International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:3, No:1, 2009

A Computer Model of Language Acquisition – Syllable Learning – Based on Hebbian Cell Assemblies and Reinforcement Learning

Sepideh Fazeli, and Fariba Bahrami

***Abstract*—**Investigating language acquisition is one of the most challenging problems in the area of studying language. Syllable learning as a level of language acquisition has a considerable significance since it plays an important role in language acquisition. Because of impossibility of studying language acquisition directly with children, especially in its developmental phases, computer models will be useful in examining language acquisition. In this paper a computer model of early language learning for syllable learning is proposed. It is guided by a conceptual model of syllable learning which is named *Directions Into Velocities of Articulators* model (DIVA). The computer model uses simple associational and reinforcement learning rules within neural network architecture which are inspired by neuroscience. Our simulation results verify the ability of the proposed computer model in producing phonemes during babbling and early speech. Also, it provides a framework for examining the neural basis of language learning and communication disorders.

***Keywords*—**Brain modeling, computer models, language acquisition, reinforcement learning.

I. Introduction

Explanation of language is one of the most important subjects when brain functions are to be investigated [1]. In contrast to great advances in cognitive science, our knowledge of human brain interactions during processing, learning and acquiring language has little improvement over the last 50 years [2]. This happened due to unavailability of animal models which allow more detailed studies of language processing using invasive but informative techniques [3]. At this juncture, computer models and simulations provide more adequate understanding of both structures and causes-and-effects which are involved in a specific task. Computer models also yields to further advances in computer science specifically in the connectionist branch of Artificial Intelligence [4].

In the field of language acquisition, computational approaches and computer simulations can be advantageous by providing a framework in order to achieve more adequate investigation into validity of proposed theorems. Moreover, using computer models help researchers to study different phases of learning, especially initial and developmental states, which are very difficult to be carried out with a child [2].

In this paper, a computer model of language acquisition using only simple associational and reinforcement learning rules is outlined. Associative neural networks are regarded as standard models for Hebbian Cell Assemblies which have been argued to support a variety of different cognitive tasks. In the field of language acquisition, cell assembly concept demonstrated its ability in sharpening our knowledge about the cortical interactions underlying language acquisition [5].

Language acquisition occurs in two stages. First, in sensory phase, a sensory template is formed by listening to and memorizing tutor's speech. Then, in sensorymotor phase, the infant tries to produce the same syllable as his tutor by reinforcing utterances which are sufficiently similar to memorized templates. The child uses his own auditory and somatosensory feedbacks to investigate the similarity. Repeated production of the sounds results in tuning feedforward connections that ultimately diminish the feedback-based reinforcement signals. This assumption is motivated by Directions Into Velocities of Articulators model (DIVA) which is a neural network model of speech production [6]. DIVA as a conceptual model provides a framework for interactions between feedforward control system and auditory and somatosensory control systems. Feedforward control system includes premotor and primary motor cortex along with cerebellum. Auditory and somatosensory control systems include both sensory and motor cortical areas [6].

To implement the proposed computer model we assume that the sensory learning is completed and concentrate only on the sensorymotor phase of learning. The results obtained from preliminary experiments show that the computer model is reasonably valid. It starts with random activation in premotor cortex and ends up producing exact syllables. An important feature of the proposed computer model which distinguishes it from other computer models is the use of Hebbian cell assembly concept and reinforcement learning within neural network architecture which are inspired by neuroscience as well as its congruity with DIVA model. Moreover, other

S. Fazeli is with Control and Intelligent Processing Center of Excellence, School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran (e-mail: s.fazeli@ece.ut.ac.ir).

F. Bahrami is with Control and Intelligent Processing Center of Excellence, School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran (e-mail: fbahrami@ut.ac.ir).

implementations of DIVA do not focus on cortical interactions during syllable learning. Furthermore, our proposed computer model has the capability to achieve more precise insight into language acquisition process in both healthy and brain damaged subjects.

This paper is organized as follows. DIVA model and conceptual approaches are introduced in Section II. The proposed method to simulate the computer model is explained in Section III. Results of proposed computer method are presented in Section IV. Final conclusions are then expressed in Section V.

II. DIVA Model and Approach

The DIVA model has been developed based on various neuroanatomical and neurophysiological studies [6]. This model is schematized in Fig. 1. Each block represents a group of neurons in human brain. In this model, projections from premotor cortex to primary cortex correspond to feedforward control of the speech articulators. Efference copy projections are from premotor cortex to auditory cortical area in the superior temporal gyrus as well as orosensory area in the supramarginal gyrus. These efference copy projections generate internal prediction of auditory and somatosensory feedbacks corresponding to each syllable. In this model, comparison between efference copy projections to the auditory cortical area and supramarginal gyrus and the auditory and somatosensoy feedbacks, results in error signals which are mapped onto the cerebellum. Eventually, based on these error signals, a reinforcement signal is transmitted by the cerebellum to modulate intrinsic plastic connections within primary cortex, as well as the projection from premotor cortex. These projections through the cerebellum to motor cortex form components of the DIVA mapping [6], [7].

Besides the assumptions used in the DIVA model, some additional functional ones are also applied in our model. To start with, in order to decrease the interfering effects of delayed auditory and somatosensory feedbacks on syllable learning, two strategies are proposed. First, the auditory and somatosensory feedbacks are set significantly weaker than efference copy signals. The second strategy is based on adaptation mechanism which produces delayed, negative images of auditory and orosensory activities in the superior temporal gyrus and supramarginal gyrus, in order to decrease delayed feedbacks interfering effects. Then it was assumed that the associational learning is asymmetric which means presynaptic activities are followed by postsynaptic activities [8].

In summary, the production of utterance starts with random activities in premotor cortex which are triggered by premotor drives and ends up producing stereotyped patterns of activity in primary cortex. The source of premotor drive is considered to be effects of basal ganglia modulation of motor cortical commands [9].

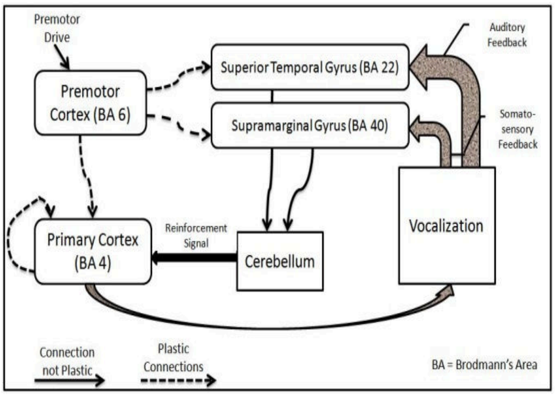


Fig. 1 Conceptual model architecture

III. Methods

In this paper, five neural populations are considered which contain several cell assemblies, corresponding to language related areas in brain. Their interactions are based on the mentioned conceptual model in section II. Each syllable produced by proposed computer model is a combination of 50 vocal features, while each primary cortex assembly represents motor related aspects of one feature, and each auditory or orosensory assembly represents sensory related aspects of one feature. Premotor cortex population consists of 250 assemblies. The cerebellum contains 5 assemblies corresponding to tutor syllables. The tutor speech consists of 5 syllables (indexed by letters A-E) and each syllable is encoded by an individual set of assemblies [8].

The output of each neural unit represents the activity rate within a corresponding cell assembly. The activity of each neural unit is encoded by the average of neural firing rates over each syllable. Neural firing rates are assumed to be constant in the course of premotor drives for each neural population except for superior temporal gyrus and supramarginal gyrus and zero during the gap between syllables. In the superior temporal gyrus and supramarginal gyrus each syllable is divided into four time stages based on the combination of efference copy, auditory and somatosensory feedback inputs received during that syllable. So that during the early part of each syllable, efference copy, which relates to the current syllable, and delayed auditory and somatosensory feedbacks from the previous syllable are received. During the middle part of each syllable only efference copy is received; and during the late part of each syllable, efference copy and auditory and somatosensory feedbacks that correspond to the same syllable are received. Finally, during the gap part of each syllable, only auditory and somatosensory feedbacks are received. The activities that are passed on to the cerebellum are calculated from the average activity in the superior temporal gyrus and supramarginal gyrus during the early and middle part of each syllable [8].

To simulate the proposed model, simple associational and reinforcement learning rules are used. The associational learning rule is based on analogies with NMDA receptor dependent long term potentiation (LTP). The synaptic strength change is calculated as follows:

(reinforcement × plasticity trace × postsynaptic activity – threshold) × presynaptic activity = (1)

$$[R\alpha(t-t^{pre})\ a^{pos}t(t) - Thre]\ a^{pre}(t^{pre}).$$

where, $a^{pre}(t)$ and $a^{pos}t(t)$ represent the activity level of the pre and postsynaptic assemblies at time t. α denotes plasticity trace that is proportional to the amount of the NMDA-receptor binding and determines the time window for neural plasticity. Multiplying plasticity trace by postsynaptic activity results in plasticity signal, $\alpha(t-t^{pre})\ a^{post}(t)$, proportional to postsynaptic calcium concentration. Furthermore, cerebellum is assumed to transmit a reinforcement signal 'R' which modulates the plasticity signal in all primary cortex assemblies. Moreover, Plasticity signals above a threshold value 'Thre' increase synaptic strength while signals below 'Thre' leads to long term depression. This threshold value is related to the average of activity in the postsynaptic assembly [8].

Among all neural populations, only primary cortex includes intrinsic excitatory connections. Also, each population includes a single inhibitory assembly which is connected to all assemblies in the corresponding population. This inhibition leads to a competition among excitatory assemblies. In addition to these local circuit mechanisms, two homeostatic mechanisms are considered. The first is normalization of synaptic strength while presynaptic normalization is applied before postsynaptic ones. The second mechanism is inhibitory plasticity which makes inhibitory connection strengths relevant to excitatory assembly activities [8].

For all neural population connections except the intrinsic primary cortex connections, initial connection strengths are based on single-projection strategy, in which each presynaptic assembly connects to a single postsynaptic assembly. This ensures the independence between any two assembly inputs in the postsynaptic populations. For intrinsic primary cortex connections to avoid correlations arising from multisynaptic pathways, a "uniform" strategy is used, in which each presynaptic assembly connects to all postsynaptic assemblies with the same strength. In addition, a zero mean Gaussian noise with a standard deviation equal to 10% of the strength of the nonzero synapses is added to all plastic connections during the initialization phase. One should notice that all negative strengths are set to zero after adding the noise [8].

IV. Results

To evaluate the validity of the proposed model, simulation of each syllable involves numerous iterations of three subroutines: 1) calculating activity patterns corresponding to a single syllable 2) applying the synaptic plasticity rule, and finally 3) updating the homeostatic mechanisms in the model. These steps are repeated for 30000 syllables.

Cerebellum conducts syllable learning by transmitting a reinforcement signal to modulate plasticity in all primary cortex assemblies. The results of reinforcement based syllable learning and also initial phase of learning are shown in Fig. 2 and Fig. 3.

Initially, primary cortex connectivity is nearly uniform while the activity pattern of sensory related areas and also primary cortex are random. Note that self connections (diagonal entries) are set to zero in order to prevent self correlations.

The progress of reinforcement learning results in similar patterns of connectivity for assemblies encoding the same tutor syllable. Thus the pattern of intrinsic primary cortex connections starts to show blocks of strong connections along the diagonal due to assemblies encoding the same tutor syllable are arranged next to each other. This pattern of intrinsic primary cortex connectivity gives rise to the production of primary cortex activity matched to the tutor template. The progress of reinforcement learning also leads to configuration of correlated pattern of activity in sensory related areas. Also, syllables are produced in a random sequence since premotor cortex is driven by the random premotor drive.

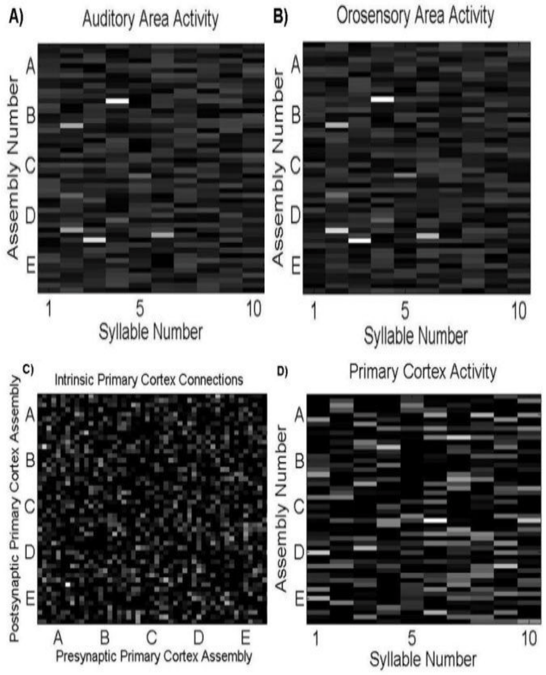
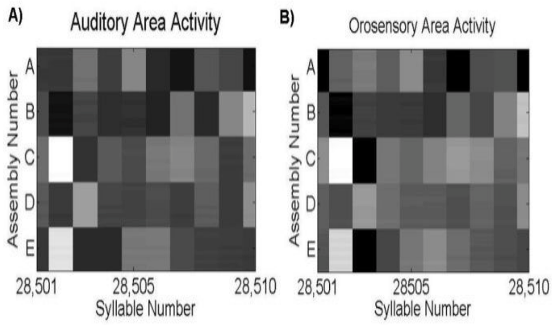


Fig. 2 Initial phase. A) Auditory area activity. B) Orosensory area activity. C) Intrinsic primary cortex connections. D) Primary cortex activity.



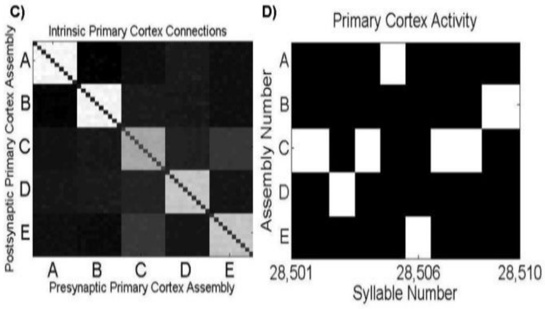


Fig. 3 After learning phase. A) Auditory area activity. B) Orosensory area activity. C) Intrinsic primary cortex connections. D) Primary cortex activity.

V. Conclusions

In this paper, a computer model of sensorymotor language acquisition using simple associational and reinforcement learning rules is outlined. This model is guided by a conceptual neural model of speech motor control, DIVA.

In proposed computer model, first, initially random premotor activities in premotor cortex are associated with auditory and somatosensory feedbacks using simple Hebbian learning. This step yields to efference copy signals. Then, efference copy signals in cooperation with auditory and somatosensory feedbacks result in indicator signals which are mapped through the cerebellum. Based on comparison between these indicator signals and stored templates, a reinforcement signal is transmitted by the cerebellum. This reinforcement signal modulates intrinsic plastic connections within primary cortex as well as the projection from premotor cortex. Finally, stereotyped sequences of primary cortex activities as well as sensory activities in sensory related areas are produced. In summary, the proposed computer model starts with random activation in premotor cortex and ends up producing exact syllables. The results, which are obtained from computer simulations, show that the computer model is reasonably valid.

The proposed computer model may be a starting point for further investigations into language acquisitions. Also, Speech disorders can be simulated by damage to neural units of the model that correspond to language related areas in the brain [9]. Furthermore, this model has great potential for studying other acquisition theories.

References

[1] H. Markert, A. Knoblauch, and G. Palm, "Modelling of syntactical processing in the cortex," *J. BioSystems*, vol. 89, 2007, pp. 300-315.

[2] P. J. Buttery, "Computational models for first language acquisition," Computer Laboratory, University of Cambridge, Tech. Rep. UCAM-CL-TR-675, Nov. 2006.

[3] M. T. Ullman, "Contributions of memory circuits to language: the declarative/procedural model," *J. Cognition*, vol. 92, 2004, pp. 231-270.

[4] A. Portoa, A. Araqueb, J. Rabunal, J. Dorado, and A. Pazos, "A new hybrid evolutionary mechanism based on unsupervised learning for Connectionist Systems," *J. Neurocomputing*, vol. 70, 2007, pp. 2799-2808.

[5] T. Wennekers, M. Garagnani, and F. Pulvermuller, "Language models based on Hebbian cell assemblies," *J. Physiology*, vol. 100, 2006, pp. 16-30.

[6] F. H. Guenther, "Cortical interactions underlying the production of speech sounds," *J. Communication Disorders*, vol. 39, 2006, pp. 350-365.

[7] F. H. Guenther, "Neural modeling of speech production," in *Proc. 4th Int. Nijmegen Speech Motor Conf.*, Nijmegen, Netherland, June 13-16, 2001.

[8] T. W. Troyer, and A. J. Doupe, "An associational model of birdsong sensorimotor learning. I. efference copy and the learning of song syllables," *J. Neurophysiol*, vol. 84, 2000, pp. 1204-1223.

[9] F. H. Guenther, and S. S. Ghosh, "A model of cortical and cerebellar function in speech", in *Proc. of the XVth Int. Cong. of Phonetic Science*, Barcelona, Spain, 2003.