Summer 8-6-2018

An Investigation of Vocal Learning Propensity in Zebra Finch (Taeniopygia guttata)

Tatsuya Hayashi
CUNY Hunter College

How does access to this work benefit you? Let us know!
Follow this and additional works at: https://academicworks.cuny.edu/hc_sas_etds
Part of the Animal Studies Commons, Behavioral Neurobiology Commons, and the Zoology Commons

Recommended Citation
https://academicworks.cuny.edu/hc_sas_etds/359

This Thesis is brought to you for free and open access by the Hunter College at CUNY Academic Works. It has been accepted for inclusion in School of Arts & Sciences Theses by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.
An Investigation of Vocal Learning Propensity in Zebra Finch (Taeniopygia guttata)

Tatsuya Hayashi
An Investigation of Vocal Learning Propensity in Zebra Finch (Taeniopygia guttata)

by

Tatsuya Hayashi

Submitted in partial fulfillment
of the requirements for the degree of

Master of Arts in

General Psychology, Hunter College

City University of New York

2018

Thesis Sponsor:

August 6th, 2018
Date

Dr. Ofer Tchernichovski
Signature of First Reader

August 6th, 2018
Date

Dr. Peter Moller
Signature of Second Reader
Abstract

The investigator attempted to predict the vocal learning propensity from either family pedigree or early vocal babbling (subsong). Pedigree trees suggest the possibility that vocal learning ability may depend on family. We failed to predict vocal learning propensity from subsong, but suggested that subsong may influence vocal learning outcome.

Keywords: songbird, zebra finches, correlational study, vocal learning propensity, imitative accuracy, family pedigree, subsong, individual difference, model-specific effect, syllable-specific effect
# Table of Contents

Abstract ......................................................................................... ii

Acknowledgments ........................................................................ vi

List of Figures ............................................................................... vii

Introduction ................................................................................ 1

Method ......................................................................................... 6

Data and Animal Care .................................................................. 6

Bird Sound Recording and Vocal Training ........................................ 6

Pedigree Tree ............................................................................... 7

Subsong Individual Difference ....................................................... 7

Correlation between Subsong and Imitative Accuracy ......................... 8

Results ......................................................................................... 9

Pedigree Tree ............................................................................... 9

Subsong Individual Difference ....................................................... 16

Correlations between Subsong and Imitative Accuracy ......................... 24

Discussion .................................................................................. 30

Appendices .................................................................................. 34

Glossary ...................................................................................... 38

The Brain Song System of Zebra Finch ............................................. 38

Acoustic Features .......................................................................... 39
Acknowledgments

I am deeply grateful to Ofer Tchernichovski for his continuous support and mentorship. He taught me everything, including the astonishing beauty of animal behavior, thorough observation of raw data, and importance of communication. I love he vocalizes “so” in euphoric manner. I am very proud of being his first Master’s student. I also want to thank the members of Tchernichovski lab; Dina Lipkind and Julia Hyland Bruno permitted me to use their recorded birdsong data and gave me the valuable comments on my research; Iva Ljubičić and the late Kirill Tokarev instructed the know-how to take care of birds and data use; Kenta Watanabe supported me emotionally. Moreover, I would like to express gratitude to Peter Moller for constructive feedback on my thesis defense and draft. His detailed editing helped me greatly to brush up this thesis. Without any of those persons, I could not have finished my Master’s degree.
List of Figures

Figure 1. Pedigree tree of zebra finches at the Tchernichovski lab. .................................................. 10-12

Figure 2. Examples of fast learners and families. .................................................................................. 14

Figure 3. The family which produced offspring with different learning speed possibly associated with the mother’s. ............................................................................................................. 15

Figure 4. Subsong sonograms exhibit within-individual stability and between-individual variability. ........................................................................................................................................... 17

Figure 5. Distributions of frequency modulation (FM) are stable within a bird but variable among birds................................................................................................................................................. 20-21

Figure 6. Distributions of syllable duration are stable within a bird but variable between birds. .... 22

Figure 7. Distributions of Wiener entropy are stable within a bird but variable between birds. .. 23

Figure 8. Distributions of Amplitude Modulation (AM). ..................................................................... 24

Figure 9. Different song models broadcast to juvenile birds. .............................................................. 27

Figure 10. Relation between subsong features and later imitative similarity to model songs. ....... 28

Figure 11. Relation between subsong features and later imitative similarity to model syllables. .......................................................... 29

Supplementary Figure 1. The probability densities in all acoustic features that can be calculated by SAP2011 on a single bird................................................................................................. 34

Supplementary Figure 2. The probability densities in all acoustic features that can be calculated by SAP2011 over subsong development. .................................................................................. 35
Supplementary Figure 3. The probability densities in all acoustic features that can be calculated by SAP2011, on different birds.

Supplementary Figure 4. Statistical evidence of robust individual difference in subsong acoustic features.

Supplementary Figure 5. The simplified diagram of the song system in zebra finch.

Supplementary Figure 6. The exemplar histogram and violin plot to visualize subsong acoustic feature.
Introduction

Male juveniles develop their songs by imitating matured songs of their father and other adult males with whom they interact. In the laboratory, juvenile males learn their songs between 30 and 90 day-post-hatch (dph), which is the sensitive period of vocal learning. This period is composed of two partially overlapping phases. During the first phase called sensory learning period, 20-50 dph, a juvenile male encodes the sensory memory of the model song he will imitate (Böhner, 1990). At the midway of this phase, the second phase called the sensory-motor learning phase initiates with bird vocal babbling. This vocal babbling called subsong, which is the focus of this study, involves the soft, highly variable and poor structure. As subsong is vocalized more frequently and louder, it is gradually turned into syllables, distinct units of sound segmented by silent gaps, and this more structured, yet variable song (“plastic song”). Eventually as syllables gets more complex structure and stable, the songs approach the model song. After 80-90dph the sexually matured males produce their song (“crystalized song”) with the high stereotypy comparable to those of older adults. The basic repeating unit of syllables sung by bird is called motif, and it is nearly identical to each other, and mutate little after 90dph (Immelmann, 1969).

Initially, subsong is vocalized by function of anterior forebrain pathway (AFP; Ölveczky et al., 2005; Aronov et al., 2008), which produces the acoustic variability in song throughout vocal development (Andalman & Fee, 2009; Charlesworth et al., 2011). Acoustic variability is used for vocal exploration that is critical for song learning, as disturbances of AFP prevent song learning (Bottjer et al., 1984; Scharff & Nottebohm, 1991; Brainard & Doupe, 2000; Haesler et al., 2007). As neural control gradually shifts to a second vocal center, HVC (Aronov et al., 2008), song patterns become stereotyped and eventually be crystallized (Hahnloser et al., 2002;
Kozhevnikov & Fee, 2007). This balance between acoustic variability and stereotypy during vocal development is temporally regulated to make immature parts of song variable, while keeping acquired parts of song stereotyped (Ravbar et al., 2012; Vallentin et al., 2016). Since AFP is required for production of subsong, it is speculated that acoustic features of subsong might reflect the individual ability of vocal exploration.

The process of vocal learning ends up in individually unique crystalized songs. Matured birdsong varies in its ‘song complexity’, which is the construct related to the number of different syllables or elements in zebra finch, and also different song repertoire in other songbirds (Neubauer, 1999; Buchanan & Catchpole, 2000; Pfaff et al., 2007). The individual zebra finch song consists of 3 to 14 different syllables (Clayton & Pröve, 1989). This song complexity is considered as a locus of preference by female, and females prefer males that sing more complex song. For example, male sedge warblers (*Acrocephalus schoenobanus*) with larger syllable type repertoires have been shown in field studies to obtain mates at an earlier date (Buchanan & Catchpole, 1997; Catchpole, 1980). Great reed warblers (*A. arundinaceus*) males with larger syllable repertoires also attract more social mates (Hasselquist, 1998). In the laboratory, females have been shown to perform more courtship displays in response to larger syllable repertoires in sedge warblers (Catchpole et al., 1984) and great reed warblers (Catchpole et al., 1986).

What is the factor contributing to the song complexity? Many studies revolve around the developmental stress hypothesis (Nowicki & Searcy, 2004), which explains the relation between song complexity and the brain development during song learning period. The song learning period is the life stage when birds are vulnerable to developmental stress, such as food insufficiency, social stressors or parasite infections that might deteriorate brain development. The proper development of the song system would act as an indicator of the birds’ ability to cope
with rough environment. Since the birds who can afford to develop better song circuit should also be able to vocalize more complex song, females can use song complexity as a sign of superior genetic and phenotypic qualities (Nowicki et al. 2002; Nowicki & Searcy, 2004; Nowicki & Searcy, 2005). The majority of experimental studies confirmed the detrimental effects of developmental stress on song circuit, song complexity and sexual attractiveness (Buchanan et al., 2003, 2004; Spencer et al., 2003, 2004, 2005a, b; MacDonald et al., 2006; Soma et al., 2006; Zann & Cash, 2008), while Gil et al. (2006) found no relations between song complexity and song nucleus size in the brain.

Aside from the song complexity, accuracy of song imitation can also be of great use as the measure of individual song difference, especially in case when birds are tutored with a single adult in laboratory. When juvenile male zebra finches reared by their mother alone and housed singly with an adult male, they usually produce a close copy of their tutor’s song (Haesler et al., 2007; Tchernichovski et al., 1999). The differences between the songs of a tutor and its pupil are often treated as errors (Derégnaucourt et al., 2004; Goodfellow & Slater, 1986; Slater & Ince, 1979). However, it should also be noted that complete imitation might not be adaptive, as imitative inaccuracy may facilitate individual recognition (Böhner, 1983). Past studies provided inconsistent evidence about whether developmental stress results in poor imitative accuracy (Brumm et al., 2009; Nowicki et al., 2002; Gil et al., 2006; Zann & Cash, 2008; Holveck et al., 2008). On the other hand, Tchernichovski and Nottebohm (1998) observed that birds with greater imitative accuracy tended to induce an earlier ovulation response in females, implicating that the imitative accuracy influences song perception by females.

Imitative inaccuracy in each individual can be attributed to three main factors (reviewed in Derégnaucourt, 2011). The first factor is the social context on song learning. For example,
song imitation was inhibited in the setting where more male siblings were present (Tchernichovski & Nottebohm 1998). Incomplete imitations were more common among early-hatched than among late-hatched chicks (Tchernichovski & Nottebohm 1998). A second factor is the characteristic of tutor songs; some songs are copied better than others as a result of particular acoustic features (e.g. syllable durations), and the ‘aesthetics’ of the songs may bias the imitative learning. For example, the isolate songs, which are the abnormal vocalization developed by the birds socially isolated from other adults, contains abnormally long notes or white-noise like sounds. Adult females prefer to mate with males who sing normal songs than isolate ones (Williams et al., 1993). The birds tutored with isolate songs, however, imitated songs in the biased manner that acoustic features of learned songs approximated those of normal songs, and this normalization was completed in 3-4 generations of tutoring (Fehér et al. 2009).

A third factor is the individual motivation or ability to learn songs: some birds might be qualified to imitate songs better than others. Chen et al. (2016) showed that juveniles that paid more attention to their tutor’s songs displayed greater imitative accuracy in adulthood. Whilst this relationship was significant for both socially and passively tutored juveniles, increases in attention linked to better imitation in socially tutored birds than in passively tutored ones (Chen et al., 2016). Furthermore, the daily light-dark change also affects the imitative accuracy. Song structure deteriorates after a night of sleep and gradually recovers through morning singing. Birds with the greater overall morning deterioration eventually produced the best copies of their tutor songs, which may underlie the balance between consolidation and plasticity of learned songs (Derégnaucourt et al., 2005). Moreover, the encoding of an internal model from the tutor may influence the imitative accuracy. Deshpande et al. (2014) exposed juveniles to the song model on only one day during their entire developmental process. Birds whose songs improved
in complexity of syllable structure during the day of model exposure turned out to be best learners regardless of the amount of singing, suggesting the importance of auditory memory encoding.

The individual ability of imitation may be ultimately governed by the genetic variation in zebra finches. Cross-fostering studies revealed the heritability of some acoustic features in innate calls, and male songs revealed a low heritability in the song traits related with song complexity, such as repertoire size or motif length, while several acoustic features of songs had substantial heritability (Forstmeier et al., 2009). They used the samples of zebra finches reared in aviaries, where there were multiple tutors singing different model songs. In this case, a juvenile bird copies different parts from different models (Eales, 1985, 1987; Mann & Slater, 1994; Slater & Jones, 1995; ten Cate & Slater, 1991; Williams, 1990), so that Forstmeier et al. (2009) could not measure the genetic influence on imitative accuracy to a single model. Therefore, present study aims to study the link between genealogical lineage and imitative accuracy by using the recorded data of songbirds tutored artificially with a single model.

Sato et al. (2016) found that the syllable duration of zebra finch subsong was biased towards their birth parents, and the individual distribution of syllable duration persisted after deafening or fostering by Bengalese finches. It is therefore assumed that the individual’s subsong features may relate to his imitation ability, and this link may be linked to the genetic variation. The goal of this study was to predict the vocal learning propensity of zebra finches from their genealogical tree and subsong. It would be of great help if we can make the prediction in imitative accuracy before tutoring juvenile zebra finches. Selection of better or worse vocal learners before development would help future studies to elucidate the social, neural and molecular substrates of vocal learning, and also its influence on female preference and fitness.
Here we show that a pedigree tree of colonies can be used to detect families with fast and slow learners. Moreover, we found that robust individual differences in subsong acoustic features existed, so that subsong can be characterized and quantified for each bird. Although no acoustic features in subsong correlated with imitation accuracy across birds learning different model songs, some acoustic features correlated with imitation accuracy in a model- or syllable- specific manner.

Method

Data and Animal Care

The investigator did not conduct the experiments, and rather analyzed the birdsong data already collected by other researchers. Those data were recorded in past experiments, which were conducted following the guidelines of the US National Institutes of Health (NIH), and were approved by Institutional Animal Care and Use Committees (IACUC) of Hunter College and City College of the City University of New York.

Bird Sound Recording and Vocal Training

All male zebra finches were kept with parents and siblings until 7 days post hatch (dph), then the father was removed to prevent song exposure during the critical period of auditory inputs. At 31-39 dph, animals were housed in sound-isolating individual boxes where they stayed for two months of their critical period for song development. Birds were given water, seed and egg powder ad libitum and kept on a 12:12 hour photo-period schedule. Their songs were recorded continuously. Some birds went through passive training, namely the exposure to 20-30 playbacks per day to a model song at random probability of 0.01 per second, from 33-40 dph until 43-50 dph. Then keys were introduced to all individual boxes, so that each bird needed to
peck the key to hear 20 song playbacks a day (active training). Each bird received the training with only one model type, composed of 1-4 unique syllable types in motif.

**Pedigree Tree**

The researchers rated song imitation speed of 403 trained males at 50-159 dph by visual inspection. 180 birds were rated as having learned model songs adequately (Mean Decision Date = 67.04 dph ±16.25), while 223 birds were rated as slow-learning (Mean Decision Date = 76.75 dph ±15.81). Their pedigree tree was generated together with 191 parent data, by using kinship2 and FamAgg packages of R software. Each bird was annotated with its song imitation speed and exposed model syllabled numbers in motif.

**Subsong Individual Difference**

The vocal babbling sound before first exposure to model songs was collected to make 2-minute wav files, where cage noise and innate calls were removed manually by using Audacity. In order to test individual difference in subsong acoustic features, subsong data of nine birds were analyzed. One of their subsongs was gathered at five different days, and another of them at five different time points within a day. The distributions of probability density were estimated for acoustic features and syllable durations calculated with Sound Analysis Pro 2011 (SAP 2011; Tchernichovski et al., 2000) by setting the advance window at 2.99ms. In order to quantify the variability within a single bird and across different birds, Kolmogorov-Smirnov (KS) statistic values were collected for each subsong group (“single bird”, “over subsong development”, and “different birds”). KS values represent the difference between two distributions in each subsong group (e.g. how different are AM distribution in subsong of R112 v.s. R373 in “different birds” group), and were calculated between all possible combinations in each subsong group. Two-sided Wilcoxon rank sum test was used to determine if there are differences in KS values
between subsong groups. This difference in KS values is thought to correspond to the difference between within-individual variability (either on a single day or across multiple days) and between-individual variability.

**Correlation between Subsong and Imitative Accuracy**

Acoustic features and syllable durations in subsong data of 66 birds were analyzed. The duration and FM for each bird were summarized by their medians, instead of means which are inappropriate for non-normal distribution. Entropy was summarized by its skewness, since the shape of entropy distribution varies across birds (see Figure 7). AM was represented by its variance, as the central tendency of AM (i.e. mean or median) always be zero (see Figure 8). Fifteen birds were trained with the model “AAAA”, 25 birds with “ABC” or “ACB”, 16 birds with “Samba”, and 10 birds with “Simple” (Sonograms are shown in Figure 8a). Model songs “ABC” and “ACB” had a different syntax but the same sets of syllables, thus they were treated collectively as “ABC” in the present study. It is noted that “AAAA” and “ABC” learners underwent both passive and active training, while “Samba” and “Simple” learners received only active training (see Bird Sound Recording and Vocal Training section of this Method). Vocalized songs were collected for “AAAA”, “ABC” learners at age 54 dph, “Samba” learners on 54 dph ($N = 16$) and on 90 dph ($N = 11$), and “Simple” learners on 80 dph. Then imitative accuracies of those birdsongs were quantified by SAP 2011 (Tchernichovski et al., 2000), using the similarity calculation with mean-value asymmetric comparison to model songs 20 times per each data point. Thereby we have calculated the similarity between vocalized motif and model motif for “AAAA” (at 54 dph), “ABC” (at 54 dph), “Samba” (at 90 dph), and “Simple” learners (at 80 dph), and the similarity between vocalized note and each model note for “Samba” learners at 54 dph. Ninetieth percentile value out of 20 similarity scores were used for analyzing each bird
song. Then Spearman test was applied to see non-parametric correlation between subsong feature similarity at later ages for each bird.

**Results**

**Pedigree Tree**

In order to find the families with better or worse vocal learning ability in the past zebra finch colonies of Tchernichovski lab, we have plotted the pedigree tree in Figure 1. Then the family with potentially great learning ability and the family with differing learning ability possibly depending on a different mother were recorded (Figures 2, 3).
**Figure 1.** Pedigree tree of zebra finches at the Tchernichovski lab.

a, Several excerpts of families from whole pedigree tree annotated with vocal learning speed. Only males with known vocal learning speed or their parents were plotted. Each rectangle or ellipse denotes one bird, with letters below each symbol is a bird ID. The solid lines denote the kin relationship, which puts the parents on top, while those connected with parents are their children. The broken line connects the same birds, which was plotted twice for graphical clarity of the tree.
b, Legend for the Pedigree tree. Rectangle denotes male, and ellipse denotes female. Empty rectangle denotes the fast learner, meaning that the bird who was evaluated to have a fidelitous imitation of model song by subjective visual inspection of the investigator. Filled rectangle denotes the slow learner, who did not exhibit the imitation by the age of subjective decision. The colors of rectangle show the number of unique syllable types in the model motif. The question mark inside rectangle means unknown vocal learning speed.
Figure 2. Examples of fast learners and families. a, The excerpt of pedigree tree showing two small families who produced children learning complex model songs quickly. b, The sonograms of 4-syllable fast learners in one of those families. Left is the sonogram of 4-syllable model song, and right are the sonograms of birds which were exposed to the model song, at the age of subjective evaluation of imitative accuracy. Horizontal direction in each sonogram denotes time course, with left preceding right, while vertical direction in each sonogram shows frequency, with top denoting higher frequency components.
Figure 3. The family which produced offspring with different learning speed possibly associated with the mother’s. a, Excerpt of pedigree tree. Note that male “o97” mated with three females, “R4386”, “R5067”, and “R4306”. When mated with female “R4386”, male “o97” produced fast-learning children, while producing slow learners with female “R4306”. b, Sonograms of model songs and 1-syllable slow learners in the tree above at the age of subjective evaluation.
Subsong Individual Difference

Are there individual differences in subsong vocalization patterns, and how can we quantify them? In order to answer these questions, we collected sets of subsong data within single individuals, and across multiple individuals. Stability within individuals and variability between multiple individuals are the prerequisite for quantifying the characteristic of each bird subsong, and would also provide the rationale to test individual subsong for its correlation with later vocal learning outcome.

Three sets of subsong sonograms were selected (Figure 4): (1) five subsong sounds collected from a single day (41 dph) and a single individual (R986), (2) five subsong sounds collected from multiple days (38-42 dph) and a single individual (R658), (3) seven subsong sounds collected from multiple individuals and varying days. Visual inspection indicates that subsongs from a single day and a single individual are similar (left column of Figure 4). Subsongs from multiple days and a single individual (center column of Figure 4) exhibited changes on 40 dph, showing incorporation of loud notes into subsong bouts. Yet other relatively soft notes were similar to each other, and loud notes were also similar to each other. Subsongs from multiple individuals exhibit variability (right column of Figure 5). Most evident frequency components in R112 and R155 were around 5 kHz, while those in R373 were around 2 kHz.
**Figure 4.** Subsong sonograms exhibit within-individual stability and between-individual variability. **Left column:** set of sonograms within a single bird (R986) on a single day (41 dph) at different time of day (labels to the left of each sonogram; 10:30, 13:11, 17:24, 19:08, and 20:19). Sonograms are similar to each other. **Center column:** set of sonograms within a single bird (R658) but on different days (labeled next to each sonogram). This bird incorporated loud notes into subsong rendition at 40 dph. **Right column:** set of sonograms across multiple birds. Birds’ IDs are labeled to the left of each sonogram. Set of sonograms seem more distinct from each other in multiple birds than that in a single bird.
To quantify and visualize within-individual stability and between-individual variability, acoustic features for each 2.99 ms of advanced window for each subsong sound file were extracted and their distributions plotted. This study focuses on four acoustic features, Frequency Modulation (FM), syllable duration (Duration), Wiener entropy, and AM (Figures 5-8), and other acoustic features are plotted in Supplementary Figures 1-3.

Frequency Modulation (FM) showed stability within a single bird (Figures 5 b,c). On the other hand, distributions across birds were variable. Shapes of violin plots show that birds PK245 and PK249 had their peaks in the right side of medians, while R112 and R373 had ones in the left side. Syllable duration (Duration) exhibited stability within a single day and a single bird (Figure 6b). Distributions of the duration changed within a single bird across multiple days (Figure 6c), yet median duration remained stable until 41dph. On the contrary, distributions of duration were quite variable across birds, both in shapes of distribution and their central tendencies (Figure 6d). Distributions of Entropy seem quite similar within a single bird and over subsong development, (Figure 7b, c), while variable across birds (Figure 7d). Violin plots to show subsong AM did not clearly show if there are within-individual stability or between-individual variability (Figure 8b-d). To quantify within-individual stability and between-individual variability, the investigator calculated the distance between the possible combinations of feature distributions in each subsong group, by using Kolmogorov-Smirnov (KS) statistics values (Supplementary Figure 4). Two-sided Wilcoxon rank sum test of KS values between “Birds” and “Days” group in panel “AM”, “Duration”, “FM”, and “Wiener entropy” (respectively $U = 191, z = 3.63, p < 0.0001; U = 148, z = 1.81, p = 0.011; U = 166, z = 2.58, p = 0.0020; U = 190, z = 3.59, p < 0.0001$), suggesting the stability within a single individual and variability between multiple individuals.
From these results, it is plausible to speculate that subsong exhibit robust individual differences, and that these differences can be captured by using the acoustic features AM, syllable duration, FM, and Wiener entropy.
Figure 5. Distributions of frequency modulation (FM) are stable within a bird but variable among birds. **a**, FM calculated at each 2.99 ms time window was overlaid as light blue lines on the exemplar sonogram of subsong. The arrow point is an example of low FM, located where the shapes of frequency trace in sonogram are flat. **b-d**, Violin plots are the smoothed probability densities of subsong FM values (x-axis). Thick parts of the plots reflect the more prevalent values of acoustic features in each subsong. Overlaid on violin plots are box plots showing minimum, maximum and 1st and 3rd quartile values for each distribution. **b**, FM distributions in subsongs within a same bird (R986) and same dph (41dph). The labels adjacent to each violin plot show the time points within a day. **c**, FM Distributions in subsongs within a bird, but across
dphs (38-42 dphs). The labels next to each violin plot exhibit the recorded day of subsongs. FM Distributions in subsongs across seven birds. The labels on violin plots are bird IDs.
Figure 6. Distributions of syllable duration are stable within a bird but variable between birds. 

Durations of syllables in subsong were plotted as red marks at the bottom of exemplar sonograms, b-d. Violin plots show the smoothed probability densities of subsong duration values. b, Duration distributions in subsong within a same bird (R986) and same dph (41 dph). c, Duration distributions in subsong within a bird across dphs (38-42 dphs). Note that while the shape of distribution somewhat changed, its median remained stable until 41 dph as compared to data in Figure 6 d. d, Duration distributions in subsong across seven birds. The distributions and central tendencies of plots are variable.
**Figure 7.** Distributions of Wiener entropy are stable within a bird but variable between birds. 

Wiener entropy calculated at each 2.99 ms time window was overlaid as yellow lines on the exemplar subsong sonogram. **b-d,** Violin plots show the smoothed probability density of subsong Wiener entropy values. **b,** Wiener entropy distributions in subsong within a same bird (R986) and same dph (41 dph). **c,** Wiener entropy distributions in subsong within a bird (R658), but throughout dphs (38-42 dphs). **d,** Wiener entropy distributions in subsong across seven birds. Wiener entropy distributions vary across birds.
**VOCAL LEARNING PROPENSITY**

**Figure 8.** Distributions of Amplitude Modulation (AM).  

- **a**, AM calculated at each 2.99 ms time window was overlaid as orange lines on the exemplar subsong sonogram.  
- **b-d**, Violin plots show the smoothed probability density of subsong AM values.  
  - **b**, AM distributions in subsong within a same bird (R986) and same dph (41 dph).  
  - **c**, AM distributions in subsong within a bird (R658), but throughout dphs (38-42 dphs).  
  - **d**, AM distributions in subsong across seven birds.

**Correlations between Subsong and Imitative Accuracy**

Given that AM, FM, duration and Wiener Entropy indicated robust individual differences, we tested the Spearman correlation between these subsong features and later imitative accuracy. Correlation analyses were performed separately on different bird groups that were exposed to different song models (Figure 9a, and 10). Note that two model songs were treated as “ABC”.

They differed in syntax, yet contained same sets of syllables. The similarities at later ages between groups exposed to different model were compared by Wilcoxon rank sum test (Figure 9b). There was a significant difference between similarity for “ABC” and for “Samba” ($U = 227.5, z = 3.09, p = .0019$), and between similarity for “AAAA” and for “Samba” ($U = 121.5, z = 2.02, p = .0429$).

Figure 10 shows the results of correlation analysis between subsong features and later imitative similarity to model motifs. Among birds who learned “AAAA”, skewness of Wiener Entropy in subsong correlated negatively with similarity to model at later stages of vocal development with a non-significant trend (Spearman $\rho = -0.49, p = 0.064$). On the other hand, a non-significant trend was observed in that skewness of Wiener entropy did not correlate with later similarity in “ABC”, “Samba” and “Simple” learners (respectively, $\rho = -0.22, p = 0.346; \rho = -0.09, p = 0.797; \rho = -0.20, p = 0.584$). Median of syllable duration correlated negatively with similarity at later ages of juveniles among “Samba” learners ($\rho = -0.67, p = 0.023$), while this relation was not evident among “AAAA”, “ABC” or “Simple” learners (respectively, $\rho = 0.12, p = 0.678; \rho = 0.36, p = 0.114; \rho = 0.06, p = 0.877$). Variance of AM correlated positively with similarity to model among “AAAA” and “Samba” learners (respectively, $\rho = 0.54, p = 0.039; \rho = 0.68, p = 0.025$), but negatively among “ABC” and “Simple” learners with a non-significant trend (respectively, $\rho = -0.43, p = 0.059; \rho = -0.58, p = 0.088$).

In order to examine the imitative accuracy at a finer level, we sampled the syllable-level similarity, meaning the similarity between each syllable in model songs and birds’ syllables. Subsong acoustic features were tested for their correlations with later syllable-level similarity among “Samba” learners (Figure 11). Subsong FM median had the non-significant trend in negative correlation with A-syllable similarity ($\rho = -0.45, p = 0.083$), while showing a positive
correlation with C-syllable and D-syllable similarity (respectively, $\rho = 0.50$, $p = 0.047$; $\rho = 0.56$, $p = 0.023$).
Figure 9. Different song models broadcast to juvenile birds. a, sonograms for models (“AAAA”, “ABC”, “Samba”, “Simple”). b, %similarity of juvenile songs to model at the age of measurement in different model learners (54 dph for “AAAA” learners 54 dph for “ABC” learners, 90 dph for “Samba” learners, and 80 dph for “Simple” learners). Wilcoxon signed rank tests found a significant difference between similarity for “ABC” and for “Samba”, and between similarity for “AAAA” and for “Samba”. “Samba” seems to be the most difficult song to imitate.
**Figure 10.** Relation between subsong features and later imitative similarity to model songs.

Panel title shows the model songs broadcast to birds. X-axis denote acoustic features of individual subsong. Each dot denotes one bird. % similarity of Y-axis is 90th percentile of the similarity values comparing the model songs and bird songs. Blue lines illustrate best-fit linear regressions.
Figure 11. Relation between subsong features and later imitative similarity to model syllables.  

**a,** The sonogram of model song “Samba” with each syllable labeled from A to D. Red line at the bottom of sonogram represents each syllable duration.  

**b,** Scatter plots of subsong FM median and the 90th percentile similarity to each syllable of Samba at 54 dph. Panel titles show the name of syllables (“A” to “D”). X-axis denotes the median of FM for individual subsong, and Y-axis denotes later imitative similarity to model song. Of particular note is that A-syllable in the model has high FM and C-syllable has low FM, though the subsong with low FM led to better A-syllable and poor C-syllable.
Discussion

The principal goal of this study was to predict the vocal learning propensity in zebra finches by investigating the pedigrees of past colonies from the Tchernichovski lab, and also by testing the influence of subsong acoustic features on later imitative accuracy. Pedigree trees of zebra finches made in this study indicate the familial aggregation of fast and slow vocal learners. Present results also provide evidence that subsong is individualized and that this individuality can be quantified using the acoustic features calculated by SAP2011. Sato et al. (2016) reported subsong individual differences only in a temporal measure (e.g., duration), while the present study found individual difference in acoustic structure (i.e., entropy and FM). As for relation between subsong and learning outcome, our results did not find the consistent evidence that the subsong features predict the later imitative accuracy of each bird, thus we do not conclude that the vocal propensity can be predicted from subsong before the vocal development. While correlations between subsong acoustic features and later imitative accuracy were found, this effect was dependent on the tutor song that juveniles heard. Additionally, subsong acoustic features also differently correlated with different types of syllables.

How can we interpret this model- and syllable-specific subsong effect on later learned song? One possible hypothesis is that the subsong acoustic features influence the tutor choice in natural situation where multiple tutors are available. Some studies have suggested that males selectively learn song components from one tutor if given a choice of two adults (Böhner, 1983; Clayton, 1987; Ward et al., 1998), while others suggested that even if a primary tutor is identifiable, juveniles also copy parts of songs from other males (Eales, 1985, 1987; Mann & Slater, 1994; Slater & Jones, 1995; ten Cate & Slater, 1991; Williams, 1990). It is suggested that exposure to multiple song models may enable a juvenile to construct a song that is unique and
thereby serves as individual signature (Williams, 1990). Results of the present study implicate that subsong features, such as FM or duration, might facilitate or inhibit the imitation of one type of model over other types. Furthermore, subsong features may lead birds to selectively copy one element out of a model song over other elements. This selective copying cannot be explained by the proximity in acoustic distances between the subsong feature and model, as refuted by the FM’s differential effect on imitative accuracy of different syllable types (Figure 11b). Since the low FM subsong led to a better imitation of a high FM syllable, the acoustic distance between subsong and model may not be the factor regulating either better imitation or tutor choice.

In our pedigree tree, we did not statistically test whether the familial aggregation was caused just by chance or not. Thus, we cannot conclude the vocal learning propensity is biased due to family. However, researchers can still pick up the families with many of either slow or fast learners, in order to examine influence of family on vocal learning in finer level. In the future, the analysis of subsong across multiple families may reveal the individual difference in subsong may be smaller within a family than that between families. After then, we may further move on to decipher the effects of genetics and fostering environment on vocal learning outcome. It is known that a part of acoustic features we tested (FM, AM, syllable duration) have genetic heritability in adult song (Forsteier et al., 2009), so those features may have the genetic heritability at subsong stage, too.

As for relations between vocal learning and brain circuits, HVC volume and its number of neurons in adults correlate with imitative accuracy, but not with song complexity (Ward et al., 1998). However, not HVC but AFP is required for subsong vocalization (Ölveczky et al., 2005; Aronov et al., 2008). HVC lesion did not significantly affect subsong acoustic features, including syllable duration, Wiener entropy, goodness of pitch, amplitude, and FM (Aronov et al., 2008).
Therefore, the correlations between subsong acoustic features found in the present study were not very likely to derive from individual differences in HVC morphology and function. Additionally, lesions of the dorsolateral division of the medial thalamus (DLM) in AFP decreased the variance of syllable duration and inter-syllable gap duration, and increased the rhythmicity in subsong (Goldberg & Fee, 2011). Those features in subsong were not analyzed in this study, which provides future directions of this kind of investigation.

Results of this study are confounded by a number of methodological limitations. Samples of the present subsong study were obtained from past data sets collected in the Tchernichovski laboratory. Birds went through different methods of tutoring; “ABC” learners experienced both passive and active training (see Method; Tchernichovski et al., 2001), while “Samba” and “Simple” learners received only active training. The present sample was also biased because data from birds that had begun tutoring before sufficient amount of subsong was recorded were excluded. Moreover, birds vary in the amount of singing, especially between 30-40 dphs, and on occasion don’t produce the subsong at all during this period (personal observation). This individual variability in subsong amount would be a potential confound in trying to predict vocal propensity based on subsong acoustic features.

Further, recording in isolate chambers may not have captured the optimal singing performance at the age when similarity measurements were taken. When a male sings to a female (directed song), variability in syllable acoustic structure and in syllable sequencing declined compared to when males sing alone (undirected song; Kao & Brainard, 2006, Kao et al., 2005, Sakata et al., 2008, Sossinka & Böhner, 1980). This is particularly remarkable in juveniles whose immature and variable undirected songs are transformed into the adult-like songs by the presence of a female (Kojima & Doupe, 2011). Undirect songs involve vocal exploration by AFP, which is
said to be the ‘practice’ state while direct songs are the ‘performance’ state (Hessler & Doupe, 1999 a,b; Kao et al., 2008; Brainard & Doupe, 2013). In our study, the recordings of undirected songs from isolate chamber were used for analysis. Thus, we may not have quantified the juveniles’ optimal performance, although we tried to estimate it by using 90th percentile value of %similarity. In addition, birds that learned models inaccurately by 100 dph from an isolated-housing tutoring method were more likely to mutate their songs after moving to the aviary (Derégnaucourt et al., 2013). Thus, it is not conclusive that birds that exhibited low %similarity scores in this study (54 dph, 80 dph and 90 dph) were poor singers in their late life.
Supplementary Figure 1. The probability densities in all acoustic features that can be calculated by SAP2011 on a single bird. Data were all from subsongs of bird “R986” across different time points on 41 dph. Violin plots showing distributions of acoustic features extracted from subsongs of different time points.
Supplementary Figure 2. The probability densities in all acoustic features that can be calculated by SAP2011 over subsong development. Data were all from subsongs of bird “R658” across different day-post-hatch (dph). Violin plots showing distributions of acoustic features extracted from subsongs of different days.
Supplementary Figure 3. The probability densities in all acoustic features that can be calculated by SAP2011, on different birds. Violin plots showing distributions of acoustic features extracted from subsongs of different birds. Each panel denote one acoustic feature.
Supplementary Figure 4. Statistical evidence of robust individual difference in subsong acoustic features. Y-axis corresponds to Kolmogorov-Smirnov (KS) statistic values in each group of subsongs. KS values represent the difference between two distributions in each group of subsong (e.g. how different are AM distribution in subsong of R112 v.s. R373 in “Birds” group). Dots on boxplots denote KS values from all possible combinations in each group of distributions in subsong acoustic features. “Birds” means the group of subsongs across birds, “Days” means the group of subsongs across different days within a bird (“R658”), and “Times” means the group of subsongs across different time points on 41dph within a bird (“R986”). P values on boxes of “Days” or “Times” were emulated by two-sided Wilcoxon rank sum test to compare KS values of each group with those of the “Birds” group.
Glossary

The Brain Song System of Zebra Finch

Birdsong is produced by the dedicated and accessible brain nuclei called song system (Nottebohm et al., 1976; Supplementary Figure 5). The song system in zebra finch brain was shown in Supplementary Figure 5. It consists mainly of the vocal motor pathway (VMP) and the anterior forebrain pathway (AFP). In VMP, HVC generates the temporal and sequential pattern of song (Hahnloser et al., 2002, Long & Fee, 2008, Yu & Margoliash, 1996), and projects to RA, which controls downstream motor neurons. The AFP has a loop structure that indirectly connects HVC and RA (Doupe et al., 2005, Gale & Perkel, 2010, Bottjer & Johnson, 1997), and gives rise to vocal learning. Neurons in HVC project to Area X, and Area X pallidal neurons send inhibitory projections to the DLM. Neurons in DLM project to LMAN, and LMAN projects both to RA and back to Area X. The output of AFP flows primarily through LMAN to RA, thus at this single bottleneck the output of AFP can be manipulated to test AFP’s effect on vocal learning (Brainard & Doupe, 2013).

Supplementary Figure 5. The simplified diagram of the song system in zebra finch. The vocal motor pathway (VMP; red lines) includes nuclei higher vocal center (HVC) and the robust
nucleus of the arcopallium (RA), which send signals to the downstream brain region (not shown) controlling patterned respiration and the syringeal vocal musculature. The anterior forebrain pathway (AFP; black lines) includes a part of basal ganglia (Area X), a thalamic nucleus (DLM, standing for dorsolateral division of the medial thalamus), and a frontal cortical nucleus (LMAN, standing for lateral magnocellular nucleus of the anterior neo striatum). P is the abbreviation for the posterior direction, and D is for the dorsal direction. Adapted from “Translating birdsong: songbirds as a model for basic and applied medical research,” by M. S. Brainard and A. J. Doupe, 2013, Annual review of Neuroscience, 36, p. 489-517. Copyright 2013 by the Annual Reviews.

**Acoustic Features**

Present study focused on four acoustic features; syllable duration, Wiener entropy, frequency modulation (FM), and amplitude modulation (AM). Syllable duration is how long a single syllable lasted (typically 0-300ms). Wiener entropy is a measure of randomness, and range from 0 to 1 (Tchernichovski et al., 2000). The Wiener entropy of a multi-harmonic sound depends on the distribution of the power spectrum: if narrow (the extreme is a pure tone), the Wiener entropy approaches zero; if broad (the extreme is a white noise), the Wiener entropy approaches one. FM is estimated based on time and frequency derivatives across frequencies. If the frequency derivatives are higher than the time derivatives, FM is low (Tchernichovski et al., 2000; Tchernichovski & Mitra, 2012). AM is the overall time-derivative power across all frequencies. AM captures changes in the amplitude of sounds. It is positive at the beginning and negative in the end of each syllable (Tchernichovski & Mitra, 2012).
Violin Plot

A violin plot is a visualization of numeric data (Hintze & Nelson, 1998). In present study, it is a box plot with the rotated kernel density plot of the data at different values. Inside box plots denote the max and minimum (when excluding outlier values), and quartile values (first and third quartile, and median) of the distribution. Outer density plots denote the all values, with thickness showing how common each value is (Hintze & Nelson, 1998). The thickest section is the mode of a distribution. The violin plot and the histogram are similar, and both show the distribution of a numeric vector. Their correspondence are plotted in Supplementary Figure 6.

Supplementary Figure 6. The exemplar histogram and violin plot to visualize subsong acoustic feature. The histogram above shows the count of different values at Y-axis, while the violin plot below shows the smoothed probability density by its vertical thickness. In both cases, X-axis show the values of the range of distribution. Here the amplitude in X-axis is used as an exemplar subsong feature. A boxplot is also plotted inside the violin plot, indicating minimum and max values, and quartile values.
References


