons in the RSN sent their currents to the output neurons through synaptic weights that were modulated by the DA-STDP. Dopamine-levels were based on the saliency of the vocalized sounds, which were defined as an averaged temporal variation of the sound spectrum. They demonstrated that this form of reward learning was key to the emergence of canonical babbling. Notably however, this model does not consider the interaction between vocalized sound and the RSN. Auditory information pertaining to the vocalized sound is not returned to the RSN, and the RSN is not self-organized. A number of studies have reported that STDP can enhance the computational capacity of the RSN [12] [13]. Norton et al. [12] showed that STDP improved the dynamic properties of the RSN, although they did not aim to model the cognitive developmental process.

In the current study, we examined how self-organization of the RSN via STDP based on sensory feedback affects babbling acquisition. We formulated an RSN model of canonical babbling learning, wherein the RSN receives auditory feedback and is organized via STDP (Fig. 1). We hypothesized that the model reflected the efficient learning of babbling via the following mechanisms, where each item number corresponds to the same number in Fig. 1:

(1). The RSN activates a vocal simulator, and the vocalized sounds are fed back to the RSN.
(2). Self-organization based on the auditory feedback refines and optimizes the dynamics of the RSN.
(3). Using the refined dynamics, reward learning enables the model to vocalize the babbling more clearly. Return to (1).

II. A RSN MODEL TO ACQUIRE BABBLING

A. Overview of the model

Fig. 2 shows an overview of the proposed model. The model consisted of the RSN, an output layer, and a vocal simulator. One hundred excitatory neurons in the RSN activate the output layer, which creates motor commands for the vocal simulator. Synaptic weights from the RSN to the output layer are modulated by DA-STDP [11], where dopamine-level is based on sound saliency. This reward-based plasticity facilitates more salient vocalization. The above-mentioned vocalization model was based on the model proposed by Warlaumont et al. [10]. Additionally, we considered self-organization of the RSN under auditory feedback in order to incorporate brain-body interaction. In the model, one hundred excitatory neurons in the RSN receive input according to the frequency spectrum of

the vocalized sound. At the same time, synaptic weights of all excitatory neurons in the RSN are modulated via STDP.

B. RSN and vocalization

We employed Izhikevich spiking neurons [14], which are based on the biophysically accurate Hodgkin-Huxley neuronal model but entail relatively low computational cost. The set parameters of neurons were almost identical to those utilized in the network described by [11]. Neurons in the RSN receive input from other connected neurons, and input that is randomly generated in the range of -6.5 to $+6.5$ mV. Excitatory and inhibitory neurons exist at a ratio of $4:1$, and each neuron is connected to another randomly selected one hundred neurons. In the current study, the delay of all synapses was uniformly set to 1 ms.

In the model, each neuron in the output layer receives input from one hundred neurons that are randomly selected in the RSN, and a random input that is identical to that of the RSN. The ratio between the excitatory and inhibitory neurons is $1:1$ in the output layer. The firing of these excitatory and inhibitory activities ($S_{\rm e}(t)$ and $S_{\rm i}(t)$, respectively) generates a motor command $m(t)$ at time t ms:

$$m\left(t\right) = g\sum_{i=0}^{100}\left(S_{\rm e}(t-i) - S_{\rm i}(t-i)\right)\left(1 - e^{\left(\frac{100-i}{\tau}\right)}\right), \quad (1)$$

where g indicates the constant motor gain and τ indicates the decay parameter. We set g to 0.05 and τ to 20 ms. The motor command $m(t)$ is generated from the firing history of the output layer for the most recent 100 ms, and is created every 1 ms after the first 100 ms of each simulation. Based on the decay, the influence of the past firing history from the recent 100 ms decreases exponentially. The vocal simulator receiving $m(t)$ produces a sound every 1 ms. We employed the Praat as the vocal simulator [15].

C. Auditory feedback

In the model, the vocalized sound is converted to a frequency spectrum, and its discretized frequency powers are fed back to the RSN (Fig. 3 (a)). The spectrum is uniformly divided into 100 bands (1–20 Hz, 21–40 Hz, etc., up to 1981–2000 Hz). The ith input to the RSN I_i is calculated based on the power of the ith frequency band E_i:

$$\begin{aligned} E_{\max} &= \max\left\{E_i : i = 1, 2, \ldots, 100\right\}, & (2)\\ I_i &= b\cdot\frac{E_i}{E_{\max}}, & (3)\end{aligned}$$

where b denotes the constant input gain and was set to 13. $E_{\max}$ is the maximum value among the frequency bands. I_i is fed back to a neuron that is randomly selected from the RSN. Therefore, one hundred neurons in the RSN receive the auditory feedback.

D. Plasticity rules

The RSN is organized using STDP, where the synaptic weights $W_{\rm RSN}$ are modulated based on correlations between presynaptic and postsynaptic neurons [16]. If neuron A fires just after the firing of neuron B, the synaptic weight from neuron A to neuron B is strengthened as A_+e^{-t/τ_+}, which is designated as the long-term potentiation. Conversely, if neuron B fires just after neuron A fires, the synaptic weight from neuron A to neuron B is weakened as A_-e^{-t/τ_-}, and this is designated long-term depression. In our settings, $\tau_+ = \tau_- = 20$ ms, $A_+ = 0.1$, and $A_- = 0.12$. Under this plasticity, the synapses involved in spike firings of neurons are strengthened, while the synapses not involved in it are weakened. A synaptic weight $w_{\rm rsn}$ is modulated as represented by the following:

$$\begin{aligned} \dot{c} &= -c/\tau_{\rm c} + {\rm STDP}(\tau_{\rm stdp})\delta(t - t_{\rm pre/post}),\\ \dot{w}_{\rm rsn} &= P_{\rm stdp}\dot{c}, & (4)\\ {\rm STDP}(\tau_{\rm stdp}) &= \begin{cases} A_+e^{-t/\tau_+} & (\text{if } \tau_{\rm stdp} > 0)\\ A_-e^{-t/\tau_-} & (\text{if } \tau_{\rm stdp} < 0),\end{cases}\end{aligned}$$

where $\delta(t)$ is the Dirac delta function that step-increases the variable c. Firings of presynaptic and postsynaptic neurons, occurring at times $t_{\rm pre}/t_{\rm post}$ change c depending on ${\rm STDP}(\tau_{\rm stdp} = t_{\rm post} - t_{\rm pre})$. This variable exponentially decays to zero where the time constant $\tau_c = 1$ sec. $P_{\rm stdp}$ is the learning rate of the STDP that controls the degree of the effect of the STDP. Too large (or too small) values of $P_{\rm stdp}$ provide rapidly-changeable (or hard-to-change) networks.

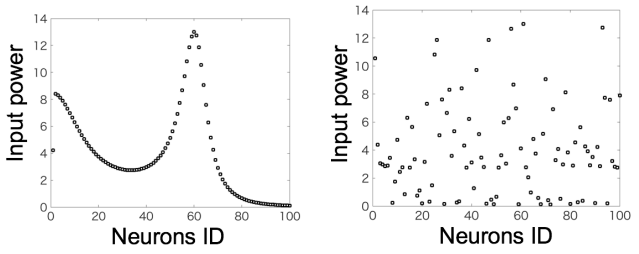
Synaptic weights $w_{\rm out}$ from the RSN to the output layer are modulated by DA-STDP. Dopamine is an organic chemical that is related to learning in the brain, and it has been reported that dopamine-modulated organization in the motor cortex promotes behavior acquisition [17]. We employed Izhikevich's DA-STDP model [11] with the modification that only long-term potentiation was considered. This model modifies $w_{\rm out}$ based on the extracellular concentration of dopamine d:

$$\begin{aligned} \dot{c}_{\rm da} &= -c_{\rm da}/\tau_{\rm c} + {\rm STDP}(\tau_{\rm stdp})\delta(t - t_{\rm pre/post}),\\ \dot{d} &= -d/\tau_d + DA(t), & (5)\\ \dot{w}_{\rm out} &= dP_{\rm da}\dot{c}_{\rm da},\end{aligned}$$

where ${\rm DA}(t)$ denotes the reward signal that is updated as ${\rm DA}(t) = 1$ when a reward is given at time t, and $P_{\rm da}$ is the learning rate of the DA-STDP. This plasticity rule is the STDP that works only when a reward is given. Under this learning rule, the output layer is organized to produce the rewarded firing patterns.

In our model, the reward is given when the saliency of the vocalized sounds exceeds a threshold, as in the previously described model [10]. Saliency is defined as the temporal variation of the sound spectrum, such that a reward is given for syllabic sounds consisting of vowels and consonants. We implemented it using a program developed by Coath et al. [18].

The reward is evaluated for the vocalized 900-ms sound every 1000 ms. The sound for the first 150 ms is discarded due to the preparation of lung pressure, and therefore, the saliency of the vocalized sound for 750 ms is evaluated in practice. The STDP is conducted at all times, i.e., every 1 ms. The threshold value is set to 4.5, and increases by 0.1 if saliency exceeds the



(a) Normal feedback (b) Scrambled feedback

Fig. 3. Coding examples of a sound for the auditory feedback. The model converts the vocalized sound into its frequency spectrum. In the case of normal auditory feedback (a), a power in each frequency band is fed to a neuron. This correspondence between the band and the neuron is unchangeable. In contrast, the neurons that receive the input are randomly selected every time in the scrambled feedback (b).

threshold in three of the past ten instances, and conversely, it decreases by 0.1 if none of the past ten instances exceed the threshold.

III. Experiments

A. Experimental setting

The vocal simulator incorporated only two muscles, the masseter and the orbicularis oris lip muscle, which are related to jaw closure and lip closure, respectively. These muscles are synchronously actuated by a scalar motor command $m(t)$. Oscillation in the one degree of freedom produces a series incorporating the vowel and the consonant, e.g., “ ababa. ” Parameters of the simulator used to synthesize vocalization were set according to the Warlaumont et al. [10] model. The vocal simulator produced a sound for 900 ms, then its saliency was evaluated as a reward. We ran this process 2000 times in all our experiments. We evaluated the model via the saliency curve 2000 times.

We conducted two experiments to investigate how the model organizes the RSN and acquires babbling. In the first experiment, we examined the effects of STDP and auditory feedback. We prepared scrambled feedback (Fig. 3 (b)) to compare with normal auditory feedback (Fig. 3(a)). The input neurons in the RSN were randomly selected every time in the scrambled feedback, while they were fixed in the normal feedback. This comparison enabled us to investigate the advantages of consistent auditory feedback. We also examined the model without STDP. The model was tested in four conditions, i.e., self-organization with or without STDP $\times$ normal feedback or scrambled feedback. We conducted 10 simulations in each condition. We set P_{stdp} to 0.05 in Equation (4), and P_{da} to 1.00 in Equation (5) in all simulations.

In the second experiment, we examined the effect of P_{stdp}, which determines the amount of weight update that is due to STDP. We set P_{stdp} to 0.100, 0.075, 0.050, 0.025, 0.010, or 0, while P_{da} was fixed to 1.000, and ran the simulation 10 times under each condition. Normal auditory feedback was employed in this experiment.

We analyzed the global dynamics of the RSN in terms of dimensional complexity. After simulation for 2000 sec, we activated the RSN without auditory feedback for 2000 ms. In this phase, the RSN did not organize using STDP or DA-STDP. Its neuronal firing history from 1000 ms to 2000 ms was analyzed via principle component analysis. We counted the number of principal components whose accumulated contribution rate reached 0.8. This indicated the number of linear spaces to explain the neural dynamics, i.e., the degree of its complexity.

B. Result 1: Effect of auditory feedback and STDP

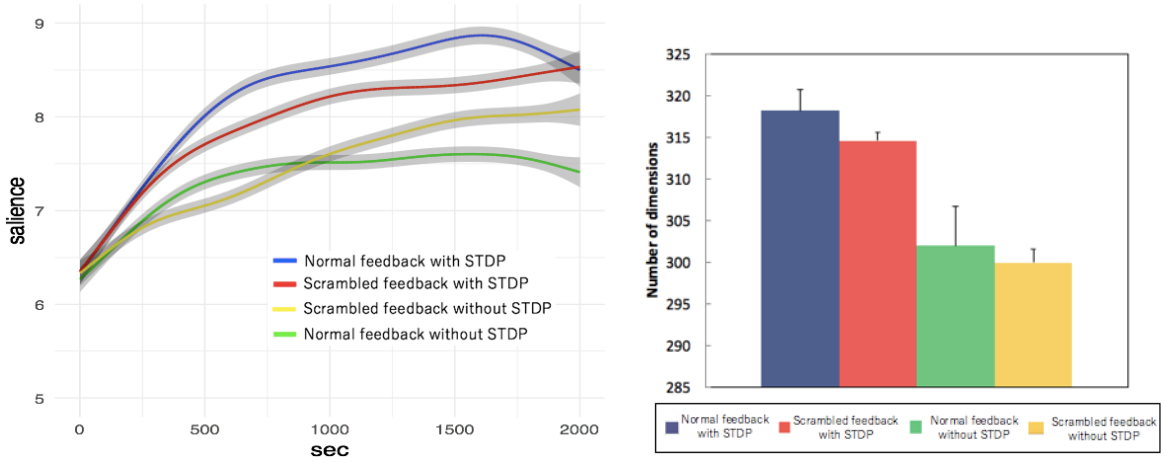
Fig. 4 (a) shows the results of the first experiment. The blue and red curves are results of the model with STDP, and the green and yellow curves are results of the model without STDP. The blue and green curves are the results of the models with normal auditory feedback (Fig. 3 (a)), and the red and yellow curves are results of the model with scrambled feedback (Fig. 3 (b)). As shown in Figure 4, the model with STDP based on auditory feedback (blue curve) exhibited the most rapid learning of babbling. The model with STDP based on scrambled feedback (red curve) also showed better learning performance than the models without STDP. These results suggest that STDP promoted the acquisition of babbling. Furthermore, adequate auditory feedback enhanced the effect of STDP. Fig. 4 (b) shows the dimensionality of the network activities in this experiment. The colors of the bars correspond to those of the curves in Fig. 4 (a). As shown in this figure, STDP increased the dimensional complexity of the dynamics of the RSNs (blue and red bars). In addition, an increase in the dimensionality of the model with STDP compared with the cases without STDP was apparent.

C. Result 2: Effect of learning rate

Fig. 5 (a) shows the results of the model with different P_{stdp} values, where the black, red, blue, purple, yellow, and green curves indicate the results where $P_{\mathrm{stdp}} = 0.100$, 0.075, 0.050, 0.025, 0.010, and 0, respectively. As shown in Figure 5, the model in which $P_{\mathrm{stdp}} = 0.050$ (blue curve) clearly exhibited the best learning performance. STDP with the high learning rate changed the dynamics of the RSN too fast, eliminating reward learning via DA-STDP. In contrast, the dynamics of the RSN were not modified when the P_{stdp} value was too small. In this setting, the model did not receive the benefit of STDP. When $P_{\mathrm{stdp}} = 0.05$, reward learning and self-organization of the RSN were well balanced, resulting in the fastest babbling acquisition. Fig. 5 (b) shows the dimensionality of the network activities in the second experiment. The colors of the bars correspond to those of the curves in Fig 5 (a). As shown in this figure, cases where P_{stdp} was > 0.05 exhibited more dimensions than cases where $P_{\mathrm{stdp}} = 0.025$, 0.010, or 0.

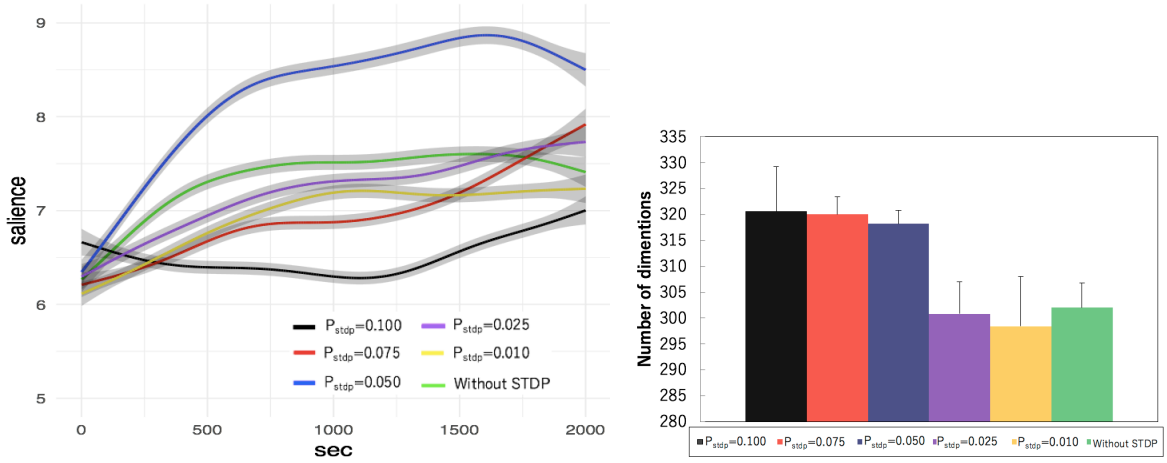
IV. Discussion

We proposed the embodied spiking neural network model, where the RSN produces babbling through reward learning, and the resulting sensory information is fed back into the RSN, which is organized via STDP. Our experimental results showed that this brain-behavior coupling, i.e. the self-organization



(a) Saliency (b) The number of dimensions of the neural dynamics

Fig. 4. Results of the first experiment. (a) Learning curves for the saliency of the vocalized sounds over time. The curves indicate generalized additive model fits, and the dark gray shading indicates 95% confidence intervals around those fits. (b) Averaged number of dimensions in each network activity. Error bars indicate standard deviation.



(a) Saliency (b) The number of dimensions of the neural dynamics

Fig. 5. Results of the second experiment. (a) Learning curves for the saliency of the vocalized sounds over time. The curves indicate generalized additive model fits, and dark gray shading indicates 95% confidence intervals around those fits. (b) Averaged number of dimensions in each network activity. Error bars indicate standard deviation.

of the RSN under sensory feedback, could promote reward learning, as shown in Fig. 4 (a) which indicates that normal feedback contributed to the promotion of the learning. Our analysis of RSN dynamics suggested that these results could be explained by the complexity of the activities of the RSNs. Self-organization under auditory feedback provided the RSN with richer dynamics, resulting in the promotion of reward learning.

This result is concordant with a report by Norton et al. [12] indicating that STDP of the RSN improves the performance of supervised learning, where activation of the RSN is used as input to train a linear regression. Several computational neuroscientists have studied the relationships between the sparseness of a neural network and the complexity of its activities [19] [20]. They have demonstrated that a sparser network, i.e., neurons with less connections, produces more complex neural activity. In the current study, STDP trimmed connections and made the RSN sparser. The RSN therefore produced more complex or richer neural dynamics, resulting in enhanced reward learning. In real infant brains, the complexity of their activities increases as they develop [21]. The results of future modeling studies investigating the relationships between sparseness of brain networks, complexity of neural activities, and performance of bodily behavior from a perspective of human development will be of interest. However, Laser et al. [22] demonstrated that STDP lowers the complexity of neural activities, resulting in improvement in the performance of supervised learning. This is contrary to the results of the current study. One reason for the discrepancy may be differences in the neuron models. We used a spiking neuron model, while Laser et al. [22] used a binary neuron model.

Notably, the above-described mechanism cannot explain

the contribution of auditory feedback. This might be because STDP under auditory feedback could provide the RSN with sensorimotor representation. Representation of the neural dynamics in the RSN should be analyzed to verify this hypothesis. We suggest that the RSN represented the articulator (e.g., mouth opening) and vocalization (e.g., vowels and consonants), as in the adult sensorimotor cortex [6]. The model without auditory feedback may correspond to infants with hearing loss. The delayed learning in this model compared to the model with auditory feedback is consistent with the delay in babbling development observed in such infants [3] [4]. The model suggests that one of the origins of the delay is an unorganized sensorimotor cortex.

We determined the critical value of the learning rate of STDP (Fig. 5 (a)). Although the complexity of the RSNs where $P_{\mathrm{stdp}} = 0.100$ and 0.075 was large, the corresponding learning performances were not very high. STDP and DA-STDP are assumed to be models of self-organization in the sensorimotor cortex and reward learning by the striatum, respectively [23]. This result suggests that too fast or too slow organization of the sensorimotor cortex, i.e., change in a state space of neural dynamics, might impede behavior acquisition. This problem arose from simultaneous organization via STDP and DA-STDP. A similar problem lies in the recently investigated deep reinforcement learning [24], where deep learning represents input while reinforcement learning is conducted using acquired representation. Changes in the representation may compromise the reinforcement learning that has occurred so far. A major solution is experience replay, whereby the representation is consolidated using memorized pairs of input and output while the reinforcement learning is stopped. This technique can separate the two plasticity phases. We can introduce the idea to our model, and this may improve babbling acquisition.

V. Conclusion

We investigated the role of the self-organization of brains in the context of the acquisition of babbling. Our study demonstrated that self-organization of the RSN under auditory feedback could promote the acquisition of babbling. The RSNs that showed high learning performance had relatively complex neural activity. These results imply that STDP increased the complexity of the activity of the RSN, resulting in enhanced babbling acquisition. Furthermore, we determined the critical value of the learning rate of STDP. Successful learning required sufficiently rapid self-organization. This elaborate mechanism of the self-organization of the sensorimotor cortex may support the behavioral development of infants.

Acknowledgment

This research and development work was supported by the MIC

References

[1] L. Byrge, O. Sporns, and L. B. Smith, "Developmental process emerges from extended brain–body–behavior networks," *Trends in Cognitive Sciences*, vol. 18, no. 8, pp. 395–403, 2014.

[2] G. Westermann and E. R. Miranda, "A new model of sensorimotor coupling in the development of speech," *Brain and Language*, vol. 89, no. 2, pp. 393–400, 2004.

[3] S. N. Iyer and D. K. Oller, "Prelinguistic vocal development in infants with typical hearing and infants with severe-to-profound hearing loss," *Volta Review*, vol. 108, no. 2, pp. 115–138, 2008.

[4] D. K. Oller and R. E. Eilers, "The role of audition in infant babbling," *Child Development*, vol. 59, no. 2, pp. 441–449, 1988.

[5] C. Stoel-Gammon and K. Otomo, "Babbling development of hearing-impaired and normally hearing subjects," *Journal of Speech and Hearing Disorders*, vol. 51, no. 1, pp. 33–41, 1986.

[6] K. E. Bouchard, N. Mesgarani, K. Johnson, and E. F. Chang, "Functional organization of human sensorimotor cortex for speech articulation," *Nature*, vol. 495, no. 7441, pp. 327–332, 2013.

[7] M. Hallett, "Plasticity of the human motor cortex and recovery from stroke," *Brain Research Reviews*, vol. 36, no. 2, pp. 169–174, 2001.

[8] Y. Yamada, H. Kanazawa, S. Iwasaki, Y. Tsukahara, O. Iwata, S. Yamada, and Y. Kuniyoshi, "An embodied brain model of the human foetus," *Scientific Reports*, vol. 6, 2016.

[9] Y. Kuniyoshi and S. Sangawa, "Early motor development from partially ordered neural-body dynamics: experiments with a cortico-spinal-musculo-skeletal model," *Biological Cybernetics*, vol. 95, no. 6, pp. 589–605, 2006.

[10] A. S. Warlaumont and M. K. Finnegan, "Learning to produce syllabic speech sounds via reward-modulated neural plasticity," *PloS One*, vol. 11, no. 1, p. e0145096, 2016.

[11] E. M. Izhikevich, "Solving the distal reward problem through linkage of STDP and dopamine signaling," *Cerebral Cortex*, vol. 17, no. 10, pp. 2443–2452, 2007.

[12] D. Norton and D. Ventura, "Preparing more effective liquid state machines using hebbian learning," in *in Proceedings of the IEEE International Joint Conference on Neural Networks*, 2006, pp. 4243–4248.

[13] F. Xue, Z. Hou, and X. Li, "Computational capability of liquid state machines with spike-timing-dependent plasticity," *Neurocomputing*, vol. 122, pp. 324–329, 2013.

[14] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Transactions on Neural Networks*, vol. 14, no. 6, pp. 1569–1572, 2003.

[15] P. Boersma and D. Weenink, "Praat: doing phonetics by computer[computer program]. retrieved may 1, 2009," 2015.

[16] E. M. Izhikevich, "Polychronization: computation with spikes," *Neural Computation*, vol. 18, no. 2, pp. 245–282, 2006.

[17] K. Molina-Luna, A. Pekanovic, S. Röhrich, B. Hertler, M. Schubring-Giese, M.-S. Rioult-Pedotti, and A. R. Luft, "Dopamine in motor cortex is necessary for skill learning and synaptic plasticity," *PloS One*, vol. 4, no. 9, p. e7082, 2009.

[18] M. Coath, S. L. Denham, L. M. Smith, H. Honing, A. Hazan, P. Holonowicz, and H. Purwins, "Model cortical responses for the detection of perceptual onsets and beat tracking in singing," *Connection Science*, vol. 21, no. 2-3, pp. 193–205, 2009.

[19] K. J. Friston, "Theoretical neurobiology and schizophrenia," *British Medical Bulletin*, vol. 52, no. 3, pp. 644–655, 1996.

[20] G. Tononi, G. M. Edelman, and O. Sporns, "Complexity and coherency: integrating information in the brain," *Trends in Cognitive Sciences*, vol. 2, no. 12, pp. 474–484, 1998.

[21] S. Lippe, N. Kovacevic, and R. McIntosh, "Differential maturation of brain signal complexity in the human auditory and visual system," *Frontiers in Human Neuroscience*, vol. 3, p. 48, 2009.

[22] A. Lazar, G. Pipa, and J. Triesch, "SORN: a self-organizing recurrent neural network," *Frontiers in Computational neuroscience*, vol. 3, no. 23, 2009.

[23] P. F. Dominey, "Complex sensory-motor sequence learning based on recurrent state representation and reinforcement learning," *Biological Cybernetics*, vol. 73, no. 3, pp. 265–274, 1995.

[24] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Patersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.