A quantitative method for analyzing species-specific vocal sequence pattern and its developmental dynamics

Raimu Imai[a], Azusa Sawai[a], Shin Hayase[a], Hiroyuki Furukawa[a],
Chinweike Norman Asogwa[a], Miguel Sanchez[a], Hongdi Wang[a], Chihiro Mori[a],
Kazuhiro Wada[a,b,c,*]

[a] Graduate School of Life Science, Hokkaido University, Sapporo, Hokkaido 060-810, Japan
[b] Department of Biological Sciences, Hokkaido University, Sapporo, Hokkaido 060-810, Japan
[c] Faculty of Science, Hokkaido University, Sapporo, Hokkaido 060-810, Japan

HIGHLIGHTS

- New method is developed for discrimination of species specificity of vocalization.
- Prevalence of the characteristic syllable transitions is calculated.
- It quantitatively evaluates inter- and intraspecific differences of bird song.
- The method allows tracking the developmental trajectories of song patterns.

ARTICLE INFO

ABSTRACT

Background: Songbirds are a preeminent animal model for understanding the neural basis underlying the development and evolution of a complex learned behavior, bird song. However, only a few quantitative methods exist to analyze these species-specific sequential behaviors in multiple species using the same calculation method.

New method: We report a method of analysis that focuses on calculating the frequency of characteristic syllable transitions in songs. This method comprises two steps: The first step involves forming correlation matrices of syllable similarity scores, named syllable similarity matrices (SSMs); these are obtained by calculating the round-robin comparison of all the syllables in two songs, while maintaining the sequential order of syllables in the songs. In the second step, each occurrence rate of three patterns of binarized "2 rows × 2 columns" cells in the SSMs is calculated to extract information on the characteristic syllable transitions.

Results: The SSM analysis method allowed obtaining species-specific features of song patterns and intraspecies individual variability simultaneously. Furthermore, it enabled quantitative tracking of the developmental trajectory of the syllable sequence patterns.

Comparison with existing method: This method enables us to extract the species-specific song patterns and dissect the regulation of song syntax development without human-biased procedures for syllable identification. This method can be adapted to study the acoustic communication systems in several animal species, such as insects and mammals.

Conclusions: This present method provides a comprehensive qualitative approach for understanding the regulation of species specificity and its development in vocal learning.

© 2016 Elsevier B.V. All rights reserved.

* Corresponding author at: Faculty of Science, Department of Biological Sciences, Hokkaido University, Room 910, Building No. 5, North 10, West 8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan.
E-mail address: wada@sci.hokudai.ac.jp (K. Wada).

1. Introduction

The acquisition of species-specific behavioral patterns and their subsequent lifelong maintenance is a crucial research theme in order to understand the diversity and evolution of complex learned behaviors. Specifically, species-specific behaviors, such as mating, foraging, nest construction, and social communication, are generally composed of a rendition of temporal sequence patterns of behavioral units. These sequential behavioral characteristics are represented with species specificity even between closely related species (Berridge, 1990; Donaldson and Young, 2008; Greenspan and Ferveur, 2000; Weber and Hoekstra, 2009; Weber et al., 2013). Therefore, the sequential behavioral pattern is a critical species-specific biosignal for the recognition of conspecifics and the discrimination of heterospecifics in nature (Leininger and Kelley, 2015; Marler and Peters, 1988; Pollack and Hoy, 1979; Woolley and Moore, 2011). However, a limited number of methods have been developed for quantitative analysis of the temporal sequence of behavioral patterns.

Over 3000 species of songbirds around the world have species-specific song patterns. These patterns comprise stereotyped acoustic units, called syllables, which are arranged in characteristic species-specific sequential orders. Similar to that of human language, songbirds learn such song features through auditory and vocal experiences in a defined sensitive period during which juvenile songbirds listen, memorize, and gradually match their own developing vocalizations to the song pattern of an adult tutor (Konishi, 1965; Tchernichovski et al., 2001; Waser and Marler, 1977). An understanding of how syllables are sequentially arranged to form songs patterns is of general interest. This is because the learning and ordering of stereotyped motor patterns are the basis of vocal communication in vertebrates including speech in human. They thus may underlie other aspects of motor control. However, there exist some limitations on the analysis of vocal phenotypes. This was especially evident in the regulation of the temporal sequential patterns of syllables in many songbird species because of the difficulty in identifying each syllable (Gardner et al., 2005; Liu and Nottebohm, 2007). Acoustic features of syllables are variable in multiple factors, such as duration, frequency modulation, amplitude, and pitch, across song renditions and even in a song bout (Tchernichovski et al., 2000). Nevertheless, for the analysis of sequence or syntax of the song, categorization of syllable types was often performed based on human visual inspection of the spectrogram shapes. Although a few studies adapted semi-automatic clustering of syllables based on their mean acoustic features (Deregnaucourt et al., 2004; Ravbar et al., 2012; Wu et al., 2008), it remains difficult to set clear standards for identifying unique song syllables. This misleads the researcher given the observer-dependent bias on the identification of syllables, and precludes the extraction of precise information on the temporal sequence pattern of syllables in the song. A similar problem also occurs in the analysis of vocal patterns in mammals. For example the ultrasonic vocalizations produced by many species of rodents consist to a great extent of acoustically variable syllables (Holy and Guo, 2005). In addition, during vocal development, highly variable acoustic features of syllables are observed at the subsong and plastic song stages before the period of vocal crystallization. This makes it difficult to extract the quantitative information on the development of the temporal sequence of syllables.

Here, we adapted a correlation matrix capturing information on the similarity between syllables in two songs, termed as "syllable similarity matrix (SSM)," to extract songs' temporal structures by quantification of specific syllable transition types rather than transitions of identified syllables. This method allows the detection of the species specificity of song temporal patterns, and quantitative tracking of the developmental trajectories without human-biased procedures in syllable identification.

2. Materials and methods

2.1. Animals

All zebra finches (ZF, *Taeniopygia guttata*) and the Bengalese finches (BF, *Lonchura striata* var. *domestica*) were laboratory bred. Other species, such as owl finch (OF, *Taeniopygia bichenovii*), star finch (SF, *Neochmia ruficauda*), Java sparrow (JS, *Padda oryzivora*), and canary (CN, *Serinus canaria*) were obtained from local breeders. The photoperiod was constantly maintained at a 13:11 h light/dark cycle with food and water provided *ad libitum*. All bird experiments were performed according to the guidelines specified by the Committee on Animal Experiments at Hokkaido University. The guidelines are based on the national regulations for animal welfare in Japan (Law for the Humane Treatment and Management of Animals; after partial amendment No. 68, 2005).

2.2. Song recording and tutoring

Song recordings were performed as previously reported (Mori and Wada, 2015). Briefly, the birds were individually housed in a sound-attenuation box. A microphone (SHURE microphone SM57, SHURE incorporated, IL, USA) was connected to a FirePod FP10 amplifier (PreSonus Audio Electronics Inc. Florida, USA). Songs were recorded at a sampling rate of 44 kHz and 16 bit amplitude resolution by using Sound Analysis Pro version 1.04 (Tchernichovski et al., 2001). Songs were then automatically saved for 24 h per day by using the Sound Analysis Pro program. Low-frequency and high-frequency noises (<0.5 kHz and >15.8 kHz, respectively) were filtered using Avisoft SASLab pro software (Avisoft Bioacoustics, Berlin Germany). Audacity's noise-canceling procedure (version 2.0.5) was used to further filter ventilator and fluorescent-delivered noises.

With respect to the song tutoring experiment, the sex of ZF and BF chicks was determined by polymerase chain reaction within 5 post-hatching day (phd), as described previously (Wada et al., 2006). Between 6 and 15 phd, before juveniles could start memorizing a tutor song, they were separated from their fathers and were kept with their mothers and siblings until fledging. Song tutoring was started at 15–20 phd and continued through over 100 phd. After fledging, birds were individually housed in a sound-attenuating box attached containing a mirror to reduce social isolation. Tutor songs were played five times in the morning and five times in the afternoon at 55–75 dB from a speaker (SRS-M30, SONY) passively controlled by Sound Analysis Pro (v1.04) (Tchernichovski et al., 2000).

2.3. Calculation of similarity score between syllables

In this study, a song bout was defined as a continuous production of songs separated by a silent period longer than four folds of the mean value of more than 30 randomly selected inter-syllable gaps. Syllables in the song bouts were used for making SSMs. More than 50 syllables in song bouts were used as a song rendition. Introductory notes in a song were not included in a song rendition. The songs of the ZF, OF, SF, BF, and JS usually contained less than 50 syllables in a song bout. In these cases, two or more song bouts were merged as a combined song rendition to contain more than 50 syllables. For the CN, songs with more than 200 syllables were used because their songs consisted of a considerable number of syllable repetitions when compared with that of the other species. Song spectrograms were formed as WAVE format files (.wav) by Avisoft