Reinforcement induced Reduction in the Number of Introductory Notes in Zebra Finches

End-semester Project Report

Name: Vasudha Kulkarni
Register number: 20191057
Date: 20th April 2023
Supervisor: Dr Raghav Rajan, Department of Biology

Introduction

The Zebra finch song is an excellent model to study the mechanism behind learned, complex motor sequences. Usually, these songs are preceded by repeated short vocalisations called introductory notes or INs (Price 1979). Such gestures have been observed in other animals as well, and multiple hypotheses try to explain these introductory gestures. Introductory gestures could enhance the communicative aspect of the display and increase its detectability (Wiley and Richard 1982); they could serve as recognition signals to the main display by providing singals for species identity (Soha and Marler 2000), or they could be a motor preparatory function for when the brain 'prepares' parameters to generate the complex movement (Rajan and Doupe 2013).

A good example of motor preparation is when an athlete dribbles the ball before she throws it into the basket, which improves the accuracy of the throw. The motor preparation hypothesis posits that it helps the brain arrive at a common spatiotemporal pattern of neural activity before the core action is carried out. INs in zebra finches are vocalisations of short duration and simple spectral composition. The motor preparatory hypothesis is supported by the observations that initial, more variable INs converge acoustically to a stereotypical last IN, which signals the readiness of song sequence generation. IN-related HVC activity in the brain progresses to a common endpoint just before the song (Rajan and Doupe 2013).

Previous studies in the lab have shown that Zebra finch INs are feedback-independent (Rao, Kojima, and Rajan 2019), and their features are shaped by both learning and biological predispositions (Kalra et al. 2021). It was also observed that INs have high beginning bout probability and self-transition probability in Zebra finches as compared to calls or song motif syllables. In this context, through this semester project, we wanted to test whether it was possible for an adult zebra finch to modulate the number of mean INs by punishing it with a burst of white noise (negative reinforcement) when it sang too many INs.

In previous studies, white noise has been used to modulate the pitch of a particular syllable in the zebra finch song and study how the song structure is maintained (Tumer and Brainard 2007; Canopoli, Herbst, and Hahnloser 2014). Similarly, in Bengalese finches, use of closely-timed bursts of white noise has been used to modulate transition probabilities between certain syllables, and show that Bengalese finches have learned contextual control over syllable sequencing (Veit et al. 2021).But these studies use computationally expensive and proprietary software and hardware, like LabView, in order to identify specific syllables and punish them.

Through this project, we wanted to test whether an adult zebra finch's mean number of INs can be reduced through negative reinforcement (white noise). First, INs had to be detected using simple measures like variance of audio data and syllable duration, and white noise had to be played after the 4th IN. The project and experiment were conducted on one adult male zebra finch – ylw97ylw29. But the analysis pipeline constructed can be extended to more birds.

Methods

Recording Songs

The bird was recorded in the lab as follows – the bird was isolated for a minimum of 24 hours within a soundproof recording box and a microphone was brought into the recording box and attached to the roof of the bird's cage to collect audio. The mike was connected to a mixer that was only used as an amplifier. On the PC, we use Alsamixer, a Linux-based graphical program used to configure the sound settings and adjust the volume settings. After setting up Alsamixer, we used the PyCBS, a python-based program that saves audio data and plots the RMS of input audio files. It allows us to choose different sound devices and can record in mono or stereo. Data was digitized at 44100Hz sampling rate. The Audio files were saved and exported in 'wav' format in the lab PC. For the experiment, a pair of speakers were connected to the PC and set up in the same soundbox to play the white noise file.

Data Analysis

The Onset Offset files were processed to first insert 'Start' and 'End' syllables before the beginning and at the end of each song bout. Song bout is defined as a series of vocalizations separated by 2 seconds of silence (Sossinka and Böhner 1980). Then, rare syllables (<2% occurrence) were deleted and the song bouts with less than 3 unique syllable types were excluded, because we wanted to focus on bouts with song motifs, and not those which only contained calls or other vocalizations.

These modified files were then used to calculate the mean number of INs, transition probability matrices, quantify different aspects of syllable properties and acoustic features (syllable duration, mean frequency, amplitude, onset time etc.), create plots and do statistical tests on the data. The Onset-Offset files were also used to calculate the duration between 4th IN and 'a' (first motif syllable), and to align the variance of syllables from recorded .wav files.

The following measures were used to quantify and estimate the song of the bird – Sequence Linearity is a measure of the ordering of notes in a song, Sequence Consistency tells us how often a particular path (or motif sequence) is followed in the songs (Scharff and Nottebohm 1991). These measures are calculated as –

$$Sequence\ Linearity = \frac{No.\ of\ unique\ syllables}{No.\ of\ unique\ transitions}$$

$$Sequence\ Consistency = \frac{Sum\ of\ typical\ transition\ probabilities}{Sum\ of\ all\ transition\ probabilities}$$

Variance of audio data

Incoming audio data can be converted into an array for the computer to process by sampling the audio signal 44,100 times a second. This gives us a long array for the incoming audio, or recorded files, which can be used to calculate variance for certain frame length or window size of the data. I chose frame length of 20 ms (i.e., array of 882 values) to calculate the variance and detect INs from other syllables. 20 ms frame length was the shortest duration that didn't cause an overflow error in the program i.e., the data processing takes longer than the speed with which new data is incoming.

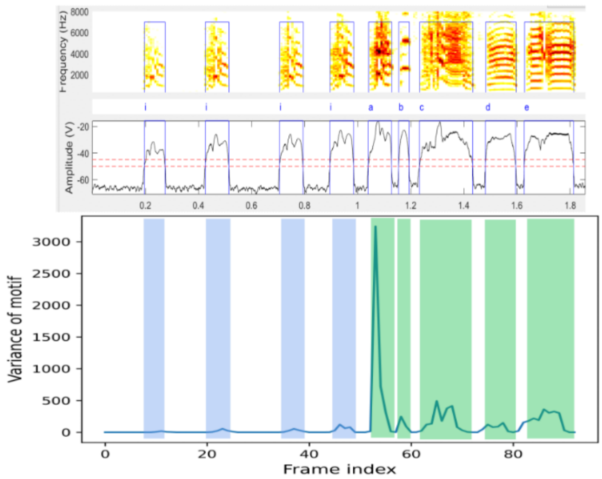


Figure 1. Spectrogram and variance plot of a recorded song file of length 2 s

Modifying PyCBS

PyCBS is a python-based program developed by Dr Raghav Rajan to record audio data and save it as .wav files. This program was modified to detect INs and punish them along with recording the sound from the microphone. The code was modified to adjust the frame size (or buffer size) and some new code was added such that variance of the incoming data could be used to detect INs based on a number of IF conditions, which are illustrated in the diagram below.

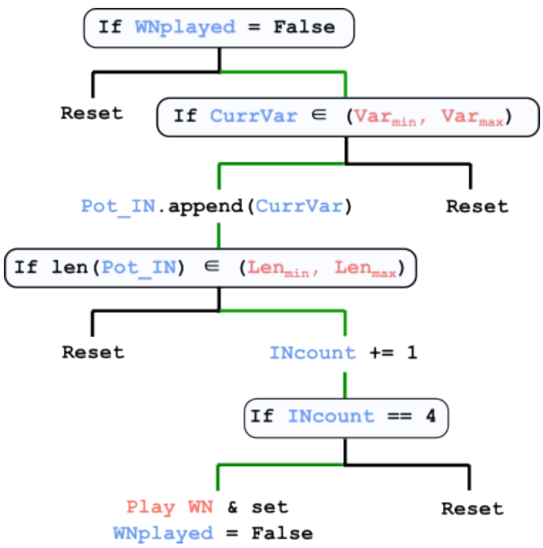


Figure 2. Illustration/flowchart of the code added to PyCBS to detect INs

Experimental protocol

The experiment with the bird was set up for 5 days (27th March 2023 to 31st March 2023), for an average of 10 hours per day (slightly varying number of hours based on technical glitches) between 7 am to 8 pm, when the program would detect INs and play white noise after the 4th IN. The hit rate was calculated as the number of times white noise was played divided by the total number of times INs were detected. The bird was recorded on the morning of 27th April before the experiment and on 1st April, after the experiment to compare the pre- and post-experiment differences in number of INs.

Estimating the performance of the detector

The performance of the detector can be calculated using measures that combine the false positive and false negative classifications by the detector.

	Label IN	Label motif
Detected IN	True Positive	False Positive
Detected motif	False Negative	True Negative

$$Specificity = \frac{TP}{TP + FN} \qquad Sensitivity = \frac{TN}{TN + FP}$$

$$FNR = 1 - Specificity \qquad FPR = 1 - Sensitivity$$

These values, along with accuracy of the classification can tell us about the performance of the detector with recorded audio files and with the bird's song.

Results

Characterising the bird

The subject of the experiment, ylw95ylw29, is about 3.5 years old, and it's song features are listed in the table below. Moreover, the bird sang very frequently throughout the day.

Feature	**Value**
No. of unique syllables	10 (7 motif, 3 call)
Mean Song duration	8.06 s
Mean no. of syllables per song bout	45.87
Sequence Linearity	0.64
Sequence Consistency	0.83
Mean no. of INs	**5.14** (now 4.5)
Mean syllable duration of INs	75.2 ms
Mean Frequency of INs	2273.04 Hz
Mean Log Amplitude of INs	37 dB

Table 1. Song features and IN properties of the bird ylw95ylw29 from 50 song bouts (30/12/22)

Variance of INs

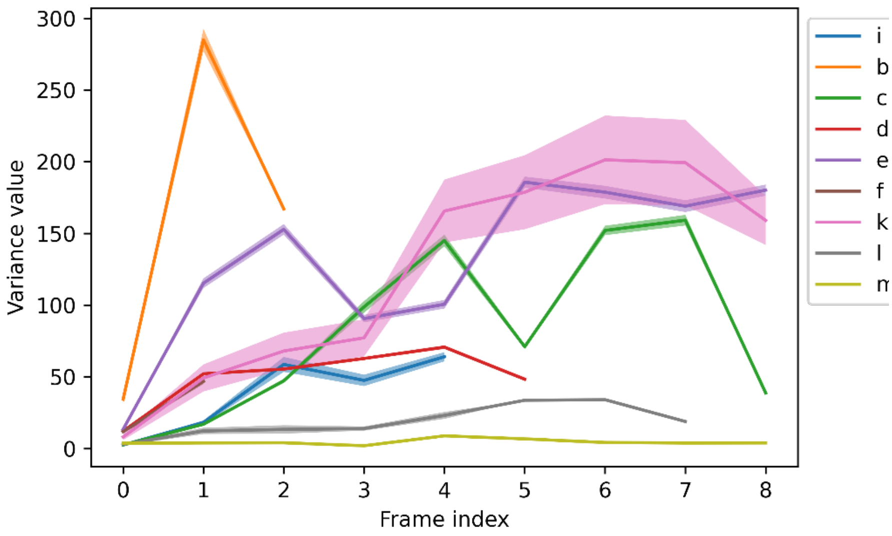


Figure 3. Variance pattern of all syllables (except 'a')

The minimum and maximum variance threshold for INs was decided based on the first 4 INs, since that is what we are interested in detecting. The 5th percentile value of the second frame (lesser variance) and the 95th percentile value of third frame (higher variance) was used to set the threshold. So, minimum variance = 2.2 and maximum variance = 57.

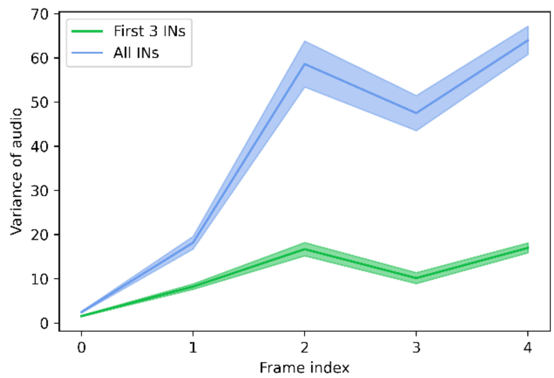
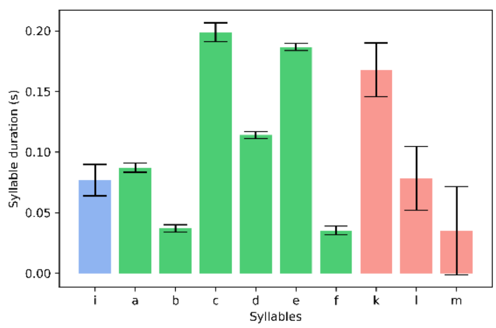
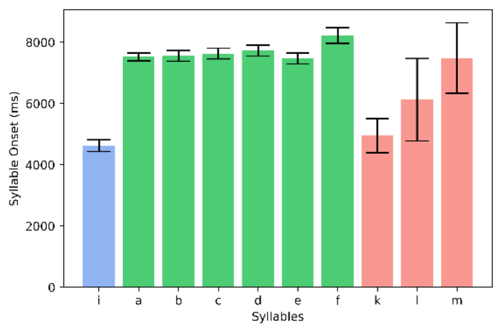


Figure 4. Variance pattern of all INs versus First 3 INs of the song bout. Lines indicate mean and shaded regions are the standard error.

Acoustic features of all syllables





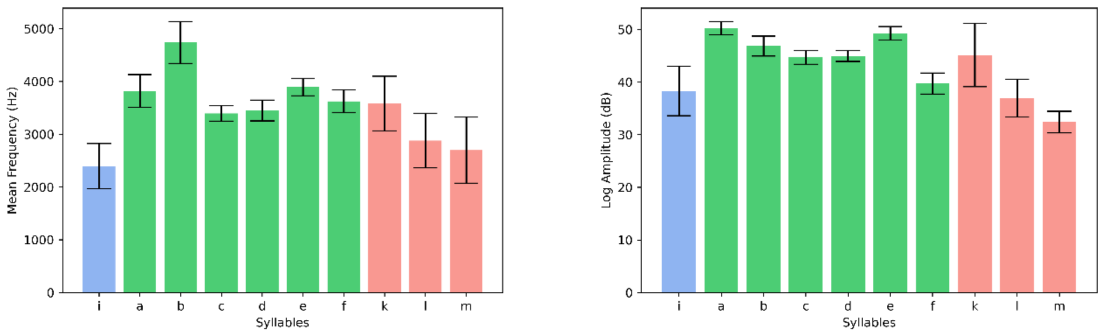


Figure 5. a) Syllable duration, b) Onset time, c) Mean frequency and d) Log amplitude of all the syllables

Performance of the detector with recorded song bouts

The detector (based on variance thresholds, syllable duration and onset time) was tested on 50 recorded song bouts and the number of False positive and False negative frames were calculated.

False positive frames – 17.4 (~4 syllables)
False negative frames – 6.3 (~1-2 syllables)

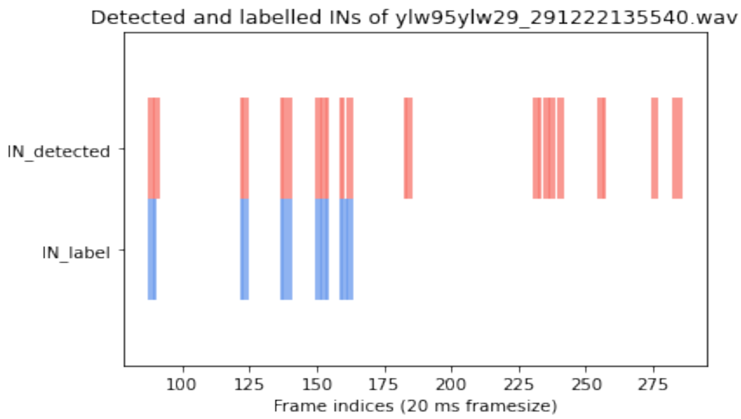


Figure 6. Example raster plot of detected INs versus Label INs for one file out of 50

Results of the Experiment with ylw95ylw29

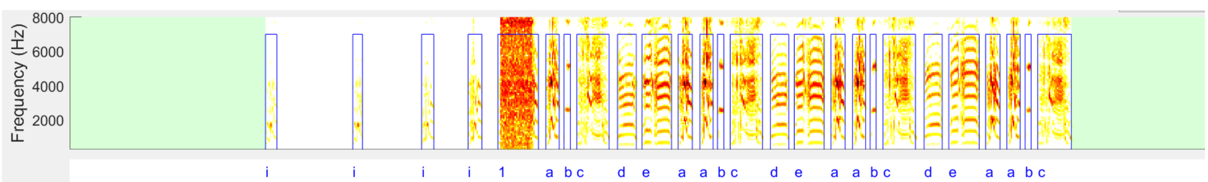


Figure 7. Spectrogram of the bird's song and the white noise played after 4 INs

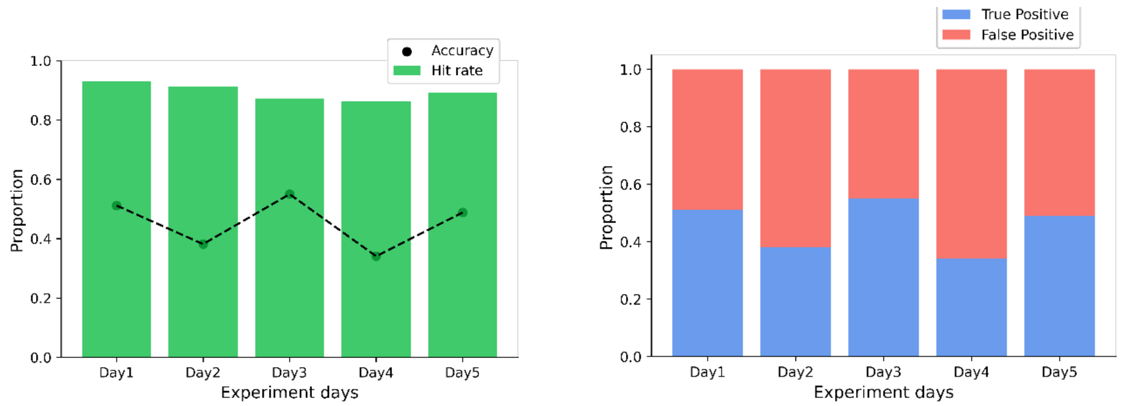


Figure 8. a) Hit rate and accuracy, b) False Positive and True Positive proportions of the detector over 5 days of experiment

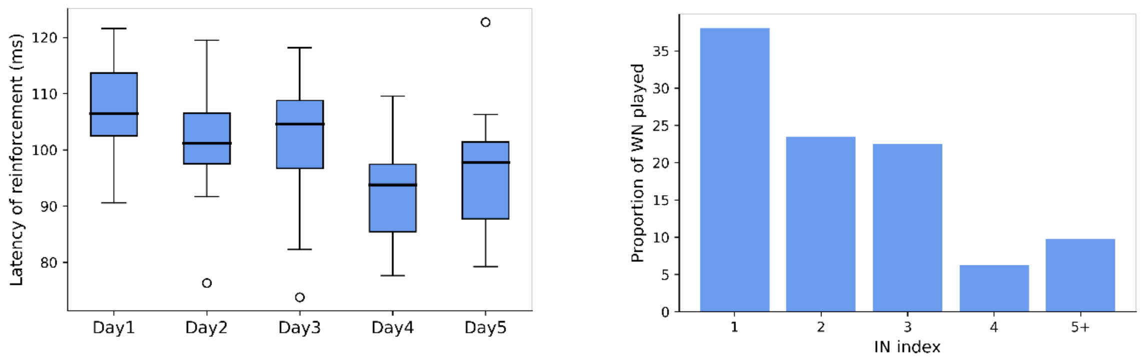


Figure 9. a) Latency of white noise from the nearest preceding IN and b) Distribution of the index of INs after which white noise was played

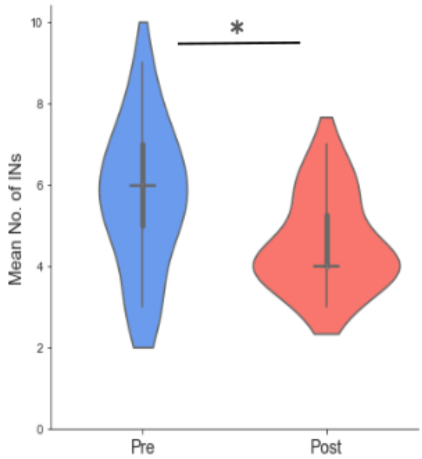


Figure 10. Mean and distribution of the number of INs before and after the experiment (N=16). MannWhitneyU test, statistic = 192.5, p-value = 0.019

Discussion

Introductory notes are an important component of song production and play a strong role in motor preparation. We wanted to test whether it was possible for an adult zebra finch to reduce the number of mean INs by punishing it with a burst of white noise, which acts as negative reinforcement.

First, we established that variance is a good and efficient measure which can be used along with other features, such as syllable duration and onset time, to detect INs and provide real-time negative reinforcement. Although, the current algorithm detects a lot of false positives, there was a significant decrease in mean number of INs produced after five days of experimental protocol in one zebra finch.

According to the pre- and post-experiment data, despite high error of the detector, it seems that the bird can associate white noise with the number of INs and learn to reduce the mean number of INs. But, this result should be viewed cautiously because the mean number of INs can vary over time — it was 5.1 on 29/12/22 and 4.6 on 10/03/23. But given that the distribution of INs has shifted post-experiment, it is a promising path, and should be expanded further with a refined algorithm and increased sample size to make the result more robust.

It is also possible that as a consequence of decreasing the mean number of INs, birds would begin a higher proportion of their songs with one or more calls. It has been shown before that songs that begin with calls have a lower mean number of INs. This was also true for ylw95ylw29, for 50 songs that were analysed, the bird began about 14% of it's songs with a mean of 3 calls, and the mean number of INs was significantly less when the song bout began with a call (Figure 11).

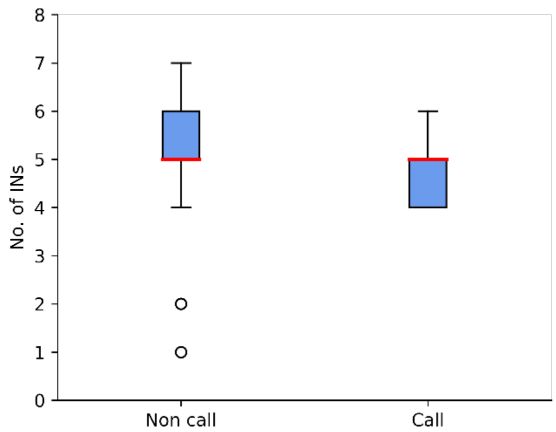


Figure 11. Distribution and median number of INs with song bouts that begin with a call versus bouts that begin with INs (N=30). MannWhitneyU test, p-value = 0.02

Future directions

The first thing which needs to be refined is the IN detection algorithm — it would be useful to consider other factors to detect INs, such as mean frequency or power spectrum. Calculating these features would require the data to be transformed using a Fast Fourier Transfrom (fft) in real time which might affect the latency of the playback. But this is an important step to reduce the number of false positive detections.

Further, we need to expand the sample size and try to extract other patterns of common behaviour and understand the neural mechanisms of how they learn to decrease the number of introductory notes. Along with the experiments, it's also important to quantify the recovery to their template song once the reinforcement has been stopped. Several studies earlier have tried to understand vocal learning by lesioning certain regions of the brain and studying how this recovery is affected.

Another interesting addition to the study would be to use the protocol to reinforce birds to increase their mean number of INs, rather than decrease them, i.e., play white noise if the bird sings less than 4 INs. The motor preparation hypothesis for INs would predict that it should be easier for the birds to increase INs rather than decrease them. It would be interesting to test this using the same birds with sufficient gap between the two tests.

Acknowledgements

I would like to thank Dr Raghav Rajan for his support and guidance throughout the project. I'm grateful to all the lab members for their support, especially Dr Sonam Chorol for her assistance during the experiments, and Mr Prakash Raut for taking care of the birds. I would also like to thank my friends and batchmates, Divyansh Gupta and Aditya Pujari, for code debugging sessions and discussions. I would like to acknowledge IISER Pune, KVPY and funding sources: IA/S/21/1/505621, CRG/2021/004690, DST/CSRI/2017/163.

References

- Canopoli, Alessandro, Joshua A. Herbst, and Richard H. R. Hahnloser. 2014. "A Higher Sensory Brain Region Is Involved in Reversing Reinforcement-Induced Vocal Changes in a Songbird." *Journal of Neuroscience* 34 (20): 7018–26. https://doi.org/10.1523/JNEUROSCI.0266-14.2014.
- Kalra, Shikha, Vishruta Yawatkar, Logan S James, Jon T Sakata, and Raghav Rajan. 2021. "Introductory Gestures before Songbird Vocal Displays Are Shaped by Learning and Biological Predispositions." *Proceedings of the Royal Society B: Biological Sciences* 288 (1943): 20202796. https://doi.org/10.1098/rspb.2020.2796.
- Price, Philip H. 1979. "Developmental Determinants of Structure in Zebra Finch Song." *Journal of Comparative and Physiological Psychology* 93 (2): 260–77. https://doi.org/10.1037/h0077553.
- Rajan, Raghav, and Allison J. Doupe. 2013. "Behavioral and Neural Signatures of Readiness to Initiate a Learned Motor Sequence." *Current Biology* 23 (1): 87–93. https://doi.org/10.1016/j.cub.2012.11.040.
- Rao, Divya, Satoshi Kojima, and Raghav Rajan. 2019. "Sensory Feedback Independent Pre-Song Vocalizations Correlate with Time to Song Initiation." *Journal of Experimental Biology* 222 (7): jeb199042. https://doi.org/10.1242/jeb.199042.
- Rao, Divya. 2022. "Progression of variable repeats of introductory notes to the stable zebra finch song.". PhD Dissertation, IISER Pune.
- Scharff, C., and F. Nottebohm. 1991. "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): 2896–2913. https://doi.org/10.1523/JNEUROSCI.11-09-02896.1991.
- Soha, Jill A., and Peter Marler. 2000. "A Species-Specific Acoustic Cue for Selective Song Learning in the White-Crowned Sparrow." *Animal Behaviour* 60 (3): 297–306. https://doi.org/10.1006/anbe.2000.1499.
- Sossinka, Roland, and Jörg Böhner. 1980. "Song Types in the Zebra Finch *Poephila Guttata Castanotis* [1]." *Zeitschrift Für Tierpsychologie* 53 (2): 123–32. https://doi.org/10.1111/j.1439-0310.1980.tb01044.x.
- Tumer, Evren C., and Michael S. Brainard. 2007. "Performance Variability Enables Adaptive Plasticity of 'Crystallized' Adult Birdsong." *Nature* 450 (7173): 1240–44. https://doi.org/10.1038/nature06390.
- Veit, Lena, Lucas Y Tian, Christian J Monroy Hernandez, and Michael S Brainard. 2021. "Songbirds Can Learn Flexible Contextual Control over Syllable Sequencing." *ELife* 10 (June): e61610. https://doi.org/10.7554/eLife.61610.
- Wiley RH, Richards DG. 1982 Acoustic communication in birds. New York, NY: Academic Press New York