Neural Networks 21 (2008) 1224–1231

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Developmental learning of complex syntactical song in the Bengalese finch: A neural network model

Yuichi Yamashita [a,*], Miki Takahasi [b], Tetsu Okumura [a], Maki Ikebuchi [b], Hiroko Yamada [b], Madoka Suzuki [b], Kazuo Okanoya [b], Jun Tani [a]

[a] *Lab. for Behavior and Dynamic Cognition, Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama, 3510198, Japan*

[b] *Lab. for Biolinguistics, Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama, 3510198, Japan*

ARTICLE INFO

Article history:
Received 2 March 2007
Received in revised form
12 February 2008
Accepted 21 March 2008

Keywords:
Recurrent neural network
Noise
Birdsong
Development
HVC
NIf
Zebra finch

ABSTRACT

We developed a neural network model for studying neural mechanisms underlying complex syntactical songs of the Bengalese finch, which result from interactions between sensori-motor nuclei, the nucleus HVC (HVC) and the nucleus interfacialis (NIf). Results of simulations are tested by comparison with the song development of real young birds learning the same songs from their fathers. The model shows that complex syntactical songs can be reproduced from the simple interaction between the deterministic dynamics of a recurrent neural network and random noise. Features of the learning process in the simulations show similar trends to those observed in empirical data on the song development of real birds. These observations suggest that the temporal note sequences of songs take the form of a dynamical process involving recurrent connections in the network of the HVC, as opposed to feedforward activities, the mechanism proposed in the previous model.

1. Introduction

Due to its similarity to human language in being a learned complex sequential behavior, birdsong, has become a widely studied topic in neuroscience. The Bengalese finch in particular learns highly complex songs that have syntactical structure, providing researchers with a good biological model for studying this phenomenon.

Fig. 1 shows a typical sound spectrogram of the song note sequences of the Bengalese finch. The song consists of several varieties of “notes”, the smallest units of a birdsong. Each note can be identified as a discrete element on a sound spectrogram and is denoted by a letter of the alphabet, for example “a”, “b” or “c”.

Note-to-note transitions follow rules. However, the transitions are not determined by the preceding note output alone, but are dependent on past sequences. This indicates that the Bengalese finch’s song sequence has a hidden state in the sense that the next transition state cannot be uniquely determined by the output note.

Note-to-note transition rules of the Bengalese finch’s song can be described using a finite state automaton (Honda & Okanoya, 1999). Normally the automaton describing a Bengalese finch’s song has probabilistic branching and recursive connections. A series of notes without branching constitute what is referred to as a “chunk”, and sequences of chunks generate diverse “motifs”. Owing to the recursive structure of the automaton describing their songs, the Bengalese finch is considered to generate an almost infinite number of different motifs. The complexity of this song structure is in contrast to the linearity of the songs produced by the Zebra finch, a bird which is nonetheless a close relative of the Bengalese finch (Zann, 1996).

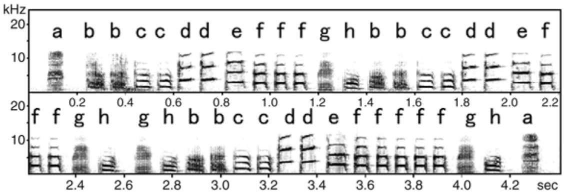

Fig. 1. Sound spectrogram of the song note sequence of Bengalese finches: Each note is identified as a discrete element on the sound spectrogram and is denoted by a letter of the alphabet.

The acquisition and production of songs are made possible by a group of discrete brain nuclei and their connecting pathways, referred to as the song system (Fig. 2) (Nottebohm, 2005). Within the song system the nucleus HVC (HVC), a premotor nucleus, plays a key part in generating temporal patterns of songs. In lesion studies of the Canary, a lesion on the HVC severely disturbs the temporal structure of songs (Nottebohm, Stokes, & Leonard, 1976).

---

* Corresponding author. Tel.: +81 48 467 1111x7415; fax: +81 48 467 7248.
E-mail address: yamay@brain.riken.jp (Y. Yamashita).

0893-6080/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2008.03.003

List of symbols:

w_{ij}	weight value from the *j*th unit to the *i*th unit
$u_{i,t}$	internal state of the *i* th unit at time *t*
$x_{i,t}$	neural state of the *i* th unit at time *t*
$y_{i,t}$	activation of the *i* th unit at time *t*
$y*_{i,t}$	desired activation value of output units at time *t*
$e_{i,t,}$	error between desired value and actual value of output activation at time *t*
E	learning error
T	length of note sequences
O	set of indices corresponding to the output units
N	total number of units
α	learning rate
θ_i	threshold of the *i*th unit
G	noise added to the internal value of non-output units
$F_{\max}$	transformed value of the unit with the highest activation in the winner-take-all computation
$F_{\min}$	transformed value of all other output units in the winner-take-all computation
A	component notes of a particular song
D	a particular song syntax with a probabilistic distribution of strings
$P_D(x)$	occurrence probability of string *x* under distribution *D*

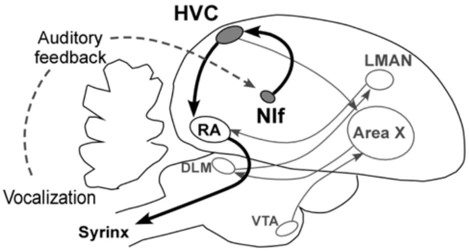

Fig. 2. Neural basis of the birdsong referred to as the song system: The NIf–HVC–RA pathway acts as the song production pathway. The other pathway (not highlighted) is responsible for song learning. LMAN: lateral magnocellular nucleus of anterior nidopallium, DLM: medial nucleus of the dorsolateral thalamus, VTA: ventral tegmental area.

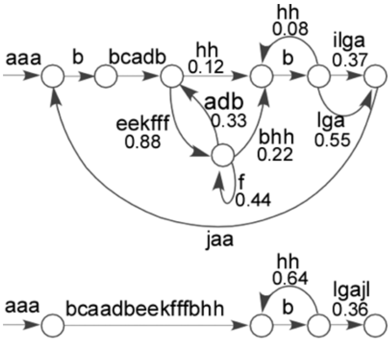

Fig. 3. Changes in a song automaton as the result of a lesion in the NIf. The upper and lower graphs correspond to pre- and post-lesion automata, respectively. Numerals indicate the probabilities of each branch. (Graphs are modified from Hosino and Okanoya (2000) with permission).

In electrophysiological studies of the Zebra finch, the activation pattern of each HVC neuron is highly context dependent and corresponds to a particular moment in a song (Fee, Kozhevnikov, & Hahnloser, 2004). Moreover HVC stimulation by a microelectrode of a singing bird causes an interruption in temporal patterns of the birdsong, whereas stimulation of the robust nucleus of arcopallium (RA), a downstream motor nucleus, affects only a particular note at the time of the stimulation (Vu, Mazurek, & Kuo, 1994). These facts strongly suggest that the HVC is a temporal pattern generator for song note sequences.

The nucleus interfacialis (NIf), one of the upstream parts of the HVC, is considered to be one of the essential regions that generate complexity in birdsongs. Lesions of the NIf reduce the branching of note-to-note transitions, however note sequences still correspond to paths on the original diagram (Fig. 3) (Hosino & Okanoya, 2000). The reduction of complexity occurs only in birds having songs that are complex and not in birds that have simple songs. Based on this observation it is inferred that the NIf, in cooperation with the HVC, provides complexity for the generation of songs.

However, the types of interaction between the HVC and the NIf that can produce complex syntactical songs and the representation of the temporal patterns of songs in the HVC remain unclear. That these questions have not yet been answered is a consequence of the technical difficulties associated with investigating the actual interactions between brain regions of singing birds and developmental changes of such interactions.

In order to overcome these difficulties, a number of models have been proposed. For example, Fee's group proposed a model in which the HVC generates temporal patterns of songs by cooperating with the RA (Fee et al., 2004; Fiete, Hahnloser, Fee, & Seung, 2004). In this model, it is assumed that temporal patterns of songs are represented as feedforward activities of the HVC, the role of which is analogous to a recording "tape" (Fee et al., 2004). It is also assumed that these feedforward activities innately exist and that the order of the notes in a song motif is acquired as a result of changes in connections between the HVC and the RA, where connections correspond to the mapping between an innately existing "tape" and each note sound. Doya and Sejnowski also developed a model based on the similar assumption (Doya & Sejnowski, 1995). Although these models are sufficient for explaining the song generation of the Zebra finch, the songs of which are very linear, they are not sufficient to explain the song generation of the complex syntactical song of the Bengalese finch.

Only a few studies exist that consider the question of NIf function. Hoshino developed a statistical model of the song learning of the Bengalese finch using hierarchical Bayesian networks (Hoshino & Doya, 2000). In this model, the hierarchical structure of the note-chunk-motif is assumed to reflect the functional hierarchy of nuclei RA–HVC–NIf and the NIf is assumed to represent and control chunk-to-chunk transitions. However, the problem of this kind of model is that it is necessary to arbitrarily set a number of hyperparameters, for example different time constants for each hierarchy. Due to the difficulty of setting these parameters, the model was only successful in reproducing simple artificial model songs, and not in reproducing song sequences of the actual Bengalese finch.

The objective of the current study is to investigate the following two questions. First, what types of neural connectivity are able to generate the complex syntactical song sequences of the Bengalese finch, which are considered to potentially generate an infinite number of motif patterns? Secondly, what is the contribution of the NIf in song production of the Bengalese finch?

In order to describe complex temporal sequences like those of the Bengalese finch's song, which have hidden states and recursive structure, systems with temporal delay and recurrent connections are superior to the linear system proposed in the previous study. Hidden Markov models (HMM), which are equivalent to probabilistic finite state automata, are one of the most popular examples of this kind of system. However, because HMMs use symbol-level abstraction, they are not sufficient for describing the neural basis for the observed phenomenon. On the other hand, although physiologically detailed models such as those that use