Vol.2010-MPS-81 No.8
Vol.2010-BIO-23 No.8
2010/12/16

*Technical Note*

# Information-Theoretic Analysis for Understanding the Behavior of Song Learning by the Bengalese Finch

Khan Md. Mahfuzus Salam,[†1] Tetsuro Nishino,[†1]
Kazutoshi Sasahara,[†2] Miki Takahasi[†2]
and Kazuo Okanoya[†2]

Songbirds have been actively studied for their complex brain mechanism of sensor-motor integration during song learning. Male Bengalese finches learn singing by imitating external models. Birdsongs are strings of sounds represented by a sequence of letters called song notes. In this study, we focused on information-theoretic analysis of these sequential data to explore the complexity and diversity of birdsong throughout song development. We present ethological data mining results for birdsong, which showed that there are 2 types of song development: linear and non-linear. We found that all male birds in the early stages of development sing complex songs, which are gradually crystallized by the elimination of extra transitions. In contrast, this is not true in the case of non-linear song development, wherein there are no practice modes, and all activity counts toward selecting, constructing, and maintaining behavioral outcomes. We also found that contour surface diagrams of the transition matrix can be a good visual representation to distinguish song features. Our results indicate that information-theoretic analysis of behavioral sequences is important for the discovery of new aspects related to behavioral science.

## 1. Introduction

Ethology is the scientific study of animal behavior for exploring mechanisms underlying diverse forms of behaviors, from unlearned stereotyped ones to learned flexible ones. Songbirds have been actively studied as a good ethological model for their complex brain mechanism of sensor-motor integration in song learning. Male songbirds learn singing by imitating external models. Birdsongs are strings of sounds represented by a sequence of letters known as a song note. Few studies have focused on developmental data from the information-theoretic viewpoint to understand the process by which songbirds learn to sing. The Bengalese finch (*Lonchura striata var. domestica*) is a domesticated strain of a Southeast Asian finch, the white-rumped munia (*Lonchura striata*). It is been a popular subject for neurobiological and ethological studies on birdsongs for its unique song features. The males of this species acquire their own songs by learning from external tutors (father or other males) during a specific period, between the nestling and fledgling stages. The learned features of the songs change very little during the matured stage of life. Their songs are used for courtship display. Two types of song features are preferred by female birds: performance- and elaboration-related traits. The performance-related trait is associated with the extent of song production, song rate, song duration, song speed, song amplitude, etc. In contrast, the elaboration-related features are associated with song rules and complexity. In the present study, we focused on elaboration-related features from 2 aspects: the first is analysis of song development and the other is comparison between the songs of the parent and progeny. Our aim is to conduct information-theoretic analysis on sequential song data to explore the complexity and diversity of birdsong throughout song development.

†1 Information and Communication Engineering, The University of Electro-Communications, Tokyo, Japan.
†2 Laboratory for Biolinguistics, RIKEN Brain Science Institute (BSI), Saitama, Japan.

## 2. Preliminaries

This section briefly introduces the theoretical foundations of birdsong and its representations, information-theoretic measures, and song model representation by k-reversible language.

### 2.1 Language and Birdsong

Humans use language to express emotions or to communicate with other humans, and language is considered unique to the human race. However, other living creatures, too, communicate vocally, for example, songbirds and dolphins, which have complex vocal communication homologous to human language.

In the 1960s, the prominent linguist C.F. Hockett divided "the design features of language" into 13 10). The proposed design features are as follows: (1) auditory-vocal channel, (2) broadcast transmission and directional reception in auditory signals, (3) rapid fading of auditory signals, (4) interchangeability in communication, (5) total feedback, (6) specialization, (7) semantics, (8) arbi-

Vol.2010-MPS-81 No.8
Vol.2010-BIO-23 No.8
2010/12/16

trariness, (9) discreteness, (10) displacement, (11) productivity (creativity), (12) traditional transmission, and (13) duality. On comparing animal and human communication, Hockett concluded that of the 13 features, only 2 – "traditional transmission" and "duality" – were not observed in any animals. Here, "traditional transmission" indicates that linguistic knowledge is passed on from one generation to the next through learning, and "duality" indicates that particular sound elements have no intrinsic meaning but combine to form structures (e.g., words and phrases) that have meaning. However, recent ethological studies have revealed that traditional transmission is found in the vocal communication of songbirds and whales, and these organisms can combine a few discrete sound elements and so even exhibit duality in communication 5). Therefore, these properties are not unique to human communication.

The song of the Bengalese finch has a more complex structure than that of other songbirds, such as zebra finches (*Taeniopygia guttata*) 3). According to recent studies, the courtship songs of the Bengalese finch have unique features and are similar to human language 5). Some research shows that the language model of this bird can be represented by a $k$-reversible automaton 12). Thus, because of the structural and functional similarities in vocal leaning between songbirds and humans, the former have been actively studied as good linguistic models. In particular, the song syntax of the Bengalese finch sheds light on the biological foundations of syntax in humans 1).

### 2.2 Song of the Bengalese Finch and Its Representation

Recent studies on Bengalese finches have shown that the songs of the male birds are neither monotonous nor random; they consist of chunks, each of which is a fixed sequence of a few song notes. The song of each individual can be represented by a finite automaton, which is called song syntax (**Fig. 1**) 5). Thus, the songs of Bengalese finches have "double articulation," which is one of the important structures of human language (i.e., a sentence consists of words, and a word consist of phonemes). Song syntax is controlled by song control nuclei in the brain. The hierarchy of song control nuclei directly corresponds to the song hierarchy 4).

Bird song analysis requires song data that have been recorded in a suitable environment. From the recorded wave data, spectrograms are obtained, and these

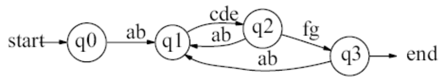


**Fig. 1** Courtship song syntax represented by an automaton

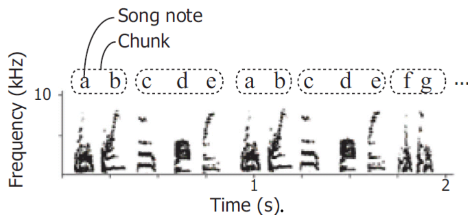


**Fig. 2** Grayscale spectrogram of a Bengalese finch's song

are used as the standard representation of the song. Below are brief explanations of some general terms used in birdsong research.

*Song note*: A song note is a symbol assigned for an independent pattern that appears in a sonogram as seen in **Fig. 2**. It is also referred to as a song element or behavioral element. From the definition, we can say the text data comprising symbols (such as *a, b,* and *c*) are called song note sequences. Song notes are analogous to phonemes in human language.

*Chunk*: A fixed sequence of song notes is called a chunk. In Fig. 2, for example, the chunks are *ab, cde,* and *fg*. Chunks are analogous to words in human language.

*Song unit*: A song unit consists of many chunks. Song units are analogous to sentences in human language.

### 2.3 Information-Theoretic Measures

In this section, we briefly discuss the information-theoretic measures that describe the features in behavioral strings and help understand their diversity.

*Transition matrix*: In birdsong research, a transition matrix is widely used to understand syntactical complexity. A transition matrix is one that shows note-to-note transition information. A transition probability matrix can be obtained by dividing the note-to-note outcome by the total number of transitions. This is the most common and important way to represent transition information, and other properties can be analyzed using this matrix.

*Linearity*: Scharff and Nottebohm introduced the measure of linearity for estimating the ordering complexity of notes 9). The linearity index score is calculated from the number of note types and transition types as follows:

$$S_{Linearity} = \frac{number of different notes/song}{number of transition types/song}$$

The linearity index score provides an estimate to predict the next note in a song when the previous note is known. In a completely linear song sequence, each note has only 1 transition type. Thus, a complete linear song has a linear index score of 1. If some notes have more transition types, the score becomes less than 1. Therefore, a lower linearity index score reflects a more syntactically complex song.

*Entropy*: Shannon introduced the concept of information entropy and discussed the details of entropy in printed English 8). Information entropy is a statistical measure of the uncertainty associated with a random variable. It quantifies, in certain sense, how much of information is produced on average for each letter of a certain text. Information entropy is denoted by

$$H(x) = -\sum_{i=1}^{n} P(x_i) log P(x_i)$$

*n-gram statistics*: An $n$-gram is a sub-sequence of n items from a given sequence. The items can be phonemes, syllables, letters, or words depending on the application. The size of the ordered list of elements is denoted by n. An n-gram of size n = 1 is referred to as a "unigram"; n = 2, as a "bigram"; and n = 3, as a "trigram." One with $n = 4$ or more is simply called an "$n$-gram". "$(n-1)$-order Markov models" are language models built from $n$-grams. In ethology, $n$-gram statistics is used to understand frequency distribution and the hierarchy of behavioral patterns.

## 3. Data Mining and Information Extraction

This section describes the application of the data mining technique to extract information from behavioral sequences.

### 3.1 Data Mining

Data mining is the process of analyzing data from different perspectives and summarizing it into useful information. Various approaches can be used to analyze data for data mining. One approach is to use an analyzing tool that allows users to analyze data from many different dimensions, categorize it, and summarize the relationships identified. In general, data mining is the process of finding correlations or patterns among relational datasets.

### 3.2 Mining Behavioral Sequential Data

In ethological studies, there are 3 steps for understanding animal behavior. First, a behavioral phenomenon is observed and recorded; second, on the basis of the recordings, a hypothesis is formulated to explain the behavior; and third, experiments are designed, performed, and evaluated to test the hypothesis 13). If these procedures are followed correctly, better predictions can be made concerning animal behavior, which in turn provide insights into human behavior. However, developing a hypothesis on the basis of recordings is not always easy if the data is vast or complex. Therefore, we apply data mining technology in ethology. In particular, we studied the application of ethological data mining from sequential animal behavior.

In general, animal behavior is recorded as sequential data of behavioral events. The same symbol is assigned to an identical behavioral event type, and the behavioral data is converted into text data, which are then used for data mining. Such sequential data of animal behavior can be analyzed for different statistical and information-theoretic measures, such as the transition matrix and the first-order Markov chain, and entropy 12). Special tools are required for dealing with complex behavioral data, and these should be used together with conventional tools. Data mining is one way to do handle complex behavioral data. The current data mining process has many well-established techniques for pattern extraction, clustering, modeling, etc 13). It can help find units from behavioral sequences and extract the rule that governs them. For obtaining significant information from animal behavior, we have to carefully select elemental data mining techniques suitable for behavioral sequences, which might be different from both word sequences in natural languages and biological sequences like DNA, and we have to then use them with proper modification in the context of ethology.

In this study, we used the song of the Bengalese finch to analyze song data during song development. The recorded songs were analyzed with sound analysis software to generate a spectrogram, which was used to convert sounds (WAV format) to texts. For the song of the Bengalese finch, we obtained a spectrogram where the different song notes were separated by a considerable gap. Each identical pattern was labeled with a similar symbol like *'a,' 'b,'* etc. Thus, the song was represented in text format. This analysis was performed manually on the basis of the phonological properties of song notes. Thus, by some pre-processing, we converted the song data into text for further analysis. For data mining from the behavioral sequence, we developed a *EUREKA*, which is a utility suite used for the following analyses:

- Information-theoretic analysis
- Extraction of probabilistic behavioral rule ($n$-gram model)
- Extraction of deterministic behavioral rule (deterministic finite automaton)

The module *StringStat* deals with the analysis of information-theoretic measures and enables the following translation in the context of ethology:

## 4. Data Mining from Birdsong

In this section, we present our results from 2 aspects. The first is the analysis of song development and the other is comparison of song property between the parent and progeny. We consider some of the measures in our analysis provided by EUREKA.

### 4.1 Description of Data

In the current study, we examined 9 juvenile birds to study song evolution and 4 parent-progeny pairs – "Sankakukoshitya-Lao," "Bakatono-RAo,," "Katsuo-LDai," "Kuroshiro-Shiro" – to compare the songs between parent and progeny. All birds were kept in the same environment inside a large cage in the laboratory. Their vocalizations were recorded to analyze their songs. Each bird was individually placed in a soundproof room, and its vocal output was recorded with a directional microphone and a DAT recorder.

It has been reported that young male birds learn songs from their fathers in the first 120–130 days after hatching 2). The principal learning period is considered the age of 60–130 days. The songs in the first 60 days after hatching are distorted.

### 4.2 Common Parent-Progeny Findings

We present our results regarding features that were common to the songs of the parents and progeny. For this analysis, we considered only the transition matrix and transition types. We also present a new technique for visual representation of the song on the basis of the transition matrix.

The birds have 7–14 different note types. Our findings show that every bird has 1 or 2 dominant notes, while other notes are used to generate variations in the song. **Fig. 3** shows the graph of rank vs. frequency for the different pattern types.

The graph (Fig. 3) is generated from the song sung by the bird *Shiro* at age 120 days. The complete song consists of 10 bouts comprising 816 song notes. The song note types are labeled as *a, b, c, d, e, f, g,* and *h*. We found that of these song notes, *d* appeared 317 times, and *f* appeared 228 times. Further, *a* appeared 123 times, and the other notes appeared less than 100 times. When we increased the value of $n$, we found that *dd* appeared in the song sequence 264 times; *ff*, 159 times; *ddd*, 211 times; *fff*, 102 times; and *dddd*, 158 times. From this simple analysis, we found that the notes d and f are the dominant song notes for *Shiro*.

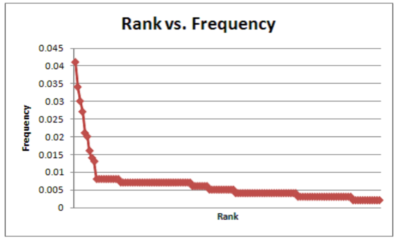


**Fig. 3** Rank vs. frequency of $n$-gram types

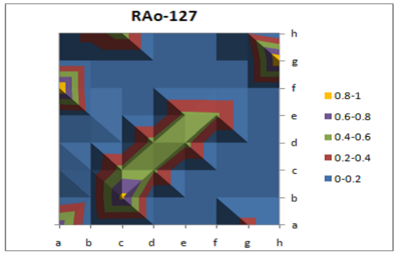


**Fig. 4** Contour surfce diagram of bird RAo

We also show that a contour surface diagram based on the transition matrix can be used for visual representation of the song. Surface charts are useful 3D chart types; they have 3 true data dimensions and can illustrate data reasonably well. Surface charts are useful to show how a variable (Z) changes according to 2 other variables (X and Y). Contour charts are a kind of surface chart containing regions colored according to the Z value. Essentially, they are 2D top views of 3D surface charts. **Fig. 4** shows a contour chart for the birdsong produced by bird *RAo* at age 127 days. Since the transition matrix shows only numbers, it is difficult to understand the song patterns from the matrix, but if we represent the song as a contour diagram, we can easily visualize its transition properties.

### 4.3 Analysis of Evolution

Here, we present the results of evolution analysis. We considered only the young bird's song data for this analysis. Although *EUREKA* provides different information-theoretic measures, we focused on linearity, entropy, and $n$-gram statistics in particular transition types. We found that different measures are not constant over time during song development.

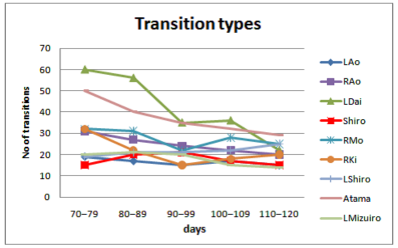


**Fig. 5** Change in number of transition types during song development

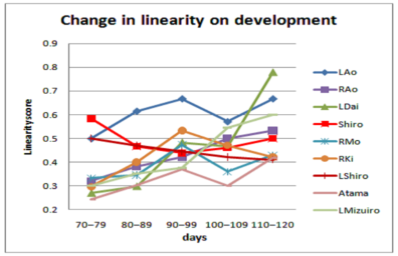


**Fig. 6** Change in the linearity index score during song development

In the early stage, most birds have a relatively large number of transition types in their songs. Eventually, the noise transitions are reduced, and the birds produce relatively small patterns. However, some birds produce songs with a relatively small pattern from an early age, and the pattern does not change later during song development. We found the song development of *Shiro* and *LShiro* to show such a trend (**Fig. 5**).

The linearity index score indicates the complexity of the song. Our investigation of this measure showed that all the birds in this study sang complex songs in the beginning, although after the development period, the crystallized songs were syntactically simpler. Again, the experimental results showed that *Shiro* and *LShiro* started singing less complex songs from the beginning, and the linearity index score of these birds changed only slightly during song development (**Fig. 6**).

Our investigation indicates that there are are two types of learning found in Bengalese finches song development: linear and non-linear. There are no practice modes in non-linear song development. All activity counts toward selecting, constructing, and maintaining behavioral outcomes. This implies that that the fundamental, self-generated activity plays an important role in the development of behavioral functions, both perception and cognition related.

### 4.4 Parent-Progeny Comparison of Songs

This section shows results of a comparison between the songs of the parent and progeny. For this analysis, the song data of both the parent and progeny were considered. We focused on different information-theoretic measures such as linearity, entropy, and n-gram statistics in particular transition types. We also show contour diagrams of the different bird families (**Fig. 8**).

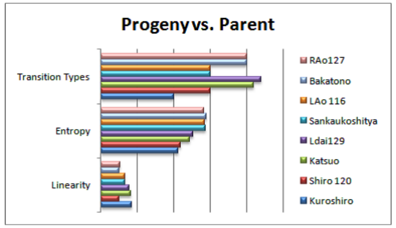


**Fig. 7** Comparison of different measures between parent and progeny

**Fig. 7** shows the comparison of information-theoretic measures between parent and progeny. It indicates that during linear learning, different measures converge to the value of the parents' song. Except the pair *"Kuroshiro–Shiro,"* all other pairs showed almost similar values of all measures between the parent and progeny. The learning process of *Shiro* was found to be non-linear. We can see from Fig. 7 that the transition types and linearity index score of this bird was very different from those of his father; in fact, *Shiro's* song was more complex than his fathers'.

Fig. 8 shows the contour diagrams of the transition matrix for the songs of the parents and progeny. The contour diagrams for the song at the early age and matured age are shown for the young birds. From Fig. 8, we can easily visualize 2 properties of the songs of the Bengalese finch: (1) Contour diagrams can visually display the unique song features of a particular bird family. We can easily distinguish different bird families from the contour diagrams. (2) Second, although there are differences in song properties at the early and matured stages, the major features are present at the early stage, and as the noise transitions reduce, the songs of the young birds eventually converge in to their fathers' songs.

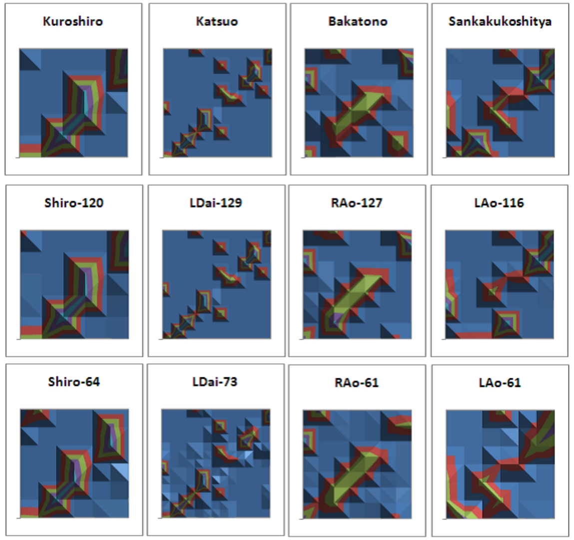


**Fig. 8** Contour diagrams of different families

## 5. Discussion & Conclusion

The main purpose of this study was to analyze birdsong for studying behavioral sequences. Application of information-theoretic analysis provides scope for closer examination of a large amount of data, whereby useful information can be extracted from them. Chatfield and Lemon 7) also concluded that information-theoretic measures can be very useful for analysis in the field of animal behavior.

Our findings showed that interesting phenomena occur during the learning process in songbirds, and we report a new kind of learning process. We also showed a new way to visually represent the features of behavioral sequences.

By applying the findings of the present study, we will be able to analyze animal behavior more precisely. For example, we could obtain a better understanding of the dominant song notes and their features and compare their sound properties and effect on female preference. Thus, we show that information-theoretic analysis of behavioral sequences can help us identify new aspects related to behavioral science.

## References

1) J. Doupe and P. K. Kuhl. *Birdsong and Human Speech: Common Themes and Mechanisms*. Annual Reviews Neuroscience 22:567–631, 1999.
2) C. K. Catchpole and P. J. B. Slater. *Bird Song: Biological Themes and Variations*. Cambridge University Press; 2nd edition, 2003.
3) E. Honda and K. Okanoya. *Acoustical and Syntactical Comparisons between Songs of the White-backed Munia (Lonchura striata) and Its Domesticated Strain, the Bengalese Finch (Lonchura striata var. domestica)*. Zoological Science 16:319–326, 1999.
4) J. Nishikawa and K. Okanoya. *Dynamical Neural Representation of Song Syntax in Bengalese Finch: A Model Study*. Ornithological Science 5:1, 95–103, 2006.
5) K. Okanoya. *Song Syntax in Bengalese Finches: Proximate and Ultimate Analyses*. Advances in the Study of Behavior 34:297–346, 2004.
6) R. Suzuki, J. R. Buck, and P. L. Tyack. *Information Entropy of Humpback Whale Songs*. Journal of the Acoustical Society of America 119:1849–1866, 2006.
7) C. Chatfield and R. E. Lemon. *Analyzing Sequences of Behavioral Events*. Journal of Theoretical Biology 29:427–445, 1970.
8) C. E. Shannon. *Prediction and Entropy of Printed English*. Bell System Technical Journal 3:50–64, 1950.
9) C. Scharff and F. Nottebohm. *A Comparative Study of the Behavioral Deficits following Lesions of Various Parts of the Zebra Finch Song System: Implications for Vocal Learning*. Journal of Neuroscience 11(9):2898–2913, September 1991.
10) C. F. Hocket. *The Origin of Speech*. Scientific American 203:88, 1960.
11) E. Thelen and L. B. Smith. *A Dynamic Systems Approach to the Development of Cognition and Action*. The MIT Press, 1994.
12) Y. Kakishita, K. Sasahara, T. Nishino, M. Takahasi and K. Okanoya. *Ethological Data Mining: an Automata-Based Approach to Extract Behavioral Units and Rules*. Data Mining and Knowledge Discovery 18(3):446–471, 2009.
13) J. Han, H. Cheng, D. Xin, and X. Yan. *Frequent Pattern Mining: Current Status and Future Directions*. Data Mining and Knowledge Discovery 15(1):55–86, 2007.
14) TNLAB, UEC & OKANOYA LAB, RIKEN Japan. *EUREKA*, URL: http://sites.google.com/site/eurekawiki; last accessed: November 1st, 2010.