PROGRESSION OF VARIABLE REPEATS OF INTRODUCTORY NOTES TO THE STABLE ZEBRA FINCH SONG

A thesis

Submitted in partial fulfillment of the requirements

Of the degree of

Doctor of Philosophy

By

DIVYA RAO

20132003

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

2021

dedicated to zebra finches

CERTIFICATE

Certified that the work incorporated in the thesis entitled **“Progression of variable repeats of introductory notes to the stable zebra finch song”** submitted by **Divya Rao** was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.

Raghav Rajan

(RAGHAV RAJAN, Ph.D.)
(Supervisor)

Date: 12th September 2021

DECLARATION

I declare that this written submission represents my ideas in my own words and where others' ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Divya.

(DIVYA RAO)

Roll No. : 20132003

Date: 12th September 2021

ACKNOWLEDGEMENTS

The work presented in this thesis has resulted from many (!) years of my Ph.D training. During this period, I have interacted professionally and personally with many people who have been of immense help and support to me, I would like to thank them (as many as I can) here.

First of all, I would like to immensely and deeply thank Raghav, my supervisor (who asked me not to miss acknowledging him). I have been extremely fortunate to be a part of his lab and being trained by him throughout my Ph.D. Right from framing several hypotheses, thinking and rethinking about them, carrying out experiments and analysis, Raghav has always been excited to provide his inputs and right there, in the lab! The interactions with him have always resulted in improving my ideas and work and resulted in a thesis, I feel proud of! I have enjoyed being trained by him for electrophysiology. Pulling out plugs and circuits to deal with noise and finding hundred ways to fix things! Endless hours waiting for the bird to sing have resulted in RA data, that I am happy for eventually making it to my thesis. I also would like to thank Raghav for the HVC data that I started analyzing during the pandemic and these have led to interesting results about progression of INs to song.

One of my first projects in the Rajan lab began with looking at deaf-bird songs. These really got me hooked on to looking at birdsong behavior, staring at spectrograms and appreciating animal behavior. It has also resulted in my first scientific publication and I would like to thank Satoshi Kojima for sharing the deaf-bird data.

One of the tedious jobs of a songbird lab is labelling birdsongs. Many colleagues of the lab (including Raghav) have contributed to labeling songs or providing me with their labelled songs for the extensive age related analysis. Firstly, I would like to thank Ananya and Shijisha who partly contributed to this experiment by recording songs from some of the birds across days. I would also like to thank Shikha who was (is) working on young bird song up tp their development. Many of the recordings for my first time point of adult bird data were obtained from her. Other people who contributed to labelling and recording songs that are part of my data are Vishruta, Sharvari, Gaurav, Harsha, (Raghav, again) and I would like to thank them. I would also like to thank Harini for the head-fixed microphone recordings which helped me make interesting comparisons.

I would like to thank Aditi Agarwal, who trained me to perform perfusion and histology. Looking at beautifully stained brain slices have motivated me to consistently keep trying even when experiments didn’t work and appreciate hard work.

I would like to thank current and past members of the lab for extensive scientific discussions and lab meetings that have been fruitful in shaping my thesis: Shikha, Harini, Aditi, Sanjana, Shivam Rashi (and probably some more I might have missed out)

I would like to thank Prakash Raut for taking good care of the birds in our lab and helping out with them whenever needed.

I would like to thank Aurnab, Collins and Sachin who were in my thesis committee and have always given interesting inputs, and encouragement to go ahead with my future experiment plans during the thesis work.

I would like to thank Department of Biology for taking me as part of the Int. PhD program. Doing a PhD in IISER has been an enriching experience. I would like to thank biology office in help with procuring birds and other requirements for the conducting experiments. I would also like to thank the IISER academic office for helping out at different stages of my PhD.

Other than direct contributions to my Ph.D a lot of people have contributed to maintaining my sanity and in keeping me “cool and calm” when experiments did not work or other random bad days. During the course of PhD I have met with many ups and downs that include laptop crashes, losing a tooth, getting married and COVID and I was fortunate to be surrounded my supportive friends and colleagues at lab and IISER.

Firstly, I would like to thank my batchmates and some of them from senior and junior batches. They have been an extremely important part of fun and food (for the lack of a better word) moments at IISER- Akhila, Amar, Aditi M, Shivani, Kashyap, Swati, Dhriti, Neeladri, Ron, Harpreet, Bharat, Anish, Mehak, Aditi, Sandip, Jay, Adarsh, Anshul, Deepak, Charu, Neha,Nilam, Shikha, Dipti.

I would like to specially thank Shikha for being a constant in the lab giving me company for so many years. As my PhD progressed, I spent a major fraction of my time in the lab and have had an amazing time. The lab has been my second home (actually first!). I would like to thank Anand Krishnan and his lab members for being our next door neighbors (in isolated G1) and people to turn up to for break, chai, etc. I would also like to thank Raghav (again), Deepa, Vinay and Anusha for the lab outings and lunches which were always fun to take time out for.

I would like to thank my family – my parents, Jnana and Deepa for being a constant support away from home, and always being supportive of what I wanted. I thank my in-laws (mother and father-in-law and the very loving Ajji) who patiently waited for me to get done with my thesis. I extend my sincere gratitude to my late father-in-law who passed away recently right before I could submit the revised thesis. In my final years of PhD he would extend his help whenever needed and refrain others from stressing me out on how much time I am left with to finish my degree. I would also like to deeply and sincerely thank my partner for life- Yashraj who has stood by me (though far, but near!) at every step of the last so many years, being my source of strength.

Lastly, I would also like to acknowledge the birds that were part of my thesis. Animal work has made me appreciate animals more and I have learned to deal with extreme care and caution while handling them.

TABLE OF CONTENTS

LIST OF FIGURES.....1
LIST OF TABLES.....4

SYNOPSIS.....5

CHAPTER 1.....9
Introduction.....9
- 1.1 Neuroanatomy of movement generation.....10
- 1.2 Movement initiation – a complex phenomenon of the brain.....11
- 1.3 Motor preparation during delayed movement initiation.....11
- 1.3.1 Properties of preparatory activity.....13
- 1.4 Movement initiation before self-timed movements.....17
- 1.5 Movement initiation of natural, self-initiated and learned complex movements remain unexplained.....19
- 1.6 The zebra finch song: a natural motor skill.....21
- 1.7 Introductory notes delay song initiation and systematically progress to song.....26
- 1.8 What is known about INs and neural activity during INs?.....26
- 1.9 What could be the function of INs?.....27
- 1.10 Understanding progression of INs to song important for testing motor preparation hypothesis.....28
- 1.10.1 How do INs progress to song?.....28

CHAPTER 2.....30
Sensory feedback independent pre-song vocalizations correlate with time to song initiation.....30
- 2.1 INTRODUCTION.....31
- 2.2 METHODS.....32
- 2.2.1 Birds and song recording.....32
- 2.2.2 Tracheosyringeal nerve cut and sham surgery.....33
- 2.2.3 Deafening.....34
- 2.2.4 Data analysis.....34
- 2.3 RESULTS.....38
- 2.3.1 Song spectral structure alone is affected after ts nerve surgery.....38
- 2.3.2 IN acoustic structure, not duration, is affected by ts-cut surgery.....40
- 2.3.3 Mean IN number before song is not affected by removal of proprioceptive or auditory feedback.....40
- 2.3.4 Progression of IN timing to song is not affected by removal of proprioceptive or auditory feedback.....43
- 2.3.5 Progression of IN acoustic features to song is not affected by removal of proprioceptive or auditory feedback.....45
- 2.3.6 IN number is reduced when calls precede the first IN of a song bout.....47
- 2.3.7 Calls just before the first IN of a song bout correlate with altered ‘initial’ state.....48
- 2.3.8 ‘Initial’ state of IN progression correlates with time to song initiation.....51
- 2.4 DISCUSSION.....51
- 2.4.1 Contributions of respiratory feedback to song initiation.....52
- 2.4.2 Long-term requirement for sensory feedback.....52
- 2.4.3 Comparison of INs to motor preparation in other systems.....52
- 2.4.4 Overt movements in other systems as motor preparation.....53
- 2.4.5 Mechanisms for IN progression to song.....53

CHAPTER 3....................55
Relationship between Introductory Notes and Song....................55
3.1. INTRODUCTION....................56
3.2. METHODS....................58
3.2.1. Birds and Song Recording....................58
3.2.2. Data analysis....................60
3.3. RESULTS....................70
3.3.1. Variability of IN intervals is not correlated to variability of upcoming-song intervals.......70
3.3.2. Variability of acoustic properties of INs is correlated to variability of upcoming-song syllables....................73
3.3.3. Age related changes to IN sequences in adults show increase in IN number and duration along with changes in temporal patterning and syllable acoustics....................75
3.3.4. Age related changes to song in adults show improved spectral stereotypy, speed and sequence regularity....................81
3.3.5. Age related changes in INs and song properties are correlated....................89
3.3.6. Temporal properties of INs are not linked to song timing, acoustics or sequencing...........91
3.3.7. The temporal progression of INs is unrelated to repetitions within and outside song.........93
3.4. DISCUSSION....................95
3.4.1. Shared control and predictive relationship between IN and song....................96
3.4.2. Differences in the control of INs and songs....................96
3.4.3. Age related changes of INs....................97
3.4.4. Age related changes to first song of the bout....................97

CHAPTER 4....................99
Neural Correlates of Introductory Note Progression to Song....................99
4.1. INTRODUCTION....................100
4.2. METHODS....................102
4.2.1. Birds....................102
4.2.2. Details of overlap with previous studies....................103
4.2.3. Assembly of microdrives and tetrodes for extracellular recording....................103
4.2.4. Surgery and electrode implantation....................104
4.2.5. Standardizing stereotaxic co-ordinates for RA....................105
4.2.6. Song and neural recording....................106
4.2.7. Data analysis....................107
4.3. RESULTS....................110
4.3.1. HVC interneurons show interval related activity that is different at the start and the end of IN progression....................110
4.3.2. Interneuron activity related to intervals does not indicate onset of next syllable.............110
4.3.3. Interneuron activity related to intervals encodes position in the sequence....................113
4.3.4. Position related activity of INs represented by activity at first, middle and last positions 114
4.3.5. Activity of HVC interneurons at different positions of INs start the same but end differently....................115
4.3.6. Multiunit activity of RA neurons is consistent across IN positions, unlike HVC.............117
4.4. DISCUSSION....................117
4.4.1. Representation of IN progression by HVC_{RA} neurons....................119
4.4.2. HVC representing past and future syllables....................121
4.4.3. Interval related activity in HVC does not arise from within HVC....................122
4.4.4. Contribution of single neurons in RA to IN progression....................122
4.4.5. Propagation of IN progression through song motor pathway....................123
4.4.6. Representation of IN progression and sequence representation in HVC....................123

CHAPTER 5..124
Discussion..124
    5.1. Summary of results...125
    5.2. Mechanism for song initiation..126
    5.4. Other functions of INs...128
    5.5. Comparing IN progression to preparatory activity..129
    5.6. Comparing representation of IN progression to preparatory activity.......................129
    5.7. Preparatory-movements before self-initiated natural movements...........................130
    5.8. Initiation of complex natural movements..130
    5.9. Future Implications...131

REFERENCES...132

APPENDIX
    p-value of statistical tests in Chapter 3
    Published Article

LIST OF FIGURES

Fig. 1.1 Movement initiation of natural, self-initiated, complex movements remain unexplained

Fig. 1.2: Neural control of song by song motor pathway and song learning pathway

Fig. 2.1: Spectrogram of a zebra finch song bout with introductory notes (INs) and motifs

Fig. 2.2: Song structure after sham surgery, tracheosyringeal nerve cut or deafening

Fig. 2.3: Changes in IN acoustic features after removal of sensory feedback

Fig. 2.4: Changes in motif syllable log amplitude after removal of sensory feedback

Fig. 2.5: Mean IN number and variability are not affected by removal of sensory feedback

Fig. 2.6: Comparison of day-to-day changes in IN properties with changes observed after surgery for sham-surgery, ts-cut and deaf birds

Fig. 2.7: Progression of intervals between successive INs is independent of sensory feedback

Fig. 2.8: Progression of IN timing and acoustic features is not altered by removal of auditory or proprioceptive feedback

Fig. 2.9: Progression of acoustic properties of successive INs is independent of sensory feedback

Fig. 2.10: Calls just before the first IN correlate with fewer INs at the start of the song bout

Fig. 2.11: Calls just before the first IN correlate with a shorter interval between the first two INs

Fig. 2.12: Changes in IN "initial" state in bouts with calls before the first IN also depend on the time between the end of the call and the start of the first IN

Fig. 2.13: 'Initial' state of IN progression correlates with song initiation delay

Fig. 3.1: Hypotheses for IN-song relationship

Fig. 3.2: Determining minimum number of bouts for IN analysis: example from one recording session of one bird

Fig. 3.3: Age of recording across birds

Fig. 3.4: Time and type of recording

Fig. 3.5: Selecting bout interval criteria for sample bird recorded on multiple days

Fig. 3.6: Introductory Notes (INs) are variable relative to the song

Fig. 3.7: Trial-to-trial variation of IN intervals is not correlated to variation of upcoming song intervals

Fig. 3.8: Trial-to-trial variation of IN acoustic properties was correlated to variation in acoustic properties of the upcoming song syllables

Fig. 3.9: The mean IN number increases in the first year of age

Fig. 3.10: Change in mean properties of IN progression with age

Fig. 3.11: Change in variability of IN progression with age

Fig. 3.12: Change in mean acoustic properties of IN syllable with age

Fig. 3.13: Change in variability of acoustic properties of IN syllable with age

Fig. 3.14: Change in mean and variability of IN temporal sequence with age

Fig. 3.15: Correlation matrix for change in IN properties

Fig. 3.16: Change in song stereotypy with age

Fig. 3.17: Change in motif properties with age

Fig. 3.18: Change in song properties beyond first motif

Fig. 3.19: Change in mean acoustic properties of song syllable with age

Fig. 3.20: Change in variability of acoustic properties of song syllable with age

Fig. 3.21: Correlation matrix for change in song properties with age

Fig. 3.22: Correlation matrix for change in song vs change in IN with age

Fig. 3.23: Correlation matrix for change in song vs change in IN with age

Fig. 3.24: Correlated IN and song changes with age are not related on given day

Fig. 3.25: Temporal progression of INs is not present in other repeats

Fig. 4.1: Standardization experiments were done to locate RA

Fig. 4.2: Custom-designed implant

Fig. 4.3: Interneurons show interval related activity that is different between first and last position of INs

Fig. 4.4: Interval related activity is not related to duration of the interval

Fig. 4.5: Interval related activity for the same period of first and last interval is not consistent across neurons.

Fig. 4.6: Interval related activity is specific to the position of IN it follows

Fig. 4.7: Spike counts are consistent across IN positions

Fig. 4.8: The pattern of activity for an IN position is similarly correlated when all middle positions are combined

Fig. 4.9: The average pattern of activity varied between first, middle and last IN positions

Fig. 4.10: Relative activity of a neuron at first, middle and last position

Fig. 4.11: Position related activity of HVC interneurons end differently for first, middle and last INs

Fig. 4.12: Extracellular recording in RA in an awake singing bird

Fig. 4.13: Activity across multiunit sites in RA start the same and end the same for all positions of INs

Fig. 4.14: Position related activity indicates representation of progression by representing past and future syllables

LIST OF TABLES

TABLE 3.1: Variable properties of INs – average properties across birds

TABLE 3.2: Correlation between IN and song intervals across trials

TABLE 3.3: Correlation between IN and song syllables

SYNOPSIS

Movement initiation has puzzled researchers since the finding that areas in the brain that control production of movements are also active several hundreds of milliseconds before the movement. Although, it was soon understood that such activity is important for the upcoming movement, we still do not understand completely what this activity encodes. A plausible theory is that the activity before movement initiation is related to the brain preparing for motor commands prior to movement execution. In the light of this theory, several experiments have been carried out to understand the nature of this "preparatory" activity and what aspects of movement may get prepared before movement onset. Chapter 1 of this thesis summarizes the findings of these experiments, which primarily focussed on simple, externally triggered, movements, like reaching with an arm to touch an object. These experiments were carried out primarily in primates and rodents. Animals were required to perform simple reaching movements as part of instruction based tasks. A major caveat of these tasks was that the ethological relevance of these movements is unknown. The field lacks a model for natural behavior which is ethologically relevant and existing results cannot be directly extrapolated to natural movements that are relatively more complex and can also be self-initiated. The song of the zebra finch is an excellent example of a natural movement with ethological relevance as the birds produce songs to court mates. These songs consist of sequences of sounds interleaved with silent gaps. Just like reaching movements involve muscles, these songs are also produced as a result of the activity of respiratory and vocal muscles. The songs are highly stereotyped every time the bird sings and have been extensively studied in the context of production of complex natural movements. Although these have not been explicitly studied in the context of movement initiation, they provide a natural counterpart to study the same. Zebra finch song is known to begin with repeats of variable introductory notes or INs. These INs are followed by the stereotyped songs. Whether these INs play a role in initiation of song has not been widely studied. Recent work on INs show that as INs are produced in repetition, the intervals between them successively reduce and the acoustic properties converge on a consistent acoustic state. These findings suggest a preparatory role of INs for song, similar to the neural preparatory activity before the start of simple movements in primates and rodents. Chapter 1 describes what is known about the song of the zebra finch and further makes prediction on the preparatory role. Finally, it suggests one of the ways to understand what INs represent is by asking how INs progress to the song. In Chapter 2, 3, 4, I describe the various related hypothesis that I tested, in order to understand the progression of INs to song.

In Chapter 2, I address an important question related to the behavioral mechanism of IN progression. Given the variable nature of INs and their occurrence before song a possible "preparatory" role could be the calibration of the vocal apparatus before song. The brain may utilize feedback from INs in order to change the state of vocal muscles and prepare them for song. This hypothesis was tested by surgical manipulation of feedback pathways. The results show that progression of INs continues without any change, following feedback removal and suggests song initiation as an internal mechanism. This further suggests that the progression of INs represents internal states changing in the brain.

In Chapter 3, I address the question of how INs may relate to the song. INs are vocalizations like song and thus may hold parallels to the song in the way it is controlled. Further the proposed preparatory mechanism for INs suggests that properties of INs and songs may be related to each other. These different interpretations of IN representation make it all the more important to understand what aspects of INs and songs are related to each other. To understand this, I compared INs and songs in two ways : (1) whether INs and songs were related on a trial by trial basis and (2) whether known changes to song (with age) are associated with corresponding changes in INs. These analyses reveal correlations in the acoustic properties of INs and song suggesting similarities. Since INs are repeated syllables, I also compared the properties of INs with the properties of song syllables that are repeated. This was done in a set of birds chosen for syllable repetition in their songs, as this is not common in zebra finches. Using this comparison, I show that the temporal speeding up of INs as they repeat is a unique feature of INs possibly important for song initiation. Changes in acoustic features of INs were also seen for other repeated syllables suggesting that the acoustic features of INs are like those of other song vocalizations. Overall, these results suggest that the unique temporal features of INs could be related to song initiation (not motor preparation), while the acoustic features are controlled similar to the song syllables.

Based on results of Chapter 2 and Chapter 3, in Chapter 4 I analyze the neural correlates of IN progression in two nuclei of the song motor pathway. I specifically focused on HVC, a premotor nucleus, important for normal song production and its downstream motor nucleus RA (robust nucleus of arcopallium). Neurons recorded during singing were analyzed for neural correlates of IN progression. I observed changes in activity pattern of HVC interneurons with respect to position of INs in the repeat as first, middle, or last IN. The differences in activity of neurons for the position revealed different neuron types and suggested a combination of neurons may represent progression of INs. Similar changes were not seen in RA. Multiunit activity in RA was highly similar for all IN positions. The IN-position based activity of neurons in HVC were indicative of previous or next syllable suggesting a mechanism of reconfiguration of HVC as it keeps track of past and future events. This is similar to what has been reported in the activity of HVC neurons in other songbird species with more variable songs.

Finally, in Chapter 5, I discuss the implications of the findings of the thesis. The findings of this study suggest that all aspects of INs may not play a role in motor preparation. However, the timing between INs and the representation of IN progression in the song motor pathway suggest ways in which the brain keeps track of time as it prepares for the movement. Further, this work shows variable sequencing related neural activity in HVC in zebra finches, similar to what is known for other song birds like bengalese finches and canaries that produce complex songs. Overall, through this work, I demonstrate many similarities and differences in the initiation of natural movements, like song, and simple reaching movements. Thus, my results suggest that INs and song initiation could be a behavior used to further understand variable motor sequencing and song initiation.

CHAPTER 1

Introduction

We perform several movements to go about our daily lives. From walking to talking, the brain engages appropriate muscles to contract or relax as we produce well-coordinated desired actions. Researchers, back in the late 19th and early 20th century, evoked twitching of various muscles in the body by electrically stimulating the frontal area of the mammalian brain (dogs, non-human primates and humans) (Fritsch and Hitzig, 1870; Leyton and Sherrington, 1917; Penfield and Boldrey, 1937; Gross, 2007; Carlson and Devinsky, 2009). Lesions within this area altered but did not abolish movements (Fritsch and Hitzig, 1870; Gross, 2007). The discovery of this excitable area, also called the motor cortex showed that higher areas of the brain are involved in the control of movements (Gross, 2007; Carlson and Devinsky, 2009).

1.1 Neuroanatomy of movement generation

How does the brain generate movements? A cascade of electrical events propagate from neurons in the brain including the motor areas, to the spinal cord (both part of the central nervous system), and in turn to the peripheral neurons innervating the muscles. The electrical inputs received at the muscles (via neuromuscular junction) drive their contraction leading to movement generation Kandel et al., 2000). All of this occurs swiftly as electrical activity conducts through a neuron in the form of units of information called action potentials lasting ~ 1ms (Kandel et al., 2000) ; and transmits from one neuron to the next via synapses in ~ 5 ms (Kandel et al., 2000). The total time for the transduction of electrical impulses over a series of neurons and synapses, and finally to synapses on the muscles determine the time for movement production. This is typically be in multiples of tens and within hundreds of milliseconds. Based on the transduction time, the activity of neurons in the brain that control movement production reliably lead the movement. Such movement related activity in humans and primates have been found to occur earliest in areas of the motor cortex namely primary motor cortex, premotor cortex and supplementary motor area (Romo and Schultz, 1987; Maimon and Assad, 2006; Kaufman et al., 2013; Ebbesen and Brecht, 2017). Although, these areas control movement generation, the specific parameters of the movement that they encode is not well-understood and is an active field of research (Shenoy et al., 2013). Interestingly, these areas also show reliable changes in activity long before (a few hundreds of milliseconds before) the transduction suggesting a complex phenomenon precedes movement generation that may affect the time for movement initiation.

1.2 Movement initiation – a complex phenomenon of the brain

How fast can the brain initiate controlled and coordinated movements? The conduction velocities of currents reaching the muscles (Donders, 1969) implicate fast movement execution times of the order of tens of milliseconds. In nature, we voluntarily produce movements either in response to an external stimulus or spontaneously (self-initiated). In behavioral experiments on non-human primates where subjects responded with a defined movement after the presentation of a stimulus, the response latency was shown to be of the order of hundreds of milliseconds, and also exhibited trial-to-trial variability (Hanes and Schall, 1996). Such long response latencies exceed the total time for the stimulus to be electrically conducted to the brain and the brain to act on the muscles. This suggested that additional neural events occur after the sensory stimulus reaches the brain and before motor cortex discharge produces the movement. These intervening events causing variable delay make movement initiation a complex phenomenon. The fact that more "thoughts" are at play before the brain executes movements, paved the way for movement initiation as an exciting field of research in neuroscience.

How are movements initiated by the brain? Addressing this question finds significance in understanding diseases like Parkinson's and speech apraxia – where movement initiation is defective (Cunnington et al., 1997; Kent and Kent, 2000), and also in various sports – where improving the speed of actions/reactions provides an advantage . Much work has been done for over a century to understand the nature of neural events in the brain that allow for movements to eventually initiate in response to a stimulus (Wong et al., 2015). However, our understanding of movement initiation in the absence of a stimulus (self-initiated movements), and further in the context of complex natural movements remains poorly understood. The following sections summarize the findings and gaps in our knowledge about movement initiation and leads to the broad question of this thesis about how complex natural movements are initiated.

1.3 Motor preparation during delayed movement initiation

What "thoughts" impede movement initiation? When responding to an external stimulus, the events leading to movement initiation have been attributed to (i) sensory perception of the stimulus, (ii) decision making processes related to transforming of sensory information to selection of a motor output and (iii) planning a structured motor output (Svoboda and Li, 2018). Although sequential, these events may overlap. Accordingly in the brain, the sensory areas, the areas involved in decision-making as well as the motor areas show changes in activity during the intervening period between the stimulus and the response. The activity in all of these areas put together, have to eventually produce the motor output in response to the stimulus. In this context, the changes in activity of the motor areas during the intervening period is of particular interest as it appears much earlier than the activity related to actual movement execution. Although, this early activity does not immediately produce the movement, results from further studies support their possible role in planning the motor output. It was shown that the early activity is indicative of the type of upcoming movement. Moreover, allowing more time for this activity to occur shortens the response time and, disrupting this activity increases response time (Tanji and Evarts, 1976). Due to its relationship with the type and timing of the upcoming movement, the early changes in activity in motor areas are called preparatory activity and suggests that motor preparation by the brain precedes movement initiation.

What happens during motor preparation? A plausible hypothesis is that the neural events that form preparatory activity contribute towards motor control in anticipation of the upcoming movement. Producing a specific movement requires motor control that allows appropriate muscles to contract or relax in sequence at the right moment. Any disruption of this control will thus cause interruption or incorrect movement. It was proposed that motor control may not be restricted to occur at the time of an ongoing movement (online control) but may manifest much earlier than the onset of the movement. In order to produce a controlled precise movement, an entirely online control would utilize sensory feedback from the muscles in the preceding moment to drive the next muscles in the sequence. Motor control in this manner would make movements slow and limited by sensory feedback. Contrary to this, movements found in nature can be very rapid such as wing-beating in birds and insects. It was also shown that removal of sensory feedback did not t change rhythm generation in fast wing-beating in locusts, although the rhythm was slower without sensory feedback (Wilson 1961). Together, it supported that motor control may be manifested prior to the execution of the movement by formulation of a motor program. Keele described motor program as "a set of muscle commands that are structured before a movement sequence begins, and that allows the entire sequence to be carried out uninfluenced by peripheral feedback" (Keele, 1968). The existence of preparatory activity in motor areas of primates and rodents supports the idea that formulation of a motor program precedes movement initiation. It further predicts that preparatory activity encodes the structuring of the motor commands in the program and acts as motor control.

1.3.1 Properties of preparatory activity

To understand the nature of the preparatory activity and the motor commands it may encode, stimulus response tasks with an instructed delay to respond have been widely used as a behavioral tool (Tanji and Evarts, 1976). These experiments have been performed in non-human primates, predominantly monkeys, and more recently in mice (Shenoy et al., 2013; Li et al., 2015). The subjects are trained to learn the instruction-based task (Fig. 1.1 A). Typically in these tasks, a stimulus (visual stimulus in case of primates; and an auditory or somatosensory stimulus in the case of mice) appears instructing what the movement response has to be. Following this, there is a variable delay period. During this period, the subjects have to remain still and remember the movement that has to be made. Finally, after the delay period, there is a go-cue that instructs the subject to make the appropriate movement in response to the stimulus. The imposed delay period in between allows separation of the sensory period from the motor period. Reaction times are measured for every trial as the time between the go-cue and the onset of the movement. The imposed delay between the stimulus onset and the go-cue has a significant role in understanding the nature of preparatory activity. It was shown that longer imposed delays shorten the reaction time. In a hand movement task, increasing the delay period up to ~100-200 ms reduced the reaction times by 30 ms, beyond which there was no effect. Preparation was suggested to begin ~100-200 ms before go-cue for that particular task (Tanji and Evarts, 1976; Churchland and Shenoy, 2007). Thus, the imposed delay provides a time window for preparation to begin while holding on to making the movement. The activity of neurons is recorded during both the imposed delay and the reaction time to determine deviations from baseline activity during this period. Such preparatory activity has been recorded in PMd (dorsal premotor) and M1 (motor cortex) areas of monkeys for hand movements (Churchland et al., 2006a), FEF (Frontal Eye Field) of monkeys for eye movements (Hanes and Schall, 1996), FOF (Frontal Orienting Field) and ALM (Anterior Lateral Motor cortex) area of mice for licking behavior (Erlich et al., 2011; Guo et al., 2014; Li et al., 2016). Findings about the nature of preparatory activity and its relationship with movement are summarized below.

- *Preparatory activity is associated with different firing rate patterns across neurons within and across brain areas*. The FEF neurons (in monkeys) show gradual increase in firing rate following stimulus onset that peaks at the time of initiation of eye movement (saccade) (Hanes and Schall, 1996b). Preparatory activity in individual neurons of ALM (in mice) may show ramping or persistent activity (Li et al., 2016). Such activity has been shown to be associated with short-term memory (Fuster and Alexander, 1971), suggesting preparatory activity may act as a bridge between past and future events

- *Preparatory activity at the level of single neurons can predict reaction times or time for movement initiation.* The variability in neural activity of PMd neurons (in monkeys) at the time of the go-cue predicts reaction times corresponding to varying periods of imposed delay (Rosenbaum, 1980; Churchland et al., 2006a). Activity of FEF neurons (in monkeys) that show an increase in firing rate before eye movements can be modeled as an activation function that predicts reaction time distributions across trials (Hanes and Schall, 1996).
- *Correlations between preparatory activity and features of the upcoming movement.* Many studies have found different properties of movement that might be correlated to preparatory activity. Preparatory activity in PMD/M1 before reaching movements (in monkeys) was shown to be tuned for spatial aspects of movement like target location – direction and distance of the target Messier and Kalaska, 2000; Churchland et al., 2006b). Preparatory activity in the same areas was also shown to relate to non-spatial aspects like grasp type, curvature of the reach, or speed of reach (Godschalk et al., 1985; Hocherman and Wise, 1991; Churchland et al., 2006c). The tuning for different parameters may arise from correlations between parameters themselves and thus pinpointing at the exact motor commands that preparatory activity may encode is difficult.
- *Disrupting preparatory activity influences upcoming movement.* Disrupting preparatory activity in monkey premotor area PMd (dorsal premotor cortex), using microstimulation around the time of go-cue, delays movement onset (Churchland and Shenoy, 2007). Similarly, disrupting preparatory activity in mice ALM by photoinhibition can cause a movement bias towards left or right lick (Li et al., 2016). In both monkey PMd and mice ALM, the effect of disruption is more pronounced for late compared to early preparatory period (late period is <100 ms before go cue for monkey hand movement; <300ms before go cue for mice lick). The early vs. late effect can reflect involvement of areas during preparation. Disruption of late preparatory activity in M1, which is downstream of PMd had no effect on movement initiation, suggesting stronger role of PMd in preparation for start of the movement (Churchland and Shenoy, 2007). In mice ALM, bilateral but not unilateral photoinhibition during early delay period causes movement bias, suggesting redundancy of preparatory activity in both hemispheres. The robustness of activity due to redundancy in both hemispheres have been modeled by representing each area involved in preparation as

an independent module and the connections between these modules allow recovery on perturbation (Li et al., 2016). This study suggests that preparatory activity may be independently driven in different areas or groups of areas involved in preparation towards movement.

- *Correlations between neural activity during preparation and movement*

  As mentioned earlier, preparatory activity at the level of single neurons show selectivity among the different types of the same movement. For example, preference among lick direction (left vs right; in mice), extent of reach (near vs far; in monkeys) or speed of reaching the target (slow vs. fast; in monkeys) (Churchland et al., 2006c; Li et al., 2016). Because such selectivity for a movement type is also shown by neural activity during the movement, two related hypotheses about preparatory activity predicted correlation between activity during preparation and movement. First, preparatory activity may be subthreshold or weaker versions of the activity during the movement. Second, preparatory activity is weak owing to stronger inhibition that is released at the time of the movement. However, it was observed (in monkeys) that across individual neurons the activity tuning or preference for a hand movement type may be similar or different between preparation and movement periods. Together with other studies, this shows that preparatory activity is not sub-threshold activity (Churchland et al., 2010) or a weakened activity related to movement execution that is released by inhibition (Kaufman et al., 2014). Averaging the preparatory and movement related activity across neurons show weak positive correlation between the two. Despite the weak correlations, firing rate patterns of population of neurons when modeled as states of dynamical system could predict neural activity during the movement. In this model the firing rate of individual neurons in the population form the axes of a state-pace. Each state is represented by firing rate of individual neurons in the population at a given time. The firing rate of the neural population from an initial state evolved over time to form trajectories in the state-space representing the change in activity. The initial state of dynamical system could predict future trajectory evolving into the firing rate pattern of the neural population at the time of the movement. The activity during the movement may directly represent movement parameters. This predictive relationship between population neural activity during preparation and movement suggests that the relationship of preparatory activity with exact movement parameters as such may be indirect. Taken together, studies suggest relationship between preparatory activity and the upcoming movement, but the exact muscle commands or movement properties encoded by preparatory activity remain unclear.

- *Preparatory activity changes over time to reach a readiness point before movement initiation.* Activity of PMd neurons (in monkeys) drops in neural variability following stimulus onset. The rate of decline is such that at the time of the go-cue the variability can predict the time left for hand movement initiation (Churchland et al., 2006a). Contrary to PMd where neural variability of single neurons changes at a consistent rate, FEF neurons that increase their firing rate before eye movements (in monkeys) show changes at different rates to reach a consistent threshold activity across trials (Hanes and Schall, 1996). The different rates predict the time for movement initiation. This fits well with an accumulator model of decision-making process for response preparation and indicates preparedness for upcoming movement may involve reaching a threshold for movement initiation. Although, single neuron activity related to eye-movement (of monkeys) fit the rise-to-threshold model, single neurons in movement related areas for hand movements (of monkeys) failed to show the same. However, analyzing populations of neurons in dynamical systems model did show a trend for readiness. The firing rates of a population of neurons during the period after stimulus onset to movement initiation can be modeled as states in a dynamical system evolving over time (Fig. 1.1 B). Here, the current state predicts the next state Churchland et al., 2010). The population state is initially variable but reaches a consistent state just before movement onset, suggesting an optimal-state space is reached before movement initiation (Shenoy et al., 2013).
- *Preparatory activity is defined as activity in motor related areas that does not cause pre-movements.* This is explained by a model where population activity of neurons is such that its output to downstream area predominantly cancels out. Support for this model comes from the observation (in monkeys) that preparatory activity in PMd is more than that in downstream M1 area. Extrapolating this model further downstream predicts negligible muscle EMG activity, resulting in prevention of immature movements (Kaufman et al., 2014).

Together, the stimulus-response studies on preparatory activity have established that motor preparation delays movement initiation and the state of preparedness affects movements. A major caveat in these studies is the specificity towards externally triggered simple movements. Attempts have been made to modify these experiments and incorporate the self-initiation aspect, as discussed further.

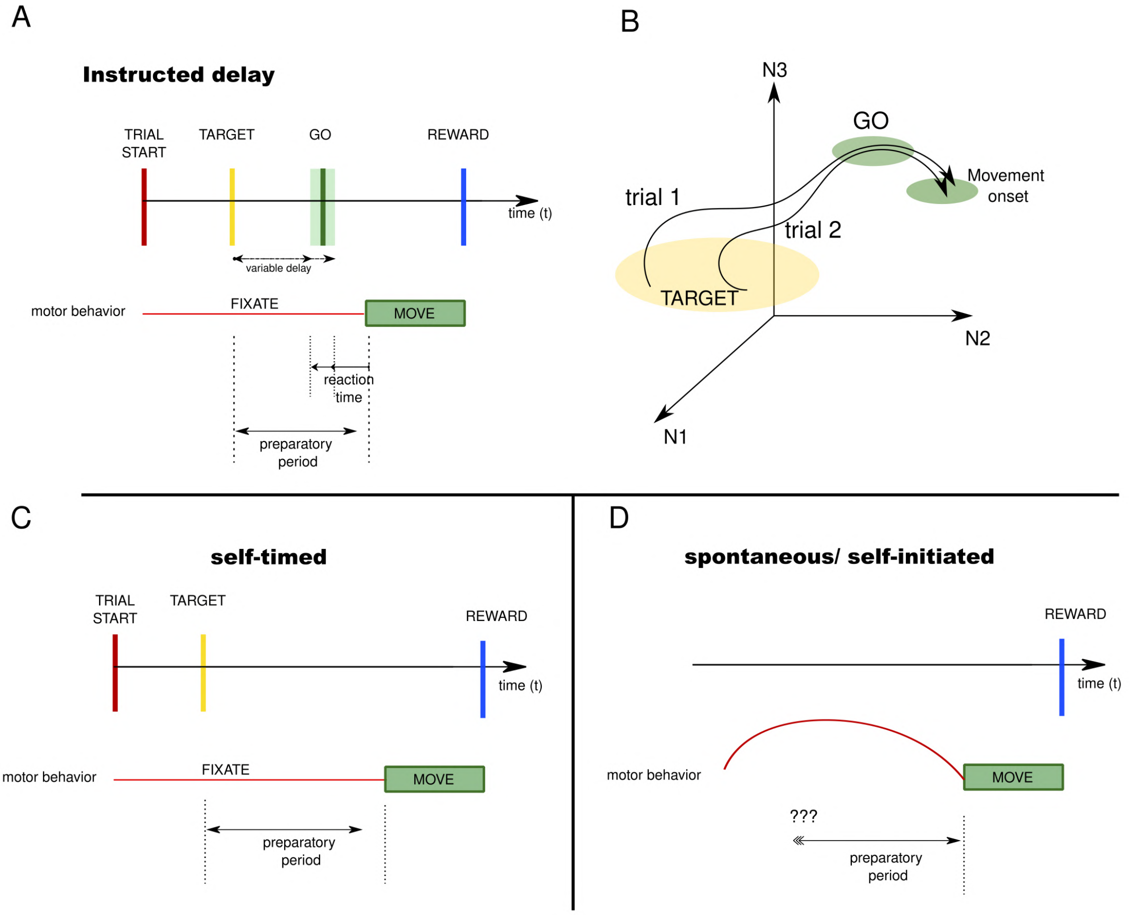

Fig. 1.1 Movement initiation of natural, self-initiated, complex movements remain unexplained
(A) Instructed delay task: Subject has to remain still (fixate) as the trial starts and the target appears; subject has to move when a go cue occurs after a variable delay in order to obtain a reward. Instructed delay allows a preparatory period to study motor planning. (B) An example population of 3 neurons involved in preparation. The axes represent the firing rates of each neuron. Preparatory activity before movement onset for two example trials is represented by the two trajectories. The ellipses represent the distribution of preparatory activity across multiple trials at the time of target presentation (yellow), go cue (green) and movement onset (green). Population activity across trials is initially variable and becomes more consistent by the time of movement onset (adaptation from previous studies). (C) In self-timed trials, subject fixates, a target appears, subject can move any time in response to obtain the reward. Preparatory period begins after target presentation and ends whenever subject moves. (D) Natural movements are more often spontaneous, self-initiated movements, have no fixation and can occur anytime as there is no instructed go cue. Preparation for these movements is unclear and remain unexplored.

1.4 Movement initiation before self-timed movements

Contrary to the movements triggered by an external stimulus, natural movements are often spontaneous or self-initiated. Examples being reaching out to a bottle of water, taking a walk, or playing the piano. In the absence of an external trigger, what in the brain determines the "initiation" of 'self-initiated' movements? In this regard, slow build-up of negative EEG potentials above SMA (supplementary motor areas) of humans were observed starting more than 1-2 seconds before self-paced finger movements (Kornhuber and Deecke, 1965). This signal in anticipation of upcoming movement was called the Bereitschaft's or readiness potential (RP). The onset of RP predominantly from SMA suggested the role of SMA in preparation and initiation of movements (Kornhuber and Deecke, 1965; Kutas and Donchin, 1980). Since then RP has been observed in many connected areas of SMA, suggesting self-initiated movements result from activity in loops of interacting areas.

Identifying the neural correlates of self-initiated movements using existing animal models is a challenge. In order to identify reliable neural activity related to a movement, multiple instances of the movement are required. Typically, this is achieved by training of tasks involving production of the desired movement. However, training animals to spontaneously produce a given movement without an external stimulus is difficult. To get around this problem, a variant of self-initiated movements called self-timed movements have been used to train monkeys and rodents (Fig. 1.1 C). The self-timed trials also called as waiting trials begin with fixation followed by a stimulus and an initial delay, similar to stimulus-response trials. Subsequent to the initial delay, there is no go-cue and the subject is allowed to respond at any time to get the reward. Thus, the movement produced is self-timed. These trials have been used to identify neural correlates that determine the time for movement initiation. Further, the tasks involve self-timed trials interleaved with stimulus-response trials of the same movement in order to compare activity related to movement initiation between the two (Romo and Schultz, 1987; Lee et al., 2006; Maimon and Assad, 2006).

In many areas suggested to be involved in self-initiated movements, the firing rates of a subset of neurons show ramping activity that builds-up towards movement onset of self-timed trials. The slow gradual ramping is reminiscent of RP. Across species (humans and monkeys) and brain areas, ramping-up or ramping-down may begin more than few hundreds of milliseconds or more than 1-2 seconds before movement (Lee et al., 2006; Maimon and Assad, 2006; Fried et al., 2011). On the contrary, a steep ramping just before the movement is observed for stimulus-response trials that catches up with self-timed trials by the time of movement onset. The selective action for self-timed trials is present in neurons of putamen (part of basal ganglia) and parietal area 5 (motor-related area) in monkeys (Lee and Assad, 2003; Lee et al., 2006; Maimon and Assad, 2006). These studies suggest initiation of spontaneous movements is a slow gradual process compared to externally-triggered movements. Further, ramp-to-threshold neurons determine the timing for initiation of self-timed movements.

The mechanism of ramp-to-threshold that triggers self-initiated movements is explained by an integration-to-bound model. According to this model, the integration of inputs from other neurons result in ramping activity that reaches a threshold firing rate and in turn triggers the movement. Support for this model comes from recordings of neural activity in rodent M2 (secondary motor cortex), an area functionally homologous to primate SMA (Rigato et al., 2014). A variant of waiting task was employed. The trial began with the rat poking its nose in a waiting port followed by playback of an auditory cue – tone 1. At any time after tone 1, the rat could poke its nose in the reward port to gain a small reward. This was called the impatient trial. Alternatively, the rat could wait longer for a second tone and immediately get a larger reward by poking into the reward port. This was called the patient trial. The impatient trial allowed the rat to perform self-timed movements at variable times. A population of M2 neurons showed ramping-rates proportional to the trial waiting time. Further, a population of transiently active neurons were identified with firing rates proportional to the trial waiting time. The activity of simultaneously recorded transient neurons when used as inputs to the integration-to-bound model, produced ramping activity similar to those experimentally observed for ramp-to-threshold neurons. This study suggests that self-initiated movements are possibly generated by internally triggered neural activity dynamics based on integration-to-bound models.

Together, the above-mentioned studies suggest initiation is different between externally-triggered and self-initiated movements. The focus of experiments on self-timed movements has been limited to determining the time for initiation. However, how preparatory activity in these areas is driven towards movement and how it relates to upcoming movements has not been studied.

1.5 Movement initiation of natural, self-initiated and learned complex movements remain unexplained

The specificity of the movements studied by training animals on artificial paradigms raise several unanswered questions about how more natural movements may initiate (Fig. 1.1 D). The caveats elaborated below limit extrapolating results from current studies to natural movements.

Although self-timed trials are considered a version of self-initiated movements, the spontaneity allowed in these tasks is restricted and may limit possible mechanisms for movement initiation. The allowed time to initiate movement is longer than that in a stimulus-response trial, but is limited compared to naturally occurring spontaneous movements. The self-timed movement is learned in response to the stimulus and thus variability allowed in these movements may be restricted by the task. Further, when movements follow a stimulus and is heavily instruction based, one would expect that a lot more aspects about the movement have already been ready as the subject orients itself to perform the task. Thus, “true” self-initiated movements will require preparation of more aspects of upcoming movement and may take longer. On the contrary, spontaneous movements may lack the need to have a targeted response and exercise less motor control, leading to shorter preparation time and more variable movements. Thus, motor preparation before self-initiated movements has different possibilities. Moreover, in instruction-based tasks the subjects are instructed to fixate or remain still at the start of trial. Without such restrictions, how preparatory activity would evolve for self-initiated movements is unclear.

The extrapolation of preparatory activity in simple movements to complex motor behaviors is also not straightforward. Adding complexity with multiple movements would mean more commands to be set in place. However, many complex behaviors are learned by practice and learning is known to improve speed and accuracy of movement pattern. Reaction times can get faster due to learning, as has been shown for simple movements. Many behaviors that we perform are heavily practiced, become habits that we perform without giving much “thought” to it. Thus, addition of complexity and learning may have different effects on preparatory period of movements. The nature of preparatory period for learned and complex motor sequence cannot be effectively predicted based on current studies.

The instruction-based training paradigms in mice and monkeys, have been extremely useful tools to understand neural mechanisms of movement initiation. They are well-controlled by the experimenter and the training allows repeatability of the behavior which is important to identify relevant neural activity. However, the ethological relevance of the movements performed is unclear. In order to test various hypotheses on brain function, there is a push in the field to study natural behaviors whose ethological relevance in the natural environment are known and are rather important to the survival of the animal (Krakauer et al., 2017). Behaviors are controlled by the brain at different levels of organization, from the cellular to the network level. Emergent properties may occur at network level that cannot be predicted by studying only cellular properties, and vice versa. Thus, different levels of organization may contribute to different aspects of the behavior. Further, the brain is known to be degenerate such that different patterns of activity may control the same behavior. Thus, the brain is capable of multiple possible mechanisms, but only a subset may be behaviorally relevant for the animal. Studying behaviors whose ethological relevance is unknown may suggest mechanisms which are actually never employed in natural behaviors. On the other hand, studying natural behaviors provide a robust framework to test relevant hypothesis and understand the role of brain in driving them. In the current studies involving highly controlled experimental paradigms, it is difficult to interpret to what extent the natural behaviors factor in. As discussed before, natural movements are also more complex and can be spontaneously initiated without external cues. The mechanisms of movement initiation has not been studied in purely natural movements. The extensive research on preparatory activity and movement initiation on simple movements have made several predictions, as mentioned in the earlier sections. These provide a strong basis and can be tested in the context of natural movements, along with addressing caveats from current studies.

1.6 The zebra finch song: a natural motor skill

Studying ethologically relevant behaviors is a way forward to fill the gaps in our understanding of movement initiation and provide useful insights. To capture the neural basis of movement initiation, one needs a natural behavior that is repeatable, well-controlled and can be easily produced as part of lab assays. Many complex motor behaviors across animals have been documented for example birdsong and bat navigation, that provide both. Of these the zebra finch song has been studied extensively in the past 70 years in the context of learned movement sequences Fee and Scharff, 2010). The zebra finch (*Taeniopygia guttata*) is a species of passerine songbirds often found living in large social groups in deserts of Australia (Zann, 1996). Being opportunistic breeders and requiring a short span of 3 months to attain maturity, they have become easy animals to be maintained in the lab (Morris, 1954; Immelmann, 1969; Zann, 1996). Several characteristics of the song make it an excellent natural behavior to study in the context of movements.

Song is a stereotyped natural motor sequence

Like arm movements are produced by coordinating muscles of the arm, birdsong is produced by coordinating the muscles of the vocal organ called the syrinx in synch with respiration Nottebohm et al., 1976; Suthers and Margoliash, 2002). It is a complex movement consisting of a sequence of vocalizations or syllables. In the zebra finch, only the adult males sing and each bird has one song also called the motif. This song is a fixed sequence of syllables stereotyped across renditions. The acoustic properties of the syllables in the sequence as well as the duration of short silent-gaps between the syllables are highly precise (Immelmann, 1969; Price, 1979). Thus, a zebra finch can naturally produce multiple instances of the same song or movement without being artificially trained on a task.

Song is a learned movement

A juvenile male learns to sing a copy of its father's song during a critical learning period that lasts upto 90 days-post-hatch. Song learning involves two overlapping phases. First is the sensory phase, involving listening and memorizing a template of the father's song. The second is auditory-motor learning, involving extensive practice using auditory feedback to match the memorized template. During auditory-motor learning, spanning ~35-90 days-post-hatch, the song temporal and acoustic properties improve. The song transitions from a babble-like subsong (with unrecognizable elements) to plastic song (with some recognizable elements similar to the tutor song) and finally a stereotyped song that marks the end of crystallisation. Konishi, 1965; Immelmann, 1969; Price, 1979). Learning song by imitation and practice is very similar to how we acquire motor skills (Fee and Scharff, 2010).

Song is ethologically relevant to the bird

The zebra finch song is part of the males' natural courtship behavior to attract females (Morris, 1954). Along with other traits like bright red bills, orange cheek patch, zebra lines on the neck, song is a sexually dimorphic feature (Morris, 1954; Nottebohm and Arnold, 1976; Zann, 1996). Song learning and production are costly traits that reflect male quality Nowicki et al., 2002; Spencer et al., 2003) thus playing a role in sexual selection. The performance of song is an important component of mate choice as females can discriminate between songs (Tomaszycki and Adkins-Regan, 2005; Woolley and Doupe, 2008; Riebel, 2009). Although aspects of the song (and other traits) preferred by females varies across individuals and is an active field of research Forstmeier and Birkhead, 2004; Riebel, 2009; Wang et al., 2021). Songs are passed on from fathers to son in a natural learning process (Immelman 1969) and normally reared females have a preference for familiar songs (Williams et al., 1993; Lauay et al., 2004; Chen et al., 2017b). Taken together, learning and producing song is significant for the animal's survival and make the zebra finch song a behavior whose ethological relevance is well-known.

Neural basis of song production is well-studied

In addition to the above behavioral characteristics of song, the neural control of song has been extensively studied. Specific neural pathways have been identified in the bird that control song production and learning Fee and Scharff, 2010) (Fig. 1.2). Further, these areas have analogous counterparts in mammals as well (Brainard and Doupe, 2013). The fine motor control during song production has been attributed to two premotor areas: HVC (used as the proper name) and RA (robust nucleus of the arcopallium). Lesioning HVC abolishes the stereotyped temporal and acoustic properties of song, and lesioning RA makes the bird mute for song (Nottebohm et al., 1976). HVC projects to RA and thus RA is the output of the fine motor control of song (Nottebohm et al., 1976). RA neurons project to downstream motor areas in the brainstem namely the hypoglossal nucleus involved in controlling muscles of the syrinx, and respiratory motor areas (Ram and Pam) that control respiratory pressure through airsacs (Nottebohm et al., 1976; Suthers and Margoliash, 2002). The stereotyped song is produced in co-ordination with breathing such that syllables are produced during expirations and silent-gaps are produced during inspirations (Wild et al., 1998; Goller and Cooper, 2004). Traditionally, this pathway starting from HVC to the muscles producing song is known as the song motor pathway. Other than direct connections from HVC to RA, these areas are also connected by an alternate indirect pathway called the anterior forebrain pathway. HVC projects to Area X in the forebrain, Area X projects to DLM (medial portion of the dorsolateral thalamus) in the thalamus, and DLM projects to LMAN (lateral magnocellular nucleus of the anterior nidopallium) in the forebrain. Finally, LMAN projects back to RA. This loop consisting of HVC – Area X – DLM – LMAN – RA is called the anterior forebrain loop and areas in this loop have been implicated in song learning (Bottjer et al., 1984; Sohrabji et al., 1990). During song development, LMAN inputs help to make specific HVC to RA connections stringer, so as to produce song that is similar to the tutor song. However, in adults, the anterior forebrain pathway is dispensable for song production (Bottjer et al., 1984; Sohrabji et al., 1990) . Together, the specialized areas in the brain for song control have made zebra finch an established model system for understanding neural mechanisms of motor sequence learning and production (Fee and Scharff, 2010).

The neural activity in HVC and RA have been characterized for control during song production (Fig. 1.2). The neurons in both HVC and RA show stereotyped neural activity corresponding to the song output. Multiple neuron types in HVC show singing related activity (McCasland, 1987; Yu and Margoliash, 1996; Vu et al., 1998) and are interconnected by dense recurrent connections (Mooney and Prather, 2005). These include two types of projection neurons – HVC$_{RA}$ and HVC$_{X}$ that send their projections outside of HVC to neurons in RA and Area X respectively, and the interneurons that retain their projections inside HVC. The singing related activity of HVC$_{RA}$ neurons is observed as a single burst at a specific time during song (Hahnloser et al., 2002). Based on this activity, network models predict a feedforward chain of HVC$_{RA}$ neurons that fire in sequence and keep track

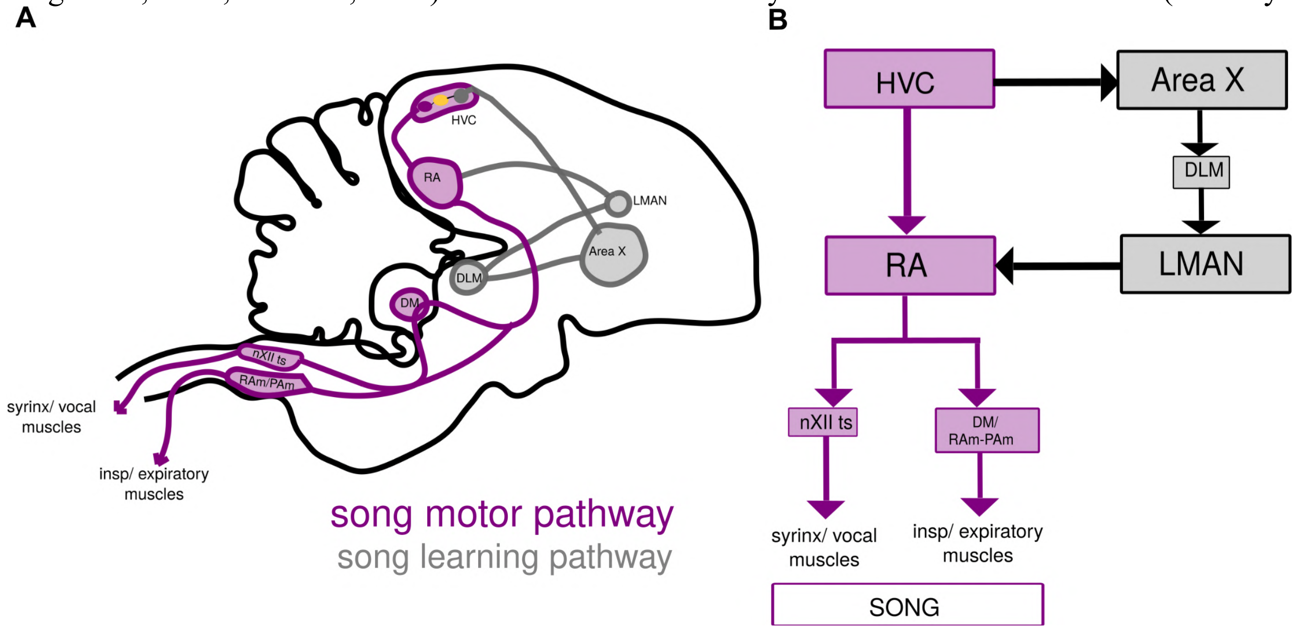

Fig. 1.2 Neural control of song by song motor pathway and song learning pathway
(A) Schematic of zebra finch brain showing areas involved in song learning and production. Three types of HVC neurons shown- HVC$_{RA}$ project to RA (magenta), HVCX project to Area X (grey) and interneurons (yellow) connect to HVC$_{RA}$ and HVC$_{X}$ but do not project outside. (B) Box diagram showing major connections between song areas. Song is produced by coordination of muscles of syrinx in synch with repiration.

of the time in the song (Leonardo and Fee, 2005; Long et al., 2010). Activity propogates downsteam from HVC_{RA} to RA neurons for motor production. The HVC_{X} neurons receive inputs from HVC_{RA} and interneurons, and can be considered as carrying a copy of the motor commands. These neurons show more than one bursts at specific times during the song. Contrary to bursting activity of projection neurons, the HVC interneurons fire throughout song with a stereotyped activity pattern (Yu and Margoliash, 1996). The interneurons receive majority of inputs from HVC_{RA} neurons and are thus considered as readouts of activity from multiple HVC_{RA} neurons. In addition to singing related neural activity of HVC neurons time-locked to the song, it was shown that localized cooling or warming of HVC alone slowed down or speeded up the song proportional to the change in temperature. As temporal properties of brain processes are temperature dependent, the temporal changes to the song resulting from localized cooling or heating of HVC infer that HVC controls the time in the song. (Long and Fee, 2008). The projection neurons in RA that receive inputs from HVC_{RA} neurons show multiple bursts again time-locked to the song Yu and Margoliash, 1996; Hahnloser et al., 2002). These multiple bursts arise due to inputs from multiple HVC_{RA} neurons. At a given time in song, an ensemble of RA neurons fire and this is thought to control the acoustic properties of vocal output (Leonardo and Fee, 2005). These findings that have widely contributed to our understanding of the fine motor control of stereotyped song, will be useful in studying song initiation.

The zebra finch song can be externally triggered or self-initiated

Other than being a well-controlled, learned and complex natural movement, the zebra finch song also has variants based on social context. The same song sequence is sung in two social contexts with minor differences. First, the song is sung in the presence of the female and is called the directed song. Second, the song is sung when alone and is called the undirected song. The directed song is marginally faster with less variability in its acoustic features compared to the undirected song (Morris, 1954; Sossinka and Böhner, 1980). Thus, the directed and undirected conditions provide natural counterparts to externally triggered and self-initiated movements respectively.

Taken together, all of these characteristics make zebra finch an excellent model that provides immense potential to understand movement initiation and preparatory activity in naturally learned complex movements.

1.7 Introductory notes delay song initiation and systematically progress to song

What do we know about song initiation? Unlike instruction based movements, initiation of song is not a quiet behavior. Birds sing in bouts and bouts begin with variable repeats of short syllables called the Introductory Notes (INs). The INs are followed by one or more renditions of the stereotyped song sequence, sometimes interleaved by a number of INs (Price, 1979; Sossinka and Böhner, 1980). The timing and acoustic properties of INs appear more variable compared to the song (Price, 1979; Rajan and Doupe, 2013), though this has not been quantified rigorously. However, the variability in properties reduces as repeats of INs approach the song (Rajan and Doupe 2013) . The interval between INs is initially longer and more variable but becomes shorter and more consistent before song. Further the acoustic properties of successive INs become more similar to that of the last IN, which is also more consistent across renditions. The transition of IN properties from variable initial state to a consistent final state is behavioral, and holds parallels to the population neural activity changes attributed to motor preparation (Shenoy et. al., 2013). In that, motor preparatory activity represented by a population of neurons started from variable initial states but converged on a consistent state at the time of the "go" cue. Owing to the similar trends, it was suggested that introductory notes before song may represent motor preparation(Rajan and Doupe, 2013).

1.8 What is known about INs and neural activity during INs?

The Introductory notes have not been extensively studied and due to its variable nature relative to the song, they have been ignored in many song related studies. However, bouts of songs are always described as beginning with variable number of INs before the song (Price, 1979; Sossinka and Böhner, 1980). In fact, the average number of INs are more during directed songs, compared to undirected songs (Sossinka and Böhner, 1980). Recent work specifically on INs make several observations on IN behavior that further support its preparatory role. These are listed below:

- Within a bout of song, multiple song renditions may be interleaved by INs. In these cases, the silence from the previous vocalization is positively correlated to the number of INs before next song (Rajan and Doupe, 2013). This could be interpreted as: more the preceding silence more the preparation time required for the next song. Thus, INs may represent degree of preparedness for song initiation.
- Similar first song syllables for pairs of renditions were shown to be positively correlated to similar last INs (Rajan and Doupe, 2013). This supports the idea that the last IN may

represent final state of preparation that influences properties of the upcoming song. This trend is analogous to levels of preparatory activity at the time of the "go" cue that influence the upcoming movement.

- Firing rate of HVC_X neurons for the last IN was significantly different from that for previous positions (Rajan and Doupe, 2013). Thus, neural activity in premotor nucleus HVC during INs may represent a distinct last state. The distinct last state is analogous to preparatory neural activity reaching an optimal state space before the movement. However, a point to note is that there is no change in the variability of neural activity, suggesting deviations from the traditional preparatory framework.
- Across trials, HVC-interneurons have similar activity for the last IN before song and the similarity to the last IN reduces for earlier IN positions (Rajan and Doupe, 2013). Neural activity in premotor area HVC may represent motor preparation similar to preparatory neural activity showing significant changes until movement onset.

Together the data on INs predict that IN progression and its neural correlate may reflect states of preparation for song that are completed before song initiation. These above observations frame various hypothesis that can be tested further for preparatory function of INs. Nonetheless, the functional role of INs is open to possibilities where preparatory role is one such possibility.

1.9 What could be the function of INs?

INs are reliably present in the beginning of the bout. Birds raised in isolation in the absence of an adult tutor also produce songs beginning with INs (Price, 1979). Moreover, INs are also present in the beginning of the bout of other songbirds (Tamura and Marler, 1962; Shiovitz, 1975; Brenowitz, 1982). These evidences suggest an important role of INs during song initiation. Apart from the suggested function for motor preparation, other suggested functions also imply an important role of INs in the beginning of the bout. Some of the speculations are discussed below:

Since INs are vocal motor outputs before song initiation, the production of INs could be a calibration process of the motor periphery. The songs produced are fast and independent of feedback in real-time (Bottjer and Arnold, 1984; Lombardino and Nottebohm, 2000; Brainard and Doupe, 2001), consistent with the motor program hypothesis. To set the vocal organ ready for the upcoming movement, it is possible that the brain utilizes feedback from IN production to improve the state of vocal muscles. Currently, we do not know whether IN production depends on peripheral feedback.

As songs play a role in communication, it is possible that INs are also a part of this communication signal. They may be used to seek the female’s attention before song or communicate some other information to other conspecifics. For both suggested roles in communication and motor preparation it is important to know whether INs are related to the upcoming song.

There may be more roles of INs, not intuitive, but may reveal themselves as we test various behavioral and neural hypotheses with our current understanding of a possible preparatory role. On the other hand, if INs are indeed movements reflecting motor preparation, it will make motor preparation all the more accessible for probing its mechanisms. Further, INs being part of a natural complex behavior may lead to untapped findings about general principles of movement initiation. These reasons make the motor preparation an important hypothesis to test for INs.

1.10 Understanding progression of INs to song important for testing motor preparation hypothesis

The progression of INs is major criteria for suggesting preparatory role for INs. Thus, a first step in understanding whether INs are preparatory activity preceding song initiation will be to understand how INs progress to song. Understanding the mechanisms behind this transition would provide insights into making predictions about what song properties may be getting prepared, if at all INs represent preparation. Further, understanding behavioral and neural correlates of IN progression will allow making predictions on how preparatory activity may evolve and eventually switch to song initiation. Such predictions arising from this study will form the basis for testing the motor preparation hypothesis for INs. Together, the reasons stated above provide strong motivation to ask the broad question of this thesis: How do introductory notes progress to song?

1.10.1 How do INs progress to song?

To address how introductory notes progresses to song, this thesis asks specific questions as summarized below:

- Does progression of INs involve real-time sensory feedback?

It is possible that INs help in establishing better control with the periphery to help in motor performance. INs could be a way for calibration of vocal muscles such that they provide feedback of the muscle state for improvement before the movement for song is executed. This hypothesis was tested by comparing IN progression before and after removal of two types of sensory feedback namely proprioceptive feedback and auditory feedback. Chapter 2 (same as (Rao et al., 2019)) describes the results of this work.

- Does progression of INs reflect correlations to properties of song?
  The properties of INs at successive position may represent aspects of songs being prepared and result in IN-song correlations. Regardless of the preparatory role, it is also possible that the relationship between properties of INs and song exists as both are part of same sequence of vocalizations. The related properties were identified by two types of correlations: (i) correlating properties of INs and song on a trial-to-trial basis and (ii) correlating age-related change in song properties to age-related changes in song. The study also for the first time characterizes the extent of change in variable properties of IN progression with age. As INs are repeating syllables, properties of repetition were also compared with other repeating song syllables in a separate set of birds. The results are described in Chapter 3.

- How do neural correlates of INs progress to song?
  Based on previous studies on INs and the results of Chapter 2 and Chapter 3 that highlighted temporal properties being unique for IN progression, interval related activity in HVC interneurons were analyzed for neural correlates of IN progression. Further position related activity related of INs at the first, middle and last position of INs were compared for HVC interneurons and multiunit extracellular recordings from RA, in order to determine representations of IN progression in the song motor pathway. The results are described in Chapter 4.

The subsequent chapters of the thesis address each of the specific hypothesis in detail, followed by a discussion on the contribution of this work in providing useful insights about Introductory notes before song and motor preparation before movement initiation.

CHAPTER 2

Sensory feedback independent pre-song vocalizations correlate with time to song initiation

2.1 INTRODUCTION

The song motif (referred to as song) of the adult male zebra finch, consisting of a stereotyped sequence of sounds (syllables) interleaved with silent gaps (Fig. 2.1), is a well-established model for understanding learned movement sequences (Fee and Scharff, 2010). How such learned movement sequences are initiated in the brain remains poorly understood. Song is learned by young birds from a conspecific tutor during a critical period (Fee and Scharff, 2010). While song is typically part of a courtship ritual for mate attraction, birds also sing when they are alone (undirected song) (Sossinka and Böhner, 1980; Zann, 1996), making this an excellent model system to study motor preparation before self-initiated, learned movement sequences.

Song is preceded by the bird repeating a short vocalization called an introductory note (IN; Fig. 2.1) (Price, 1979; Sossinka and Böhner, 1980). Each song bout consists of a variable number of such INs followed by multiple repeats of the song. It was previously shown that intervals between successive INs and the acoustic properties of successive INs progress from a variable initial state (first IN in each song bout) to a more consistent ‘ready’ state (last IN in each song bout) just before the start of each song (Rajan and Doupe, 2013). Given the similarity to the reduction in variability associated with neural preparatory activity before the onset of simple movements (Shenoy et al., 2013), INs may represent vocalizations that help prepare the zebra finch brain to produce song. However, the mechanisms underlying IN progression to song remain unclear.

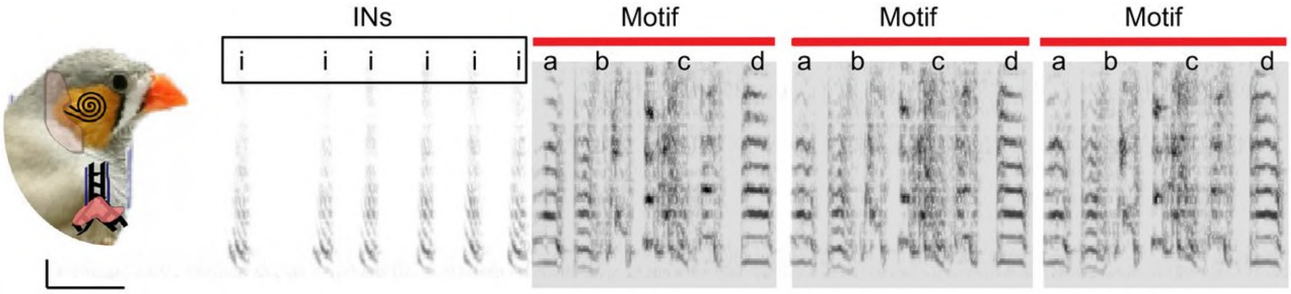

Fig. 2.1 Spectrogram of a zebra finch song bout with introductory notes (INs) and motifs. A male zebra finch showing two important sensory feedback sources, namely auditory feedback from the cochlea and proprioceptive feedback from the syringeal muscles. Spectrogram of the song of an adult male zebra finch. ‘i’ denotes INs and ‘a’ to ‘d’ represent the syllables of the song motif (gray shading). Scale bar: 200 ms (horizontal) and 1000 Hz (vertical).

One possibility is that real-time sensory feedback could drive the progression of INs to song. Sensory feedback is important for song learning and maintenance in zebra finches Konishi, 1965; Nordeen and Nordeen, 1992). In a related species, the Bengalese finch, recent work has shown that removal of auditory feedback changes the repeat number of individual syllables within song (Wittenbach et al., 2015). As INs are also repeating syllables, real-time sensory feedback could drive IN progression to song. Consistent with this hypothesis, previous studies disrupting proprioceptive feedback or auditory feedback have reported changes in the number of INs before song in some birds (Bottjer and Arnold, 1984). However, these changes have not been quantified rigorously and the specificity of these changes to removal of feedback has not been determined.

In order to assess the role of sensory feedback, I analyzed the number and properties of INs soon after removal of two important forms of sensory feedback, namely proprioceptive feedback from the syringeal muscles (Bottjer and Arnold, 1984; Vicario, 1991; Williams and McKibben, 1992) and auditory feedback (Konishi, 1965). I found that mean IN number before song and progression of INs to song were not affected by removal of either form of feedback. Further, the progression of INs to song was not affected by removal of neural input to the syringeal muscles. Finally, I found fewer INs and a quicker transition to song when the first IN was produced soon after calls (non-song vocalizations that are different from INs and song syllables). These data demonstrate that INs do not provide real-time sensory feedback during the progression of INs to song. Rather, INs may reflect internal neural processes, potentially involved in getting the zebra finch brain ‘ready’ to produce the learned song sequence.

2.2 METHODS

Experimental procedures performed at IISER Pune were approved by the Institute Animal Ethical Committee in accordance with the guidelines of the Committee for the Purpose of Control and Supervision of Experiments on Animals (CPCSEA, New Delhi). Experiments performed at UCSF (CA, USA) were approved by the UCSF Institutional Animal Care and Use Committee in accordance with NIH guidelines.

2.2.1 Birds and song recording

All birds (n=42) used in this study were >100 days post-hatch at the time of the experiment and were either purchased from an outside vendor (n=13) or bred at IISER Pune (n=16) or UCSF (n=13). Birds were kept in separate sound isolation boxes (Newtech Engineering Systems, Bangalore, India, or Acoustic Systems, Austin, TX, USA) for the duration of the experiment. All songs were recorded by placing a microphone (AKG Acoustics C417PP omnidirectional condenser microphone or B3 lavalier microphone, Countryman Associates, CA, USA) at the top of the cage. For birds in the tracheosyringeal nerve surgery (ts-cut) and sham-surgery groups (see below), I kept the position of the microphone the same for recording songs before and after surgical manipulations. Signals from the microphone were amplified using a mixer (Behringer XENYX 802) and then digitized on a computer at a sampling rate of 44,100 Hz using custom-written software. Songs were recorded in 'triggered' mode before and after surgery, such that data were saved when the microphone signal crossed a pre-set threshold. Along with the data that crossed the threshold, 1–3 s of data before and after threshold crossing were also saved. For a subset of birds, data were saved in 'continuous' mode, i.e. all the data for the entire recording period. All songs were recorded in the 'undirected' condition. Songs of three of the birds used for the analysis of calls and their influence on song initiation were also used in a previous study for analysis of INs before song (Rajan and Doupe, 2013). The influence of calls on song initiation was not considered in the previous study. For the analysis of day-to-day changes in IN number and properties, I used data from 14 birds that were recorded on two different days (range: 1–3 days apart). Of these 14 birds, one bird was also used at a later time point for ts-cut surgery with a new set of pre- and post-surgery recordings and nine birds were also used for analysis of the influence of calls on INs. Pre-surgery recordings for 18/21 birds (n=5 ts-cut, n=6 sham surgery and n=7 deaf) were performed 0–2 days before surgery. For the remaining three birds (n=3 ts-cut), pre-surgery recordings were made 18, 14 and 5 days before surgery, respectively.

2.2.2 Tracheosyringeal nerve cut and sham surgery

Tracheosyringeal nerves were surgically cut using previously described protocols (Konishi, 1965; Vicario, 1991; Williams and McKibben, 1992). Briefly, birds (n=9) were deeply anesthetized by intramuscular injection of ketamine (30 mg kg−1), xylazine (3 mg kg−1) and diazepam (7 mg kg−1). Absence of a response to toe pinch was used to assess the depth of anesthesia. Birds were then placed on a platform with the ventral side facing up. A rolled tissue under the neck served to stretch and give easy access to the throat. Feathers were plucked and an incision of ∼10 mm was made. The trachea was exposed by removal of fat tissue. Using fine forceps, the tracheosyringeal (ts) nerve bundle on either side of the trachea was pulled away from the trachea and part of the nerve (n=9 birds, median length cut 4 mm, range 2–7 mm) was cut out on both sides using spring scissors (Fine Science Tools, Foster City, CA, USA). The skin was then glued using tissue adhesive (Vetbond, 3M). For sham surgeries (n=6), the same procedure was followed but the ts nerves were not cut. The fat tissue above the ts nerves was either removed completely or only pushed to the side to observe the ts nerves before gluing back the skin. In two of the sham-surgery birds, some cuts were made on the thick membrane enclosing the esophagus. Birds typically resumed singing within 10 days of surgery. I considered songs produced on the second day of singing after surgery (sham surgery: 2–5 days and ts-cut: 3–10 days after surgery) for analysis because of the higher number of songs. For one bird, I did not have pre-surgery songs in the undirected condition, so I excluded this bird from analysis involving comparison with the pre-surgery condition (Figs 2.2-2.9). Data from this bird were included only for analysis of the influence of calls on the number and properties of INs (Figs 2.10 and 2.11). For the ts-cut group, birds with first motif syllables that were significantly different in duration from the IN were chosen; this made it easier to recognize the onset of the motif after surgery. The number of INs was not a consideration while choosing birds for the sham-surgery or ts-cut groups. For the first two birds in which I attempted ts nerve surgery, I did not observe any changes to song after surgery. While doing the other ts nerve surgeries, I realized that I had not cut the ts nerve in these two birds and so both birds were assigned to the sham-surgery group. The rest of the sham-surgery birds were chosen at random based on availability in the colony. IN number and song were not taken into consideration.

2.2.3 Deafening

Deafening was done by bilateral removal of the cochlea under equithesin anesthesia using previously described protocols (Konishi, 1965; Kojima et al., 2013). All of the deaf birds (n=7) were also used in a previous study that examined the effects of deafening on song (Kojima et al., 2013). Here, I specifically analyzed the effects of deafening on IN number and properties. As I was interested in the role of real-time auditory feedback in progression from INs to song, I analyzed IN number and properties for songs recorded 1 day post-deafening. Birds for deafening were chosen on the basis of their motif structure; the number of INs was not a consideration.

2.2.4 Data analysis

All analysis was performed using custom-written scripts in MATLAB.

2.2.4.1 Song Analysis

Audio files were segmented into syllables based on a user-defined amplitude threshold. Syllables with less than 5 ms between them were merged and syllables with a duration shorter than 10 ms were discarded. Individual syllables were given labels in a semi-automatic manner. They were first assigned labels based on a modified template-matching procedure (Glaze and Troyer, 2006) or clustering based on acoustic features calculated using Sound Analysis Pro. Clustering was done using KlustaKwik (http://klustakwik.sourceforge.net/). Labels were then manually checked for all files.

The repetitive sequence or song motif for each bird was identified. Song bouts were defined as groups of vocalizations with at least one motif syllable that were separated from other such groups by more than 2 s of silence (Sossinka and Böhner, 1980). For a subset of birds (n=7 deaf birds; n=6 birds for analysis of call-song bouts and n=7 birds for analysis of day-to-day changes in IN number and properties) with triggered recordings, a number of files did not have 2 s of silence before the first vocalization in the file. However, as these were triggered recordings, I assumed that there was silence before the start of the file too and so I considered files with >0.5 s silence at the beginning of the file as valid bouts. For a given bird, I used the same criterion before and after surgery to ensure that the criterion did not affect the results. Syllables that were produced in isolation outside of song bouts were identified as calls. All kinds of calls (distance calls, short calls and intermediate calls) (Zann, 1996) were combined together.

As described earlier (Price, 1979; Sossinka and Böhner, 1980; Rajan and Doupe, 2013), syllables that were repeated at the beginning of a bout were considered as INs. Calls were not considered as INs. Similar to a previous report (Zann, 1993), 76.2% of all the birds analyzed (n=32/42) produced only one IN type. The rest of the birds produced two IN types (n=10/42). For all the analyses described, I combined the multiple types of INs together unless specified.

For ts-cut birds, syllables and INs lost their characteristic acoustic structure and were reduced to harmonic stacks without any modulation (Fig. 2.2 A, middle). However, durations of individual syllables and INs remained the same as pre-surgery (Fig. 2.2 A, B, middle, Fig. 2.3). In these birds, syllables were labeled using cluster analysis as described above and INs and motif syllables were matched to pre-surgery INs and motif syllables by examining plots for duration versus mean frequency for all syllables.

On average, I analyzed 124 song bouts per bird (median 98 song bouts per bird; range 11–428 song bouts per bird).

2.2.4.2 Temporal and Spectral Similarity

I quantified changes in song after removal of sensory feedback using temporal and spectral similarity. Temporal similarity was calculated as the maximum of the cross-correlation function between the normalized amplitude envelopes of a pre-surgery template motif and other pre-/post-surgery motifs (n=9 randomly chosen motifs from pre-surgery and n=10 randomly chosen motifs from post-surgery) (Roy and Mooney, 2007). The template motif was proportionally stretched

±20% to account for differences in duration of the entire motif. As a measure of random temporal similarity between any two zebra finches, I calculated temporal similarity for motifs from 10 random pairs of birds (n=10 motifs each). Spectral similarity (% similarity) was calculated using Sound Analysis Pro (five motifs pre-surgery were compared with five other motifs pre-/post-surgery) (Tchernichovski et al., 2000). Random spectral similarity was measured for 10 random pairs of birds (n=5 motifs each).

2.2.4.3 Characterization of IN Progression

In each song bout, the last set of consecutive INs with inter-IN intervals <500 ms before the first motif syllable were considered for counting IN number (Sossinka and Böhner, 1980; Rajan and Doupe, 2013). All the analyses were restricted to such sequences of INs present at the beginning of each bout.

Intervals between INs were measured as the duration between the end of an IN to the start of the next IN. The first interval was the interval between the first two INs satisfying the above criteria. The last interval was measured as the interval between the last IN and the first motif syllable. As a measure of the progression of IN timing, I quantified the ratio between successive IN intervals across all IN sequences. Ratios were averaged across bouts to obtain a mean ratio for each bird. A ratio <1 indicated a speeding up of successive intervals as shown earlier (Rajan and Doupe, 2013)). The coefficient of variation (CV) was measured as the standard deviation divided by the mean.

To characterize acoustic properties of INs and their progression to song, I used the acoustic distance to the last IN and the ratio of the distance of successive INs, respectively. The acoustic distance is an inverse measure of similarity in acoustic properties between an IN and all last INs (Rajan and Doupe, 2013). I calculated four acoustic features, namely duration, log amplitude, entropy and mean frequency, for each IN using the MATLAB code for Sound Analysis Pro (http://soundanalysispro.com/matlab-library). For each day, I randomly chose 50% of the last INs as the reference. The distance of the remaining last INs and the corresponding first INs in the same bouts was measured as the Mahalanobis distance of the IN from the reference last INs in the 4-dimensional space formed by the four acoustic features. As a measure of acoustic progression of INs, I calculated the ratio of distances of successive INs from the last IN for each IN sequence at the beginning of a bout (50% of the bouts were excluded as the last INs from these bouts were chosen as the reference). A ratio <1 indicated that successive INs became closer in distance (or more similar) to the last IN, as seen in intact birds (Rajan and Doupe, 2013).

2.2.4.4 Analysis of the Influence of Calls on the Number and Properties of INs

The influence of calls on IN number and properties was analyzed in 16 normal, unmanipulated birds. Bouts where the first IN began <2000 ms after the end of a call were considered as call-song bouts. Bouts with only INs at the beginning were considered as IN song bouts. Birds with a minimum of five IN song bouts and five call-song bouts were considered for this analysis. For each bird, the mean number of INs in IN song bouts was subtracted from the number of INs in each call-song bout. This gave the deviation from mean for number of INs in each call-song bout. Since the time for onset of first IN after the call could occur at any point within 2000 ms, the normalized number of INs in call-song bouts were binned based on interval between call and IN. For each bird, the data was then binned at 100 ms resolution starting at 40 ms after the end of the call to 1940 ms after the end of the call. The first bin edge was kept at 40 ms as minimum time of interval between call and IN across all bouts taking all birds was between 40-50 ms. Across all birds, I fitted an exponential function (MATLAB fit function) to characterize the dependence of this change in IN number on time between the end of the call and the start of the first IN. Similarly, I also fitted exponential functions to the change in the interval between the first two INs and change in acoustic properties of the first IN (Fig. 2.12).

For many of the feedback-deprived birds, I did not have enough call-song bouts to carry out a similar analysis. Instead, I divided call-song bouts into two categories: (1) bouts where the first IN started <200 ms after the end of the call and (2) bouts where the first IN started >200 ms after the end of the call. A 200 ms period was chosen based on the exponential fit (Fig. 2.10B) and data availability in the feedback-deprived birds. I calculated mean IN number, mean and variability of the interval between the first two INs and acoustic distance of the first IN for both these bout categories and compared them with the corresponding properties for IN song bouts (Figs 2.10 and 2.11). Only birds with >3 call-song bouts in both of these categories were considered for analysis. Further, I combined data for ts-cut and deaf birds as the previous results showed that both manipulations had no effect on IN number and properties.

2.2.4.5 Statistics

I did not perform any power calculations to determine sample sizes for each group. However, sample sizes are comparable with other studies. As detailed below , I used non-parametric tests for most of the statistical comparisons because of small sample sizes (<10). I used parametric tests repeated-measures one-way ANOVA and repeated-measures two-way ANOVA only for analysis of the effect of calls on the number and properties of INs and the analysis of progression of IN features after feedback removal, respectively. Birds were excluded from the analysis only if there were no pre-surgery undirected song recordings or if there were too few song bouts (number specified in earlier sections). Both of these conditions were established before the start of the analysis. There were no blind investigations done for both the choice of animals in each group and the analysis of data. However, IN number or other properties were not considered while choosing birds as described in an earlier section.

Wilcoxon signed-rank test was used for paired comparisons of temporal similarity (Fig. 2.2B), spectral similarity (Fig. 2.2C), changes in IN/motif syllable acoustic features (Fig. 2.3; Fig. 2.4), mean IN number (Fig. 2.6D), IN number CV (Fig. 2.6E) and progression in IN features (Fig. 2.8). For comparing progression in IN timing and IN acoustic structure after removal of feedback (Figs 2.7D,E and 2.9D), I used repeated-measures two-way ANOVA using IN position (first versus last) as one factor and time (pre-surgery versus post-surgery) as the second factor (MATLAB code from https://in.mathworks.com/matlabcentral/fileexchange/6874-two-way-repeated-measures-anova).
For comparing IN number and properties in bouts where calls preceded the first IN, I used repeated-measures one-way ANOVA (Figs 2.10C,D and 2.11). If the ANOVA p-value was <0.05, I used a post hoc Tukey–Kramer test to identify groups that were significantly different (Figs 2.10C,D and 2.11). For comparing changes in IN number and properties after surgery with day-to-day changes, I used Kruskal–Wallis test (Fig. 2.5). Pearson's correlation coefficient was used to assess the correlation between first IN properties and time to song initiation (Fig. 2.13).

All of the tests used and the associated p-values are provided in Table S1 of Appendix. A significance level of $p=0.05$ was used throughout.

2.3 RESULTS

To test whether sensory feedback plays a role in song initiation, I analyzed the number and progression of INs after removal of either proprioceptive (n=8 birds) or auditory feedback (n=7 birds). As a control, sham surgeries were performed in a separate group of birds (n=6). As I was interested in self-initiated movement sequences, I only analyzed undirected songs produced when the bird was alone.

2.3.1 Song spectral structure alone is affected after ts nerve surgery

Proprioceptive feedback was removed by bilaterally cutting the ts nerve (Bottjer and Arnold, 1984; Vicario, 1991; Williams and McKibben, 1992) (ts-cut; n=8 birds, see METHODS) and auditory feedback was eliminated by bilateral removal of the cochlea (Konishi, 1965) (deaf; n=7 birds, see METHODS).

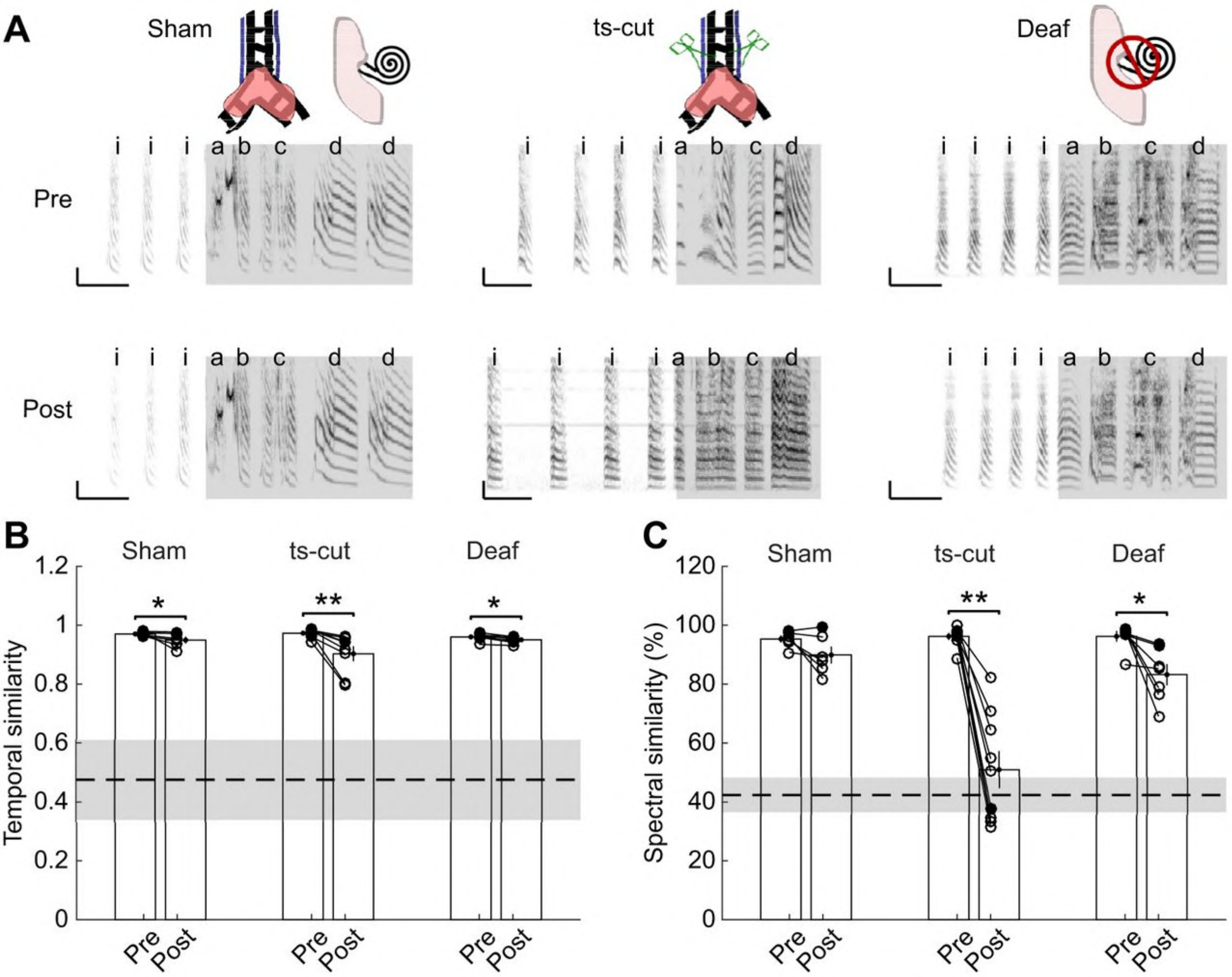

Fig. 2.2 Song structure after sham surgery, tracheosyringeal nerve cut or deafening. (A) Representative spectrograms of song for individual birds before and after sham surgery (left), tracheosyringeal (ts) nerve surgery (middle) or deafening (right). Scale bar: 200 ms (horizontal) and 1000 Hz (vertical). 'i' denotes INs and 'a' to 'd' represent song motif syllables. (B,C) Temporal similarity (B) and spectral similarity (C) to pre-surgery song for songs produced before and after sham surgery (n=6 birds), ts nerve surgery (n=8 birds) or deafening (n=7 birds). Each circle represents one bird and lines connect data from the same bird before and after surgery. Bars and whiskers represent means±s.e.m. across birds. Dashed lines and shading represent means and 95% confidence intervals for similarity between random birds. Filled circles represent values for the birds shown in A. *p<0.05, **p<0.01, Wilcoxon signed-rank test.

Despite small changes, songs of birds subjected to sham surgery remained more similar to pre-surgery songs than expected by chance in both spectral and temporal structure (Fig. 2.2). As described earlier (Konishi, 1965; Price, 1979), only minor changes to song characteristics (both temporal and spectral) were seen after deafening. Songs post-surgery remained more similar than expected by chance to pre-surgery songs (Fig. 2.2). The ts nerve contains both efferent and afferent nerves, carrying motor input to the syringeal muscles and proprioceptive feedback from the syringeal muscles, respectively (Bottjer and Arnold, 1984). Nerve cuts disrupted both efferent and afferent nerves, resulting in the loss of song spectral structure immediately after nerve cut surgery (Fig. 2.2A,C) (Vicario, 1991; Williams and McKibben, 1992; Roy and Mooney, 2007). However, as described earlier, song temporal structure remained more similar than expected by chance to that before ts nerve surgery (Fig. 2.2A middle and Fig. 2.2B) as motor input to the respiratory muscles was not affected (Bottjer and Arnold, 1984; Vicario, 1991; Williams and McKibben, 1992; Roy and Mooney, 2007). Thus, consistent with earlier studies (Bottjer and Arnold, 1984; Vicario, 1991; Williams and McKibben, 1992; Roy and Mooney, 2007; Vallentin and Long, 2015), I also report similar effects on song when both proprioceptive and auditory feedback were altered by cutting ts nerve, and when auditory feedback was disrupted by deafening.

2.3.2 IN acoustic structure, not duration, is affected by ts-cut surgery

I next quantified changes to the acoustic structure of INs after removal of either proprioceptive or auditory feedback. Similar to changes in song syllable structure (Fig. 2.2 A, middle column), INs also became harmonic stacks after surgery in ts-cut birds as seen by increased goodness of pitch and decreased frequency modulation (Fig. 2.2 A middle, Fig. 2.3 E, F, $p<0.05$, Wilcoxon signed-rank test). Despite this change, I could identify INs because IN duration, mean frequency, entropy and amplitude did not change significantly after surgery (Fig. 2.3A–D ts-cut; see METHODS for details of IN identification procedure in ts-cut birds). The position of INs at the beginning of the bout was also maintained (Fig. 2.2 A middle). Post-deafening, INs were softer and had reduced mean frequency (Fig. 2.3 B,D, $p<0.05$, Wilcoxon signed-rank test). However, song syllables were also softer after deafening (Fig. 2.4, right column, $p<0.05$, Wilcoxon signed-rank test), suggesting that these changes could have been the result of a change in microphone position after surgery. No significant changes in IN acoustic structure were seen after sham surgery (Fig. 2.3, $p>0.05$, Wilcoxon signed-rank test).

2.3.3 Mean IN number before song is not affected by removal of proprioceptive or auditory feedback

I next analyzed the mean and variability of IN number before each song (Fig. 2.5 A–C; see METHODS). The mean number of INs before song (Fig. 2.5 D) and the variability in IN number (measured by the CV – Fig. 2.5 E) did not change significantly soon after surgery in sham-surgery, ts-cut and deaf birds (Fig. 2.5 D and E, $p>0.05$, Wilcoxon signed-rank test). In fact, changes in

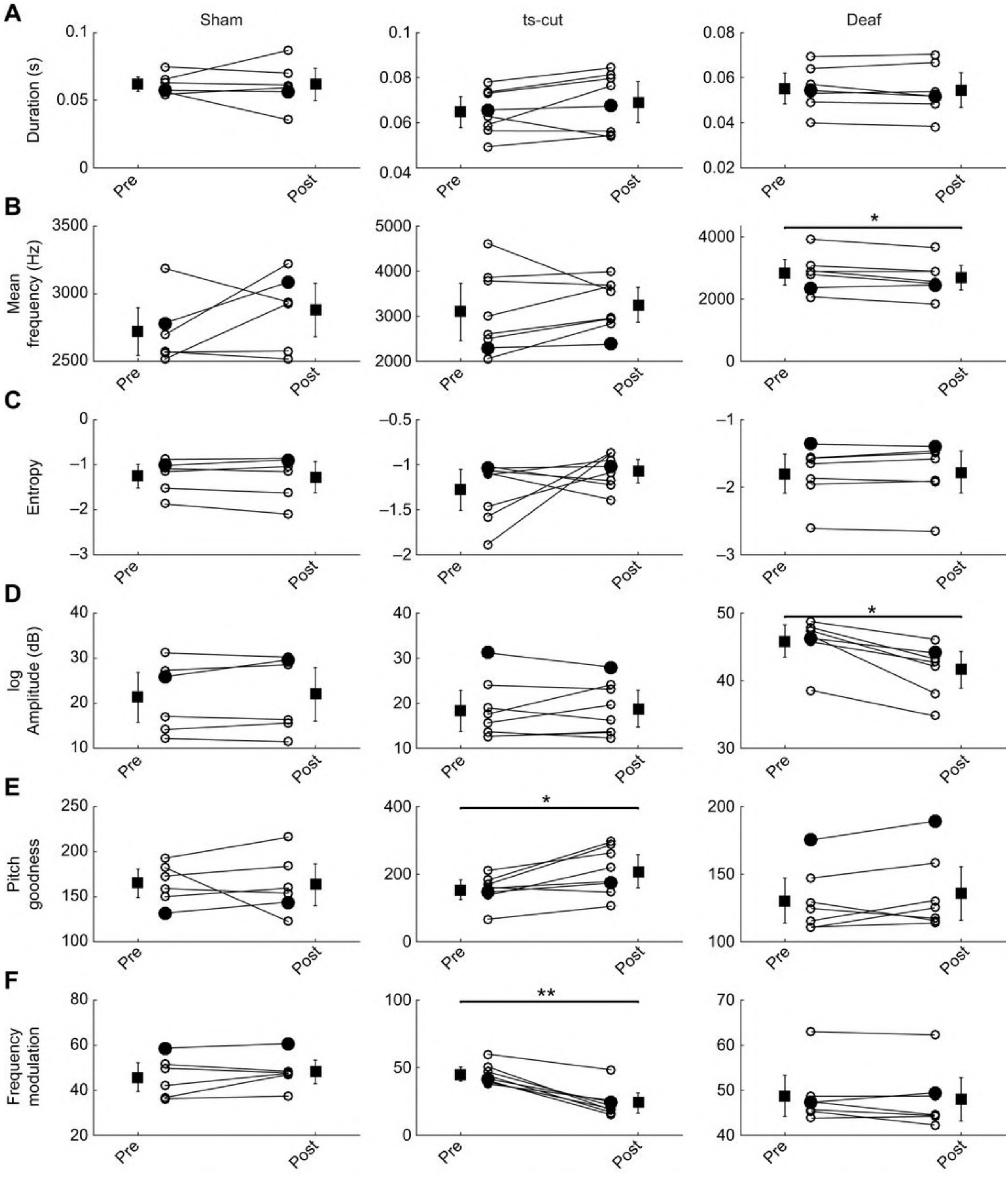

Fig. 2.3 Changes in IN acoustic features after removal of sensory feedback.(A–F) Acoustic properties of INs before and after sham surgery (left column, n=6 birds), ts nerve surgery (middle column, n=8 birds) or deafening (right column,n=7 birds). Circles represent individual birds and lines connect data from the same bird. Squares and whiskers represent means±s.e.m. across birds. Acoustic features plotted are duration (A), mean frequency (B), entropy (C), log amplitude (D), goodness of pitch (E) and frequency modulation (degree) (F). The filled circles in the columns from left to right represent the values for the birds shown in Fig. 2.2A–C, respectively. *p<0.05, **p<0.01, Wilcoxon signed-rank test.

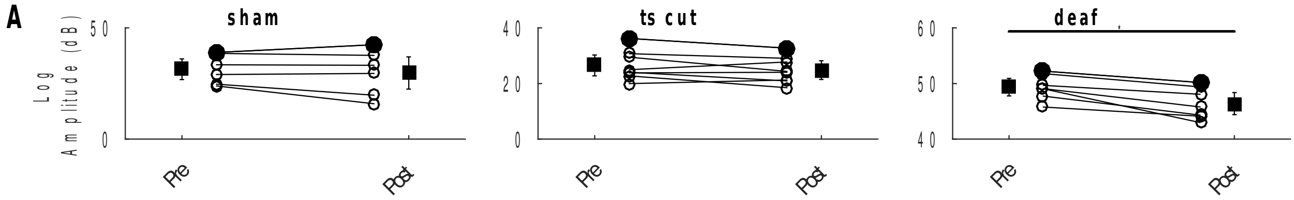

Fig. 2.4 Changes in motif syllable log amplitude after removal of sensory feedback. (A) Log amplitude of motif syllables pre and post sham surgery (n=6 birds, left column), ts-cut surgery (n=8 birds, middle column) or deafening (n=7 birds, right column). Circles represent average across all motif syllables for individual birds and lines connect data from the same bird. Squares and whiskers represent mean ± s.e.m. across birds. The filled circles in the first, second and third columns (left to right) represent the values for the birds shown in Figure 2,2 A-C respectively. * represents $p<0.05$, Wilcoxon sign-rank test.

mean IN number post-surgery for feedback-deprived birds were not different from day-to-day changes in IN number seen in normal, unmanipulated birds (Fig. 2.6 A). This further strengthened the conclusion that IN number was unaffected by removal of proprioceptive or auditory feedback.

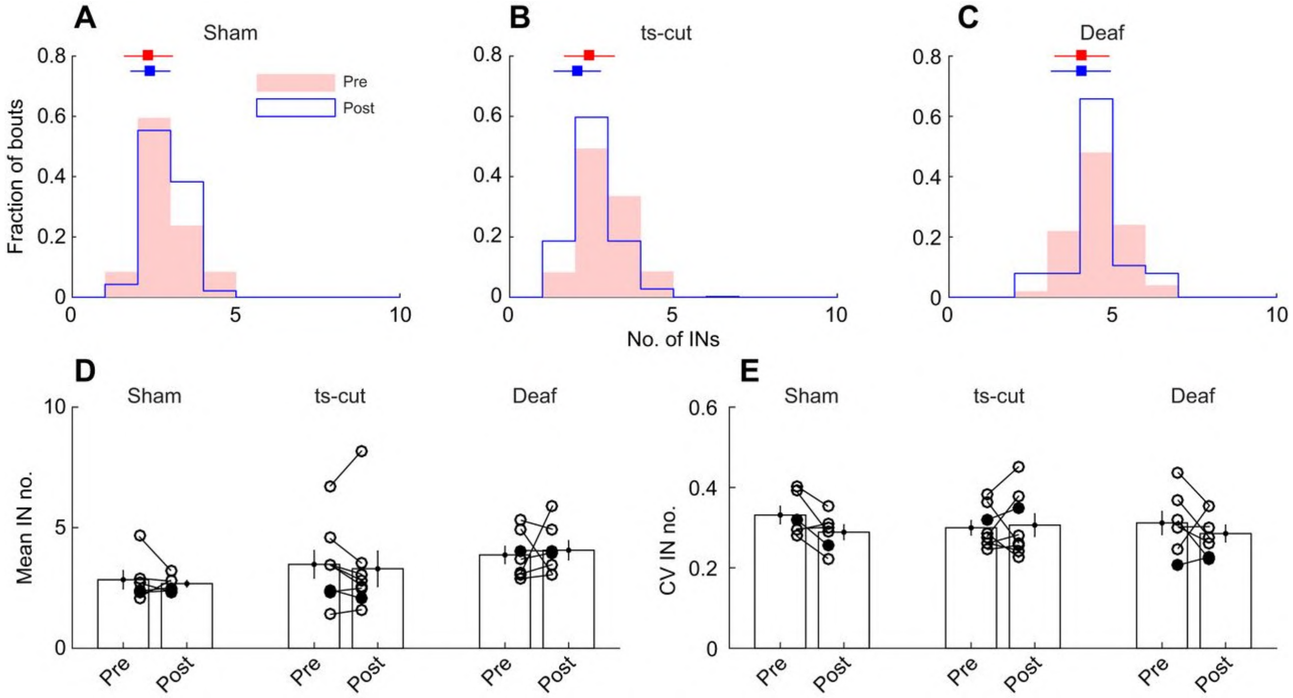

Fig. 2.5 Mean IN number and variability are not affected by removal of sensory feedback. (A–C) Distribution of number of INs for a representative bird before (red shaded) and after (blue outline) sham surgery (A), ts nerve surgery (B) and deafening (C). Squares and whiskers represent means and s.d. of the distributions before (red) and after surgery (blue). (D,E) Mean IN number (D) and coefficient of variation (CV) of IN number (E) before and after surgery for all sham surgery (n=6 birds), ts-cut (n=8 birds) and deaf birds (n=7 birds). Circles represent individual birds and lines connect data from the same bird. Bars and whiskers represent means±s.e.m. across birds. Filled circles represent the birds shown in A–C. $p>0.05$, Wilcoxon signed-rank test between pre- and post-surgery.

2.3.4 Progression of IN timing to song is not affected by removal of proprioceptive or auditory feedback

It was previously shown that progression of INs to song is accompanied by changes in both the timing and acoustic structure of INs within a bout (Rajan and Doupe, 2013). I first considered IN timing. Specifically, intervals between successive INs progress from a longer, more variable first interval to a shorter, more stereotyped interval between the last IN and song (Rajan and Doupe, 2013). This progression in IN timing was unchanged after surgery in sham-surgery, ts-cut and deaf birds (Fig. 2.7 A–C). After surgery, the interval between the first two INs remained longer and more variable than the interval between the last IN and song in ts-cut, deaf and sham-surgery birds (Fig. 2.7 D,E; $p<0.05$ for first versus last, repeated-measures two-way ANOVA). Importantly, removal of feedback did not alter either the mean or variability of both the first interval and the last interval (Fig. 2.7 D,E; $p>0.05$ for pre- versus post-surgery, repeated-measures two-way ANOVA). A number of other aspects of IN timing were also not affected by removal of auditory or proprioceptive feedback and changes in IN timing post-surgery were similar to day-to-day changes seen in

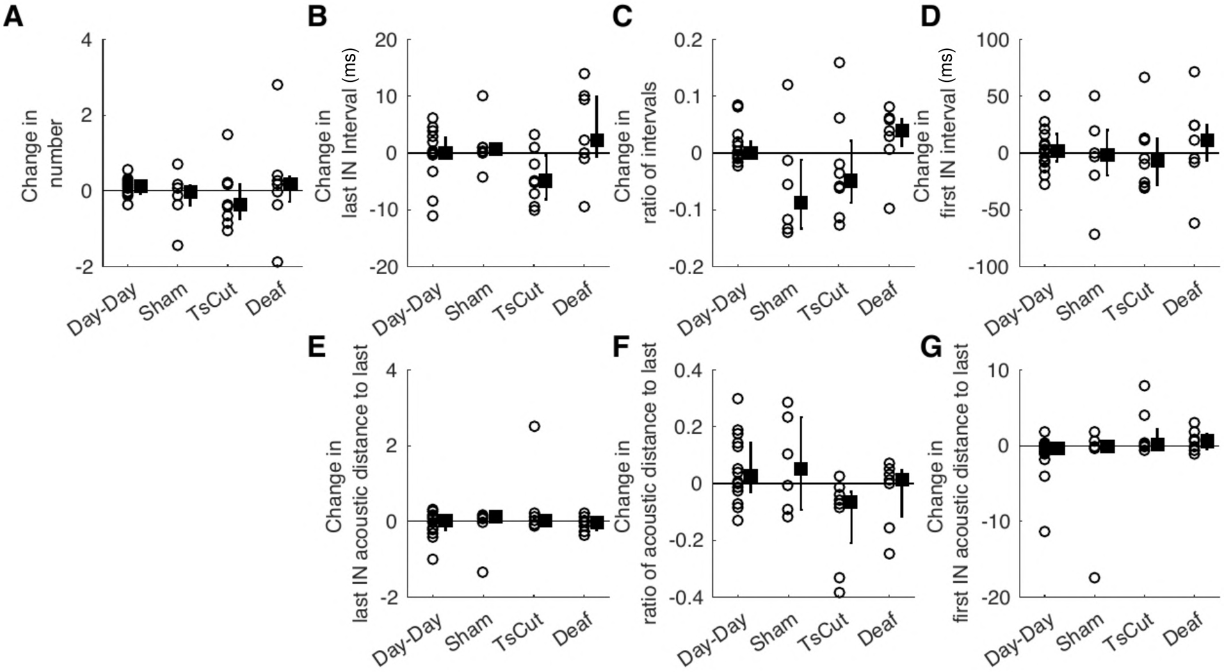

Fig. 2.6 Comparison of day-to-day changes in IN properties with changes observed after surgery for sham-surgery, ts-cut and deaf birds. (A), (B), (C), (D), (E), (F) and (G) Comparison of change in IN properties before and after shamsurgery (n=6 birds), ts-cut surgery (n=8 birds) or deafening (n=7 birds) with day-to-day changes in IN properties in n=14 birds. Changes in IN number (A), Last IN interval (B), Ratio of successive intervals (C), first IN interval (D), last IN acoustic distance to last IN (E), ratio of acoustic distances of successive INs to the last IN (F) and first IN acoustic distance to the last IN (G) are plotted. Circles represent individual birds. Squares and whiskers represent median and inter-quartile range across all birds. (A)-(G) $p>0.05$, Kruskal-Wallis test.

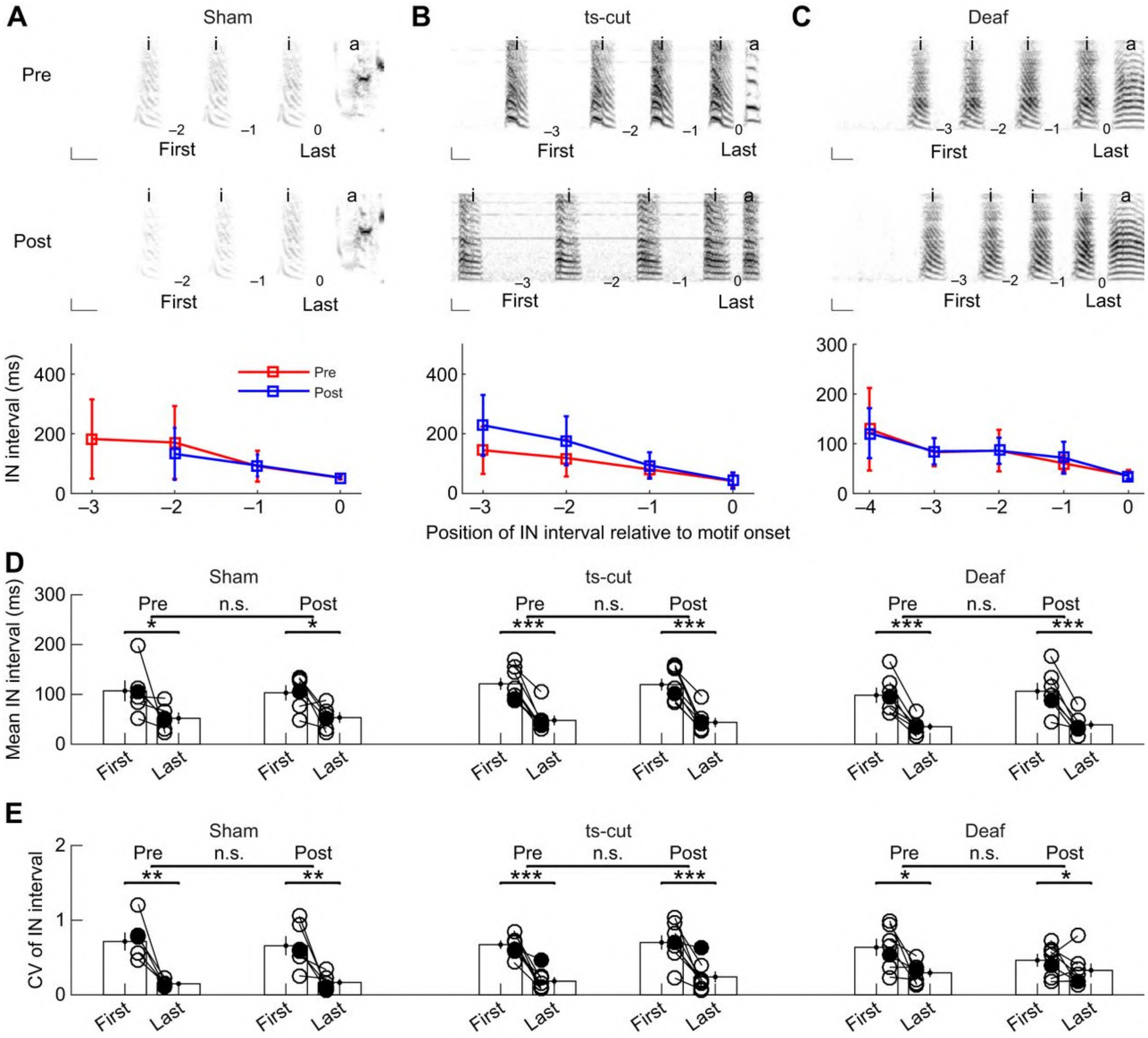

Fig. 2.7 Progression of intervals between successive INs is independent of sensory feedback. (A–C) Top: spectrograms of a sequence of INs before the first motif syllable before and after surgery. The position of each IN interval relative to the motif onset and the first and last interval is marked. Bottom: interval between successive INs versus position of the interval relative to motif onset for three example birds before (red) and after (blue) sham surgery (A), ts nerve surgery (B) or deafening (C). Squares and whiskers represent mean and s.d., n>5 bouts. Scale bars for A–C: 50 ms (horizontal) and 1000 Hz (vertical). (D) Mean interval between the first two INs in a bout and mean interval between the last IN and song across all birds before and after sham surgery (left, n=6 birds), ts nerve surgery (middle, n=8 birds) or deafening (right, n=7 birds). Circles represent individual birds and lines connect data from the same bird before and after surgery. Bars and whiskers represent means±s.e.m. across all birds. Filled circles represent the birds shown in A–C. *p<0.05, **p<0.01, ***p<0.001, repeated-measures two-way ANOVA. (E) CV of the interval between the first two INs in a bout and the interval between the last IN and song across all birds before and after sham surgery (left, n=6 birds), ts nerve surgery (middle, n=8 birds) or deafening (right, n=7 birds). Circles represent individual birds and lines connect data from the same bird before and after surgery. Bars and whiskers represent means±s.e.m. across all birds. Filled circles represent the birds shown in A–C. *p<0.05, **p<0.01, ***p<0.001, n.s. represents not significant (P>0.05), repeated-measures two-way ANOVA.

unmanipulated birds (Figs 2.6 B–D and 2.8 A – see METHODS). These results showed that the timing of INs and their progression did not depend on intact sensory feedback.

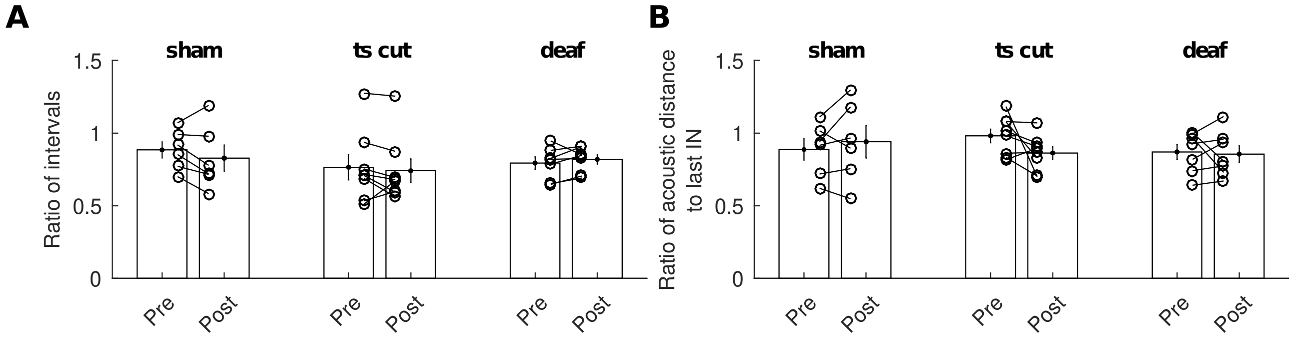

Fig. 2.8 Progression of IN timing and acoustic features is not altered by removal of auditory or proprioceptive feedback. (A), (B) Ratio of successive inter-IN intervals (A) and ratio of acoustic distances of successive INs from the last IN (B) for all birds before and after sham-surgery (n=6 birds), ts-cut surgery (n=8 birds) or deafening (n=7 birds). Each circle represents data from an individual bird and lines connect data from the same bird. Bars and whiskers represent mean ± s.e.m. across birds. $p>0.05$, Wilcoxon signrank test between pre and post surgery.

2.3.5 Progression of IN acoustic features to song is not affected by removal of proprioceptive or auditory feedback

Similar to IN timing, the acoustic structure of INs has also been shown to progress to a consistent last-IN state just before song (Rajan and Doupe, 2013). Although individual INs in each bout looked very similar (Fig. 2.9 A–C top), it was previously shown that the first IN is less similar to the last IN across bouts (Rajan and Doupe, 2013). I quantified this by calculating the similarity between the first IN and the last IN before and after surgery (acoustic distance to the last IN: the smaller the distance, the greater the similarity and vice versa; see METHODS; see Fig. 2.9A–C for representative examples for sham-surgery, ts-cut and deaf birds). As I was interested in the progression, I calculated similarity to the last IN on the same day (pre-surgery last IN for pre-surgery and post-surgery last IN for post-surgery; see METHODS). For each day, half of the last INs across all bouts were randomly chosen as a reference. The rest of the last INs and all the first INs were then compared with this reference using the acoustic distance as an inverse measure of similarity (see METHODS). The first IN was significantly different from the last IN (larger distance – Fig. 2.9 D) before and after surgery in deaf birds ($p<0.05$ for first versus last IN, repeated-measures two-way ANOVA). The average magnitude across birds for first IN was more distant than the last IN in ts cut birds for both before and after surgery, although not statistically significant ($p=0.0544$ for first versus last IN, repeated-measures two-way ANOVA). This difference was also small in sham-surgery birds both before and after surgery and did not reach significance (Fig. 2.9 D,

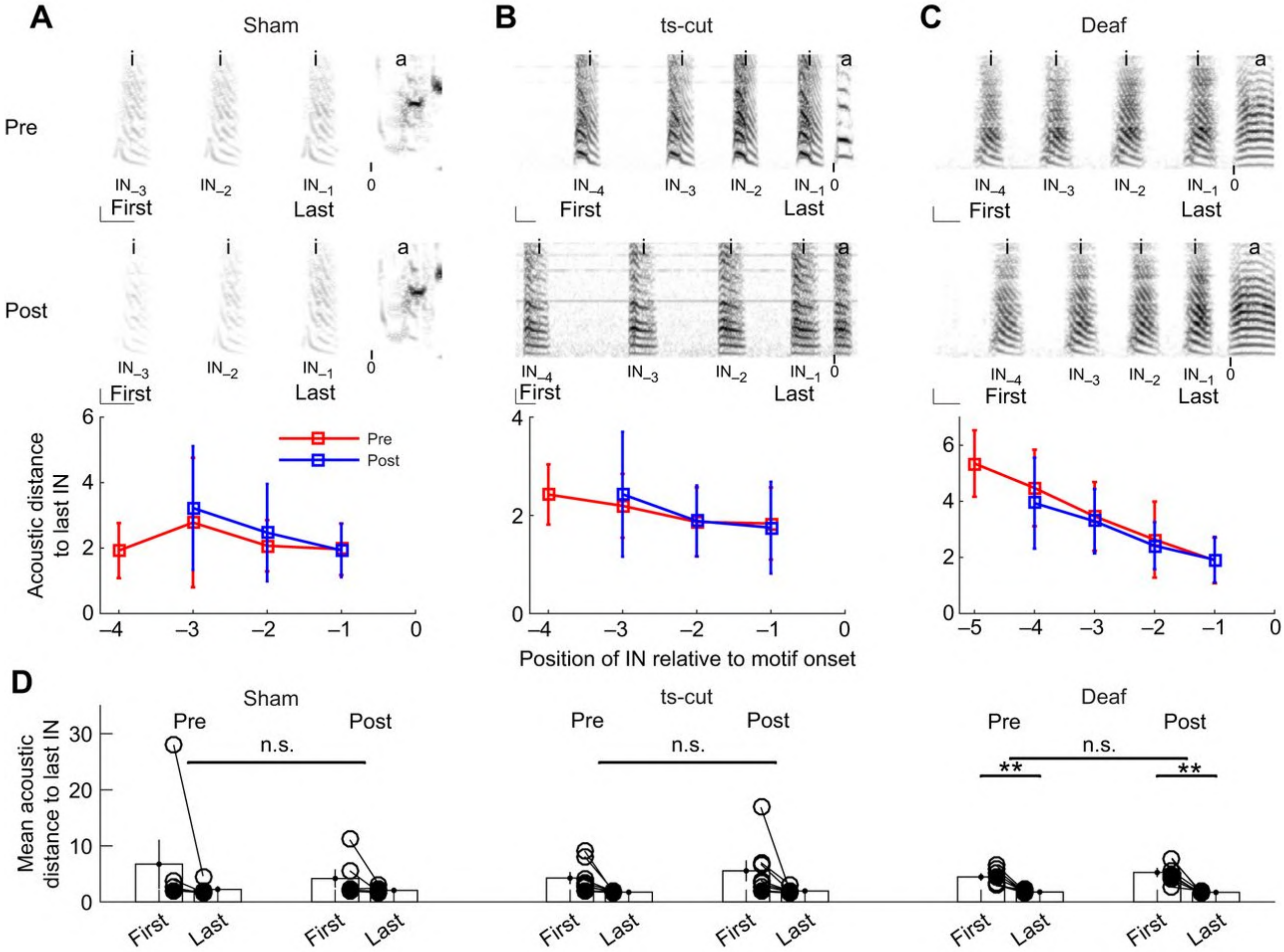

Fig. 2.9 Progression of acoustic properties of successive INs is independent of sensory feedback. (A–C) Top: spectrograms of a sequence of INs before the first motif syllable before and after surgery. The position of each IN relative to motif onset, and the first and last INs are marked. Bottom: acoustic distance of successive INs from the last IN for three example birds before (red) and after (blue) sham surgery (A), ts nerve surgery (B) or deafening (C). Squares and whiskers represent means and s.d., n>5 bouts. Scale bars for A–C: 50 ms (horizontal) and 1000 Hz (vertical). (D) Mean acoustic distance for the first IN and last IN in a bout across all birds before and after sham surgery (left, n=6 birds), ts nerve surgery (middle, n=8 birds) or deafening (right, n=7 birds). Circles represent individual birds and lines connect data from the same bird before and after surgery. Bars and whiskers represent means and s.e.m. across all birds. Filled circles represent the birds shown in A–C. **p<0.01, n.s. represents not significant (p>0.05), repeated-measures two-way ANOVA.

p=0.2469 for first versus last IN in sham-surgery birds, repeated-measures two-way ANOVA). Importantly, in all groups of birds, removal of feedback did not affect any of the measures of progression (p>0.05, pre- versus post-surgery, repeated-measures two-way ANOVA). These results showed that INs still progressed from a first IN that was significantly different from the last IN to a more consistent last IN even in the absence of auditory or proprioceptive feedback. A number of other aspects of IN acoustic structure progression to song were also not affected by removal of auditory or proprioceptive feedback and remained similar to day-to-day changes seen in unmanipulated birds (Figs 2.6 E–G and 2.8 B – see METHODS). As mentioned earlier, ts-cut birds lacked neural input to the syringeal muscles in addition to the loss of proprioceptive feedback from the syringeal muscles. The continued progression of IN acoustic features suggested that this progression is a result of changing respiratory drive, as neural input to the respiratory muscles remained intact in these birds.

Overall, these results show that IN number and progression are not dependent on intact sensory feedback (auditory and proprioceptive). This suggested that IN progression to song is controlled by internal neural processes.

2.3.6 IN number is reduced when calls precede the first IN of a song bout

If IN progression is controlled by internal neural processes, I predicted that any behavior associated with change in levels of neural activity just before the first IN will influence the number of INs before song. Calls are other non-song vocalizations produced by the bird outside song. These are partially learned or unlearned vocalizations that are acoustically distinct from song and are initiated by separate neural pathways (Simpson and Vicario, 1990; Zann, 1996; Vicario, 2004). Many aspects of calls are controlled by song motor nuclei and increased neural activity is seen in many of the song motor nuclei before and during calls (Simpson and Vicario, 1990; Yu and Margoliash, 1996; Hahnloser et al., 2002; Kozhevnikov and Fee, 2007; Long and Fee, 2008; Benichov et al., 2015; Vyssotski et al., 2016; Danish et al., 2017; Rajan, 2018). Further, it was previously shown that the presence of higher levels of preparatory activity in the premotor nucleus HVC before the first IN when calls precede the first IN of a song bout (Rajan, 2018) . Given the changes in neural activity when calls are present before the first IN, I asked whether the presence of calls just before the first IN influenced the number of INs before song. I tested this in a separate set of unmanipulated birds (n=16) by examining the number of INs in song bouts where calls preceded the first IN (call-song bouts – see METHODS. Calls occurred before the first IN in a small fraction of bouts and the time between the call and the first IN was variable (Fig. 2.10 A; mean±s.e.m. interval between end of call and start of first IN: 468.6±48 ms, mean±s.e.m. CV of interval between end of call and start of first IN: 0.85±0.07; n=16 birds). Compared to bouts that directly begin with INs, I observed fewer INs in bouts when calls occurred before the first IN (n=16 birds; mean±s.e.m. for IN song bouts: 3.7±0.24, for call song bouts: 3.4±0.25, p=0.03, Wilcoxon signed-rank test). This reduction was dependent on the time between the end of the call and the start of the IN: the shorter the time, the greater the reduction (Fig. 2.10 B, adjusted R^2=0.31 for an exponential fit, see METHODS). Thus, grouping song bouts where the first IN began <200 ms after the end of a call had fewer INs when compared with song bouts with only INs at the beginning or song bouts where the first IN began

>200 ms after the end of the call (Fig. 2.10 C, $p<0.05$, repeated-measures one-way ANOVA and post hoc Tukey–Kramer test).

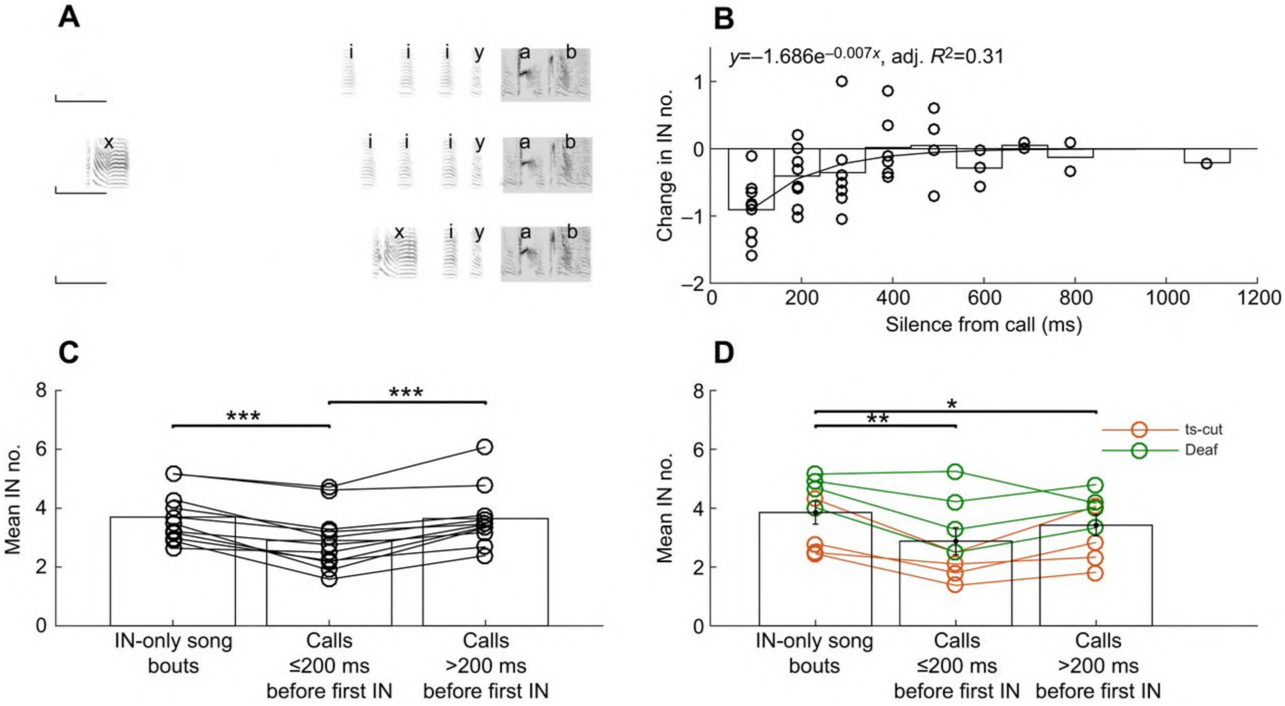

Fig. 2.10 Calls just before the first IN correlate with fewer INs at the start of the song bout. (A) Spectrograms of the start of an IN song bout (top) and two call song bouts with the call occurring well before the first IN (middle) or just before the first IN (bottom). 'i' and 'y' represent INs, 'a' and 'b' represent motif syllables and 'x' represents a call. Scale bars: 200 ms (horizontal) and 1000 Hz (vertical). (B) Silence between the end of the call and the beginning of the first IN versus change in IN number relative to the mean IN number in IN song bouts (n=16 birds). Each circle represents one bird. Bars represent means across birds and the line represents an exponential fit to the data ($y=-1.686e^{-0.007x}$, adjusted $R^2=0.31$). (C,D) Mean IN number in IN song bouts, call bouts with calls occurring ≤200 ms before the first IN and call bouts with calls occurring >200 ms before the first IN for unmanipulated, feedback-intact birds (n=11 birds) (C), and for feedback-deprived birds (D; orange – post-ts-cut, n=4 birds; green – post-deafening, n=4 birds). Circles represent individual birds and lines connect data from the same bird. Bars and whiskers represent means±s.e.m. across birds. *$p<0.05$, **$p<0.01$, ***$p<0.001$, repeated-measures one-way ANOVA.

Further, this was also true in feedback-deprived birds (Fig. 2.10 D, $p<0.05$, repeated-measures one-way ANOVA and post hoc Tukey–Kramer test) indicating that the influence by calls is independent of sensory feedback. These results showed that the presence of calls just before the first IN of a song bout correlated with fewer INs in both feedback-intact and feedback-deprived birds and further strengthened the conclusion that IN progression to song may be controlled by internal neural processes.

2.3.7 Calls just before the first IN of a song bout correlate with altered 'initial' state

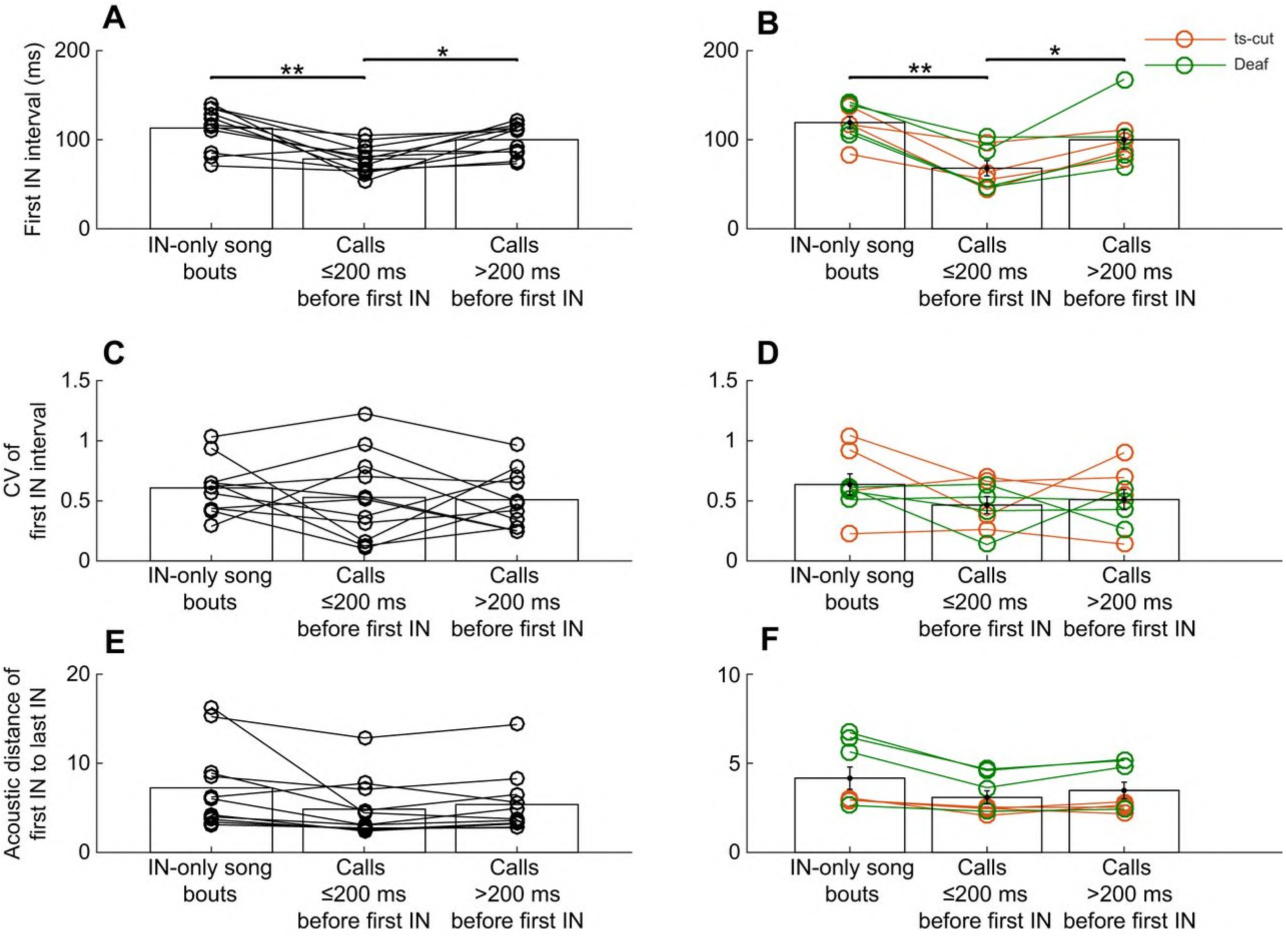

Fig. 2.11 Calls just before the first IN correlate with a shorter interval between the first two INs. (A,B) Mean duration of the interval between the first two INs in IN song bouts, call bouts with calls occurring ≤200 ms before the first IN and call bouts with calls occurring >200 ms before the first IN for unmanipulated, feedback-intact birds (A; n=11 birds) and for feedback-deprived birds (B; orange: post-ts-cut, n=4 birds; green: post-deafening, n=4 birds). Circles represent individual birds and lines connect data from the same bird. Bars and whiskers represent means±s.e.m. across birds. (C,D) Variability of the interval between the first two INs in IN song bouts, call bouts with calls occurring ≤200 ms before the first IN and call bouts with calls occurring >200 ms before the first IN for unmanipulated, feedback-intact birds (n=11 birds) (C) and for feedback-deprived birds (D; orange: post ts-cut, n=4 birds; green: post-deafening, n=4 birds). Circles represent individual birds and lines connect data from the same bird. Bars and whiskers represent means±s.e.m. across birds. (E,F) Mean acoustic distance of the first IN from the last IN in IN song bouts, call bouts with calls occurring ≤200 ms before the first IN and call bouts with calls occurring >200 ms before the first IN for unmanipulated, feedback-intact birds (n=11 birds) (E) and for feedback-deprived birds (F; orange: post-ts-cut, n=4 birds; green: post-deafening, n=4 birds). Circles represent individual birds and lines connect data from the same bird. Bars and whiskers represent means±s.e.m. across birds. *p<0.05, **p<0.01, repeated-measures one-way ANOVA.

Given that both IN timing and acoustic features progress towards a consistent ‘ready’ state just before song, I hypothesized that calls might reduce IN number by speeding up this progression. Consistent with this idea, song bouts where the first IN began <200 ms after the end of call had a significantly shorter interval between the first two INs when compared with song bouts with only INs or song bouts where the first IN began >200 ms after the end of the call (Fig. 2.11A,B; $p<0.05$, repeated-measures one-way ANOVA followed by post hoc Tukey–Kramer test). This was true both in feedback-intact (Fig. 2.11 A) and feedback-deprived birds (Fig. 2.11 B). In feedback-intact birds, the decrease in interval between the first two INs in bouts with calls was correlated with the time between the end of the call and the start of the first IN, though the strength of the correlation was moderate (Fig. 2.12 A, $R^2=0.19$). In contrast to the changes in IN timing, neither the variability of the interval between the first two INs nor the acoustic structure of the first IN showed any differences based on whether calls were present before the first IN or not (feedback-intact birds: Fig. 2.11C,E; feedback-deprived birds: Fig. 2.11 D,F; $p>0.05$, repeated-measures one-way ANOVA). However, in feedback-intact birds, relative to bouts with only INs, the acoustic structure of the first IN after a call was more similar to that of the last IN (Fig. 2.12 B). The change in acoustic structure was correlated with the time between the end of the call and the start of the first IN, but the strength of the correlation was weak (Fig. 2.12 B, $R^2=0.1$). Overall, these results showed that the presence of calls correlated with a change in IN timing (shorter interval between the first two INs), potentially causing the reduction in IN number before song.

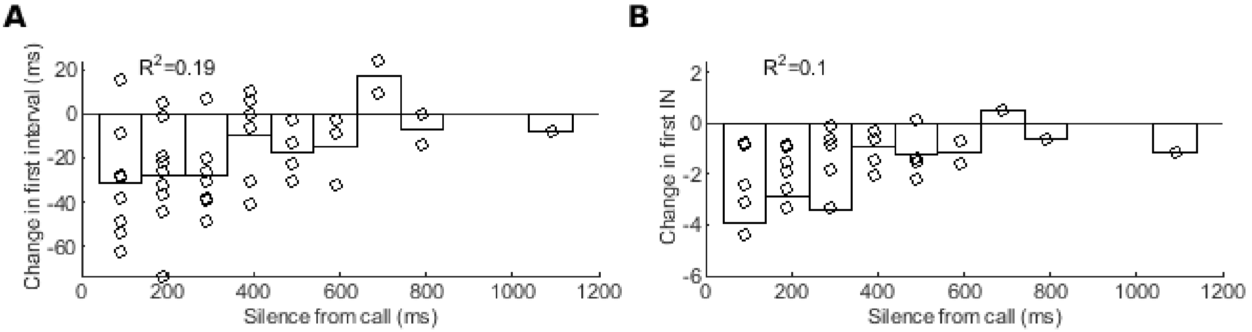

Fig. 2.12 Changes in IN "initial" state in bouts with calls before the first IN also depend on the time between the end of the call and the start of the first IN. (A) Silence between the end of the call and the beginning of the first IN vs. change in interval between the first two INs in call song bouts relative to the interval between the first two INs in IN song bouts. Each circle represents one bird. Bars represent mean across birds (n=16 birds) and the R^2 represents the strength of correlation. (B) Silence between the end of the call and the beginning of the first IN vs. change in acoustic distance of the first IN to the last IN in call song bouts relative to acoustic distance of the first IN from the last IN in IN song bouts. Each circle represents one bird. Bars represent mean across birds (n=16 birds) and the R^2 represents the strength of correlation.

2.3.8 'Initial' state of IN progression correlates with time to song initiation

The results suggested that the progression of IN timing and acoustic features are controlled by internal neural processes possibly related to motor preparation. In other systems, neural preparatory activity is strongly correlated with the time to movement initiation: the greater the progress of preparation, the shorter the time to movement initiation (Churchland et al., 2006a; Shenoy et al., 2011, 2013). Similar to this, I found a significant correlation between the length of the interval between the first two INs and the time to song initiation in all birds (see example data from one bird in Fig. 2.13 A; across all 16 birds, mean r=0.77, range=0.57–0.90). How similar the first IN was to the last IN was also correlated with the time to song initiation, albeit to a weaker extent in 14/16 birds (see example from one bird in Fig. 2.13 B; significant in 14/16 birds, mean r=0.32, range=−0.39–0.62). These data suggested that IN timing and acoustic features reflect internal neural processes, possibly involved in preparing the zebra finch brain for song initiation.

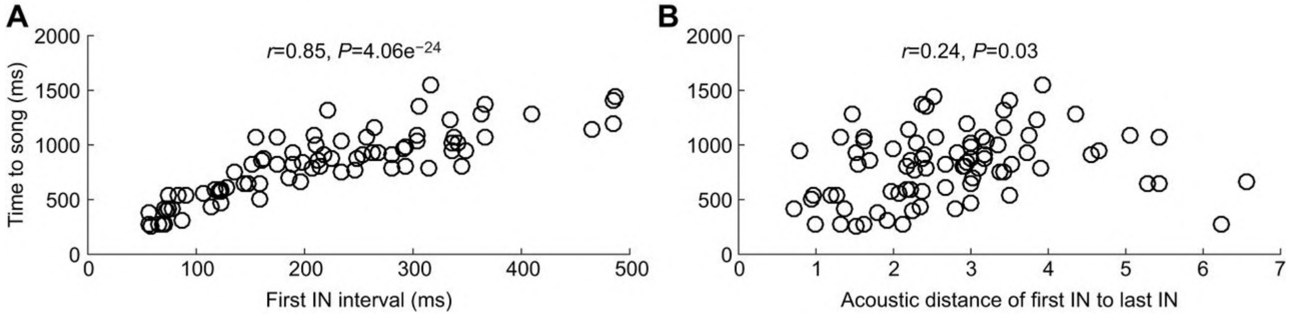

Fig. 2.13 'Initial' state of IN progression correlates with song initiation delay. (A,B) Correlation between the interval between the first two INs (A) or the acoustic distance of the first IN to the last IN (B) and the time to the start of song (onset of the first motif syllable) for one bird. Circles represent data from individual bouts. Pearson's correlation coefficient was calculated.

2.4 DISCUSSION

In this study, I show that real-time auditory and/or proprioceptive feedback is not required for initiation of adult zebra finch song. I also show that the progression of INs, the repeated pre-song vocalizations, from a variable initial state to a more stereotyped final state is also independent of real-time sensory (auditory and/or proprioceptive) feedback. Further, I show, in both feedback-intact and feedback-deprived birds, that fewer INs are present when the first IN of a song bout occurs within 200 ms of the end of a call (other non-song vocalization). In such cases, IN timing was closer to the final state. Finally, the 'initial' state of IN progression was correlated with the time to song initiation. Overall, these results demonstrate that the progression of INs to song does not require real-time sensory feedback. Rather, progression of INs to song is controlled by internal neural processes possibly involved in preparing the motor system for song initiation.

2.4.1 Contributions of respiratory feedback to song initiation

One feedback that I did not alter is respiratory feedback from the air sacs (Méndez et al., 2010). However, previous work strongly suggests that respiratory feedback does not contribute to IN initiation. First, one earlier study showed that disrupting respiratory pressure during short syllables (of the order of 60 ms) did not disrupt song progression (Amador et al., 2013). Given that INs are short syllables of the order of 60 ms, INs may not require real-time respiratory feedback for progression to the next syllable (or song). Second, unilateral disruption of vagal feedback mostly affected syllables at the end of a song (Méndez et al., 2010). Finally, sparse, patterned neural activity of one class of neurons in the premotor nucleus HVC during singing was also not affected by removal of sensory feedback including respiratory feedback (Vallentin and Long, 2015). All of these data suggest that respiratory feedback does not play a role in IN progression.

2.4.2 Long-term requirement for sensory feedback

Song production in adult birds does not depend on real-time sensory feedback (Konishi, 1965; Bottjer and Arnold, 1984) and this study further shows that song initiation also does not depend on real-time sensory feedback. However, long-term song maintenance does require intact sensory feedback, as shown by song degradation starting many weeks after deafening (Nordeen and Nordeen, 1992; Williams and McKibben, 1992; Horita et al., 2008). Similarly, it is possible that sensory feedback could be necessary in the longer term for maintenance of IN progression to song (this study focused on songs produced within 10 days of removal of feedback). It would also be interesting to see whether song degradation seen at later time points after deafening is linked to (or caused by) a change in IN progression to song. If INs represent preparatory vocalizations, such a link would be expected as small changes in the neural preparatory state in primates are correlated with changes in features of the upcoming movement (Churchland et al., 2006c; Afshar et al., 2011).

2.4.3 Comparison of INs to motor preparation in other systems

Preparatory neural activity has been described as a slow change in neural activity, starting as early as 1 s before the start of a movement (Tanji and Evarts, 1976; Romo and Schultz, 1987; Lee and Assad, 2003; Churchland et al., 2006b; Maimon and Assad, 2006; Murakami et al., 2014; Li et al., 2015; Chen et al., 2017; Gao et al., 2018). One important characteristic of this preparatory activity appears to be a decrease in variability across trials (Churchland et al., 2006a, 2006b). The decrease in variability as INs progress to song (Rajan and Doupe, 2013) is very similar to the decrease in variability in neural activity seen before the start of a movement. Together with the current data showing that sensory feedback is not important for progression of INs to song, INs may represent preparatory activity. Additionally, earlier studies have shown the presence of preparatory neural activity in song control areas well before the first IN of undirected song bouts (Hessler and Doupe, 1999; Kao et al., 2008; Woolley et al., 2014; Rajan, 2018) and directed song bouts (Daliparthi et al., 2018). Thus, INs may reflect a continuation of this preparatory activity that begins hundreds of milliseconds before the first IN.

2.4.4 Overt movements in other systems as motor preparation

The results suggest that overt vocalizations (INs) serve as preparatory activity. Previous studies describing neural preparatory activity in primates and rodents before the onset of a movement have not described similar overt movements as motor preparation (Tanji and Evarts, 1976; Romo and Schultz, 1987; Churchland et al., 2006a; Murakami et al., 2014; Chen et al., 2017a; Gao et al., 2018). However, all of these studies have involved training animals to perform a task and animals are rewarded for maintaining stable posture without movements until a ‘go’ signal is provided for movement initiation. Therefore, overt preparatory movements, if present during the early stages of learning, would not be reinforced. This raises two interesting questions for further experiments. (1) Are overt movements present at early stages of learning in primates and rodents too? (2) Given that songbirds learn their song with internal reinforcement cues that only reinforce similarity to the tutor song (or tutor song memory) (Fee and Scharff, 2010), are INs learned similar to song learning? Additionally, there are human studies showing the presence of small eye movements (microsaccades) and small limb movements while waiting for a ‘go’ cue to perform an eye or limb movement (Betta and Turatto, 2006; Cohen and Rosenbaum, 2007; Corneil and Munoz, 2014). Changes in pupil size have also been shown to correlate with preparatory activity (Wang et al., 2015). This suggests that overt movements like INs may be more common before the start of naturally learned movements and may reflect motor preparation.

2.4.5 Mechanisms for IN progression to song

The results show that sensory feedback is not essential for IN progression to song. Rather, the properties of INs correlate with the time to song initiation. How do the properties of INs change to progress to song? In the current study, I showed that the presence of calls prior to the first IN was correlated with shorter intervals between the first two INs and fewer INs before song. Similarly, shorter intervals between the first two INs have also been observed when neural preparatory activity in the premotor nucleus HVC precedes the first IN (Rajan, 2018). As calls are also associated with increased neural activity in many song control areas (Hahnloser et al., 2002; Kozhevnikov and Fee, 2007; Long and Fee, 2008; Benichov et al., 2015; Vyssotski et al., 2016; Danish et al., 2017), the intervals between successive INs may reflect a history of increased activity within these interconnected motor regions. The shorter interval might also lead to short-term plasticity that could facilitate song initiation by speeding up IN progression. Such short-term plasticity has been observed in the inputs to the premotor nucleus HVC (Coleman et al., 2007). Further experiments disrupting short-term plasticity or disrupting activity in motor control regions during IN production could help us to understand the mechanisms of IN progression to song.

Overall, the results of this study show that real-time sensory feedback is not essential for INs to progress to song. Rather, changes in IN properties just before song initiation may reflect internal neural processes, potentially involved in preparing the zebra finch brain for initiation of the learned song sequence.

CHAPTER 3

Relationship between Introductory Notes and Song

3.1. INTRODUCTION

The initiation of song bouts in the zebra finch is dominated by repeats of Introductory Notes (INs) preceding the song (Price, 1979). Different functions have been proposed for INs that suggest different relationships between INs and songs. Recent work on the behavioral patterns of INs, drawing parallels with primate brain preparatory activity before movements, hint at INs having a preparatory role that can predict the nature of the upcoming song (Rajan and Doupe, 2013). This "preparatory" hypothesis (Fig. 3.1 A) suggests that INs could be low-intensity versions of the upcoming song, similar to low amplitude body gestures before a dive or a discus throw. Alternatively, studies of INs in other songbird species like rufous-sided towhees describe an alerting function of INs due to the high detectability and low degradation sound properties over distances, relative to the song Richards, Douglas, 1981). This "alerting" hypothesis (Fig. 3.1 B) suggests that INs could be distinct movements from the upcoming song, like ball-bouncing before a tennis serve or running before a javelin throw.

The above proposed functions for INs are related to the repetition and the presence of INs in the beginning of the bout. These features of INs make them distinct from the song. But, INs are also vocalizations and are produced in sequence with the song. Although there is an innate tendency in the zebra finch to produce INs first and then the song (Price, 1979), it remains unclear whether these vocalizations are just like song. This "vocalizations" hypothesis (Fig. 3.1 C) suggests that INs may be similar to the syllables of the song in the same way as syllables within the song are related to each other (Glaze and Troyer, 2007).

Fig. 3.1 Hypotheses for IN-song relationship
(A)-(C) 'i's represent INs and the space in between represent the intervals. 'abcd' represent song. Similar colored alphabets of varying shades represent related syllables in the sequence, different colored alphabets represent unrelated syllables (A) "Preparatory" hypothesis suggests 'i's could be preparatory vocalizations that are low intensity versions of upcoming 'abcd'. The intensity increases as preparation progresses – the later 'i's become more similar in color to 'abcd'. (B) "Alerting" hypothesis suggests 'i's could function as an alerting signal that have properties different from upcoming 'abcd' – the 'i's have a separate color from 'abcd'. (C) "Vocalizations" hypothesis suggests 'i's together with 'a','b','c','d' form a sequence of vocalizations that share similar properties among them – all letters have the same color.

The above listed speculations about what INs could be make it all the more important to understand whether INs and songs are related to each other. Understanding the relationship between properties of INs and songs will also have further implications for the brain control of both these vocalizations.

What is currently known about the relationship between INs and songs? The properties of INs are relatively more variable compared to the stereotyped song (Chapter 2). INs vary in the number of repeats, the timing and the acoustic properties across renditions. On the contrary, the song consists of learned sequence of notes produced in a fixed order with high temporal and acoustic precision, and thus minimal variation (Price, 1979; Glaze and Troyer, 2012). Do variable properties of INs explain variation in the following song? Within a bird, similar last INs before the song between two trials is related to similar first song syllables, suggesting IN-song relationship (Rajan and Doupe, 2013). But, the exact relationship between INs and song remains poorly understood. Here, I address the question by directly comparing INs and songs in two ways: (1) correlating variation in properties of INs and song across trials. (2) using the known changes to song (with age), correlating changes in the IN properties with changes to the song.

In adults, the post-learning period (after 90 dph or days-post-hatch) up to one year of age is associated with further refinement of the learned song (Pytte et al., 2007; Glaze and Troyer, 2013). The fine changes in the song suggest a better control of song parameters, possibly due to central or peripheral changes with age. If IN sequences are related to the upcoming song, properties of INs are expected to change with age and show similar trends as age-related changes to song. Chapter 2 (same as (Rao et al., 2019)) shows that the properties of INs are maintained within a short time-span of 5 days. Do IN properties change long-term? Typically, in the first year, songs increase in tempo, improve in regularity of temporal properties of notes and gaps and structural properties of notes, and increase in the consistency of transitions between pairs of syllables in the sequence (Brainard and Doupe, 2001; Pytte et al., 2007; Glaze and Troyer, 2013; James and Sakata, 2019). The temporal and structural properties of INs may also show similar changes due to central or peripheral changes with age, irrespective of its relationship to song. Taken together, properties of INs may change in the first year due to its relationship to the song that follows or due to age-related effects that may influence the independent control of INs and song. However, the properties of INs which directly relate to song can be separated as these properties will also show similar relationship to the song on a trial-to-trial basis. Thus, I analyze both correlations of properties across trials and the correlation of change in properties with age to pinpoint relationship between IN and song.

I show that acoustic properties of syllables but not intervals are related between INs and songs across trials (or bouts; bouts and trials are used interchangeably). Like songs, INs also show changes in the first year of age. The number of INs increases with age along with a change in the temporal pattern of the IN sequence, the amplitude modulation of INs and the overall time to start song. As INs and songs show correlated changes with age, it suggests that both IN and song properties are controlled by factors that are influenced by age. Further, on a given day, the number of INs, the time to start the song and the temporal pattern of INs do not relate to the temporal-spectral pattern of song, the speed or the regularity in sequencing of the upcoming song. Together, the analysis on IN-song relationship suggest that the acoustic properties related to amplitude of INs and songs may share similar control . However, the mechanisms controlling the number and timing of INs are different from song. Further comparisons of IN repetitions with call and song syllable repetitions in a subset of birds confirmed that variability in repeat number and acoustic progression are common properties shared with other repeated syllables. However, the temporal progression during INs is an exception and points towards a mechanism of control for IN timing which is independent of the upcoming song.

3.2. METHODS

Bird recording procedures performed at IISER Pune were approved by Institute of Animal Ethical Committee in accordance with the guidelines of the Committee for the Purpose of Control and Supervision of Experiments on Animals (CPCSEA, New Delhi, India). Bird recording procedures at UCSF (CA, USA) were approved by the UCSF Institutional Animal Care and Use Committee in accordance with NIH guidelines.

3.2.1. Birds and Song Recording

The birds (*n*=40) used in this study were either purchased from an outside local vendor (*n*=15) or bred at IISER Pune (*n*=19) or UCSF (*n*=6). Songs of the UCSF birds were recorded as part of a different study characterizing INs (Rajan and Doupe, 2013). Although the study compared how similar the pairs of last INs were to pairs of first song syllables, the direct relationship between individual properties of INs and songs was not compared. Beyond 90 days-post-hatch (dph), zebra finches have stereotyped songs and are considered as adults. Only the songs recorded from adult birds were part of this study (> 89 dph or days-post-hatch). The exact age of purchased birds was not known. At the time of purchase they were within 4 months old (claimed by the vendor) and had red beaks. Juvenile birds below 60 dph typically have black beaks. The age for these birds was assumed to be 60 dph at the time of purchase. Thus estimated age for purchased birds might have been underestimated at the most by two months. Note that this caveat did not greatly affect the age-related analysis as the error in age is smaller than the duration of age groups considered for comparing the age-related changes..6/15 purchased birds were part of the age-related study. Undirected songs were recorded in “triggered” or “continuous” mode as described in Chapter 2 – Methods. Three sets of birds were used for different analyses and the overlap between these sets are as explained below.

3.2.1.1. *IN-Song correlations*

20 birds, recorded on multiple days (2-8 sessions) ranging in age from 89-1087 dph, were analyzed for age related changes in INs and songs. The microphone position at the roof of the cage was maintained for the same bird on different days of recording, as much as possible. The age for purchased birds (n=6/20) was assumed to be 60 dph at the time of purchase, as mentioned above. A subset of birds and sessions (14/20) overlapped with those analyzed for day-to-day changes in IN properties in Chapter 2. One bird (1/20) overlapped with the birds used for ts-cut surgery, but the recording sessions analyzed here are from well before the surgical procedure. A subset of birds (5/20) were recorded earlier at UCSF and have sessions from two nearby days, <5 days apart, included in this analysis.

3.2.1.2. *Song recordings with head-fixed microphone*

Recording songs from suspended microphones, as in the previous set of birds, have a caveat that the amplitude received at the microphone depends on the relative position between the bird and the microphone. An artifact in amplitude measurements may arise due to the change in position of the bird across trials, and influence the correlation analysis between acoustic properties of INs and songs. To control for the amplitude, the correlation analysis of acoustic properties was repeated in undirected songs from 5 birds recorded with head-fixed microphone. These birds were recorded earlier as part of a different study on amplitude related changes in the courtship song of the male based on the distance from the female (Suri and Rajan, 2018). However, the relationship between INs and songs was not compared. One undirected recording session from each of these birds was analyzed to verify the correlations between acoustic properties of INs and song syllables across trials, obtained in the previous set of birds.

3.2.1.3. *Analysis of repeat syllables*

Songs of 17 birds in the colony were identified that specifically repeated at least one syllable within the motif. One session from each of these birds was used to compare the syllable repetition properties among the different types of syllables produced by the bird namely INs, motifs and calls. The session from one bird overlapped with that analyzed for IN- song correlations. One of the birds recorded in this set (1/17) was bred and recorded earlier at UCSF.

3.2.2. Data analysis

All the analyses were performed using custom-written scripts in MATLAB (www.mathworks.com).

Audio files were processed and vocalizations were labeled as mentioned in Chapter 2 – Methods. Consistent with Chapter 2, a bout was defined as a period of vocalizations separated by 2 seconds of silence. The 2 seconds criteria for identifying bouts was used to select bouts for analysis of continuous data. However, many audio files recorded in triggered mode did not have 2 seconds of silence before the first syllable in the file. This can happen if initial vocalizations in the bout were soft and did not cross the trigger. Thus, the audio files saved with 1-2 seconds buffer time before the trigger onset may have less than 2 seconds of silence before the first syllable in the file. For such files, silence was assumed before the start of the file as the set trigger-threshold was not crossed. The bout criteria for triggered recordings was reduced to include enough audio files for analysis and the criteria ranged from 500-1500 ms across birds. When analyzing the same bird across sessions, a consistent bout criteria was maintained (explained later in section: 3.2.2.3 *Picking bouts on different days with a common bout criteria*).

3.2.2.1. *Determining minimum number of bouts for IN analysis*

The number of INs varied across bouts and the number of bouts sung by the birds varied across sessions and birds (range [4-430], median 74). It was important to characterize how the number of bouts affected the estimate of IN number for a session, and to determine the minimum number of bouts beyond which the estimate did not change considerably. In 15 birds, for one session each with >100 song bouts, a range of different number of bouts were sub-sampled and the corresponding average standard deviation across 1000 iterations of Monte Carlo simulations was calculated. (Fig 3.2– Simulations from a sample bird session). The minimum number of bouts for IN analysis was identified as the number of bouts beyond which increasing the number of samples did not change the average standard deviation by more than 0.01. The minimum number of bouts was 10 for a majority of the birds and 15 for most birds (14/15, and one bird it was 20). Hence, across all days and birds, only sessions with at least 15 song bouts were selected for IN analysis. Two sessions from a bird got excluded after the criteria was applied.

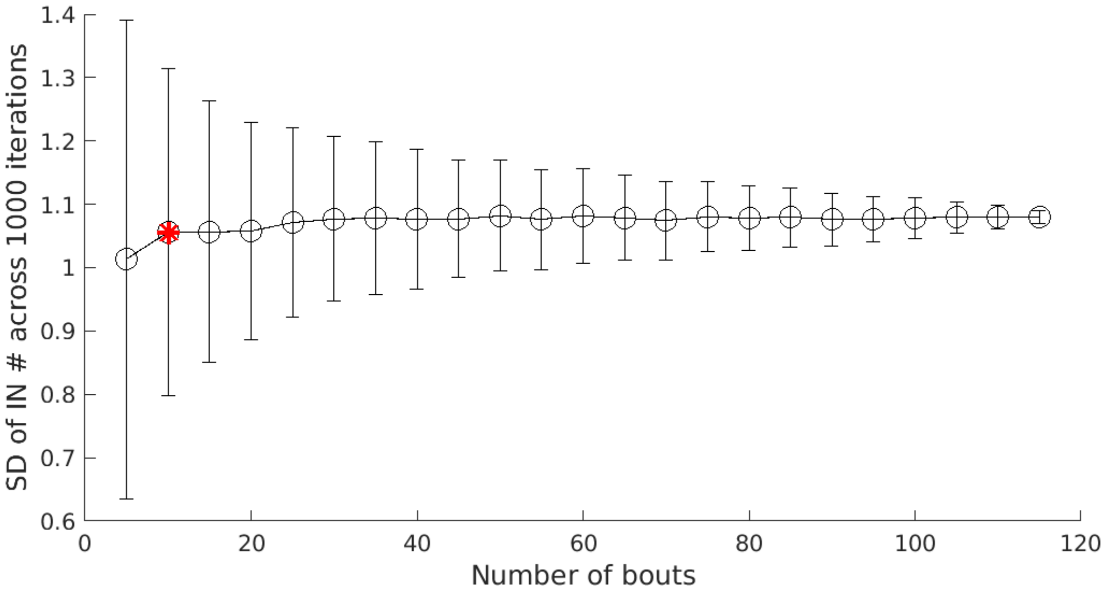

Fig. 3.2 Determining minimum number of bouts for IN analysis: example from one recording session of one bird

Standard deviation of IN number averaged across 1000 simulations (y-axis) was calculated for different number of bouts (x-axis) sub-sampled from total number of bouts in the recording session (117 bouts). Circles and whiskers represent mean and standard deviation respectively for sample standard deviation measured 1000 times. Red star marks the number of bouts (15) beyond which increasing the sample number of bouts increased the average standard deviation of IN number by less than 0.01.

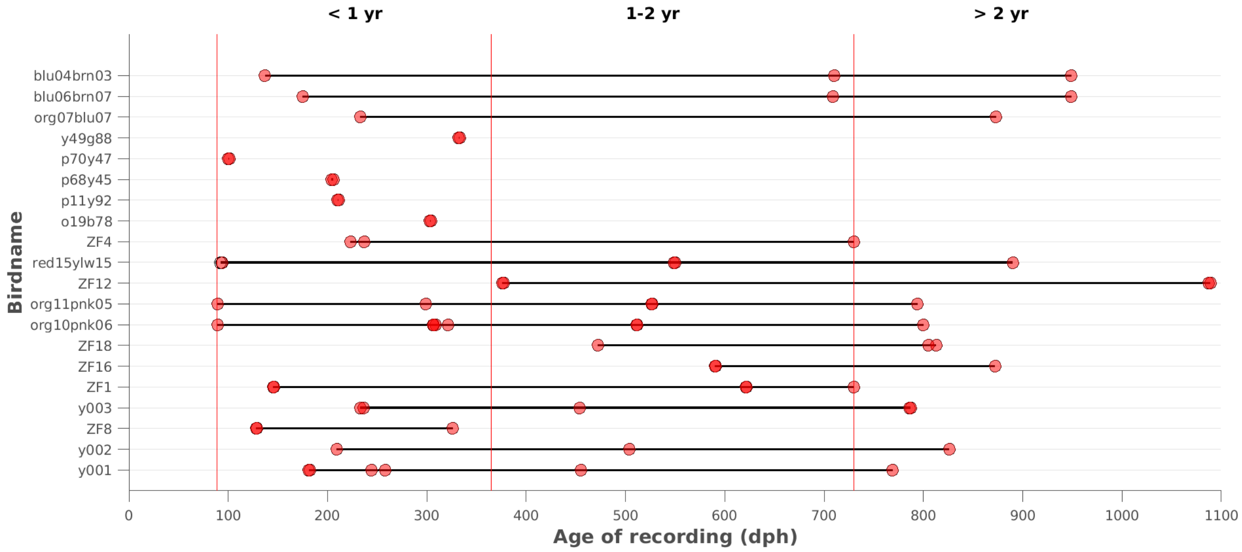

Fig. 3.3 Age of recording across birds

Birds (y-axis) recorded on multiple days of age (x-axis). Each line joined by circles is data from one bird. Each circle represents an individual recording session. The nearby days in some birds are too close to be distinguishable.

3.2.2.2. Age and time of recording

To estimate age-related changes in IN and song properties, birds were recorded at multiple time points covering a range of ages from <1yr to ~ 3 yrs of age (89-1087 dph) (Fig 3.2.). The difference between days of recording also spanned a wide range (1-812 days). This allowed for measuring the change in IN and song properties over short-term (day-to-day) and long-term (between many days

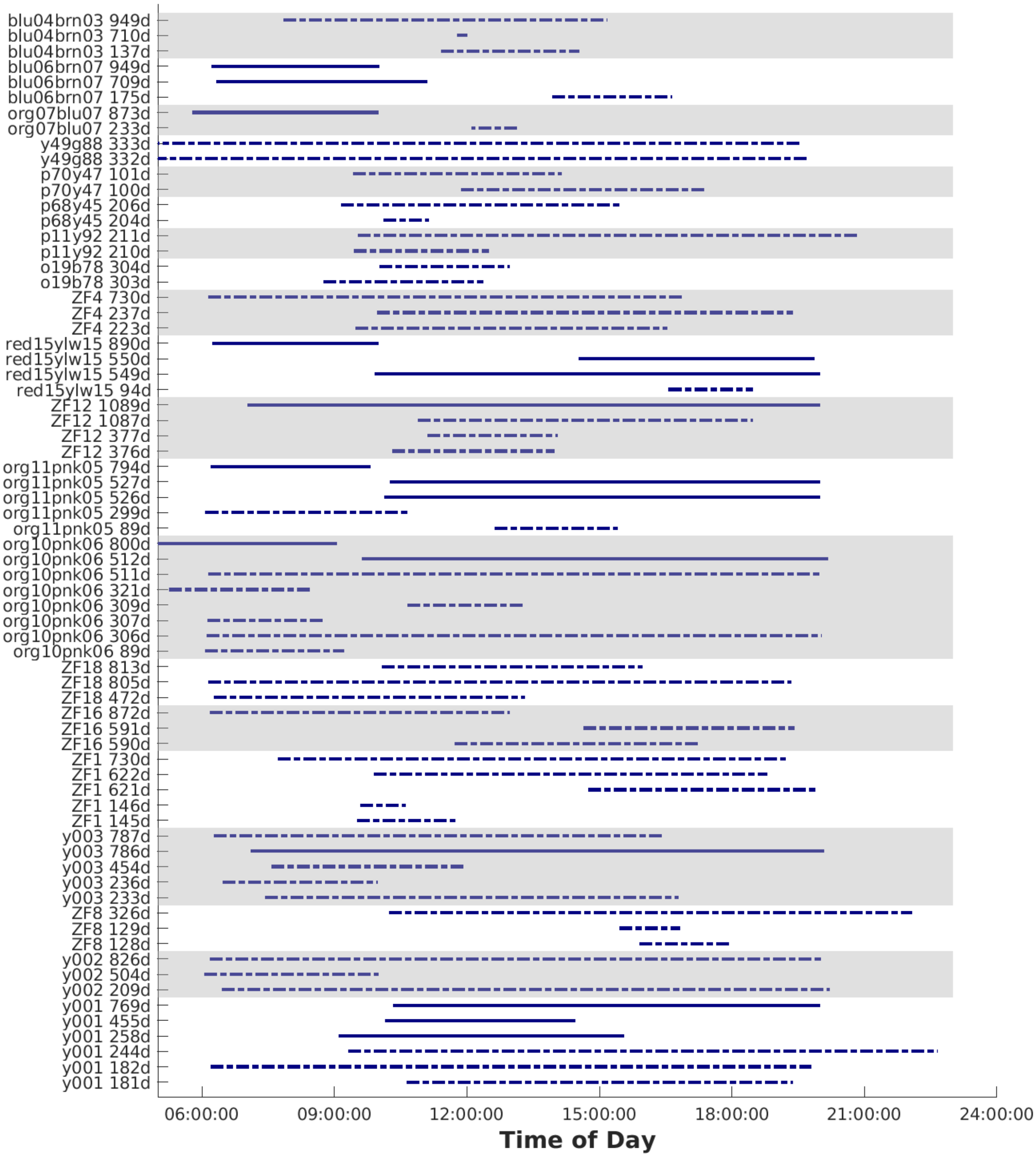

Fig. 3.4 Time and type of recording
Horizontal lines represent the time of recording sessions (x-axis) for different birds on different days (birdname and age mentioned on y-axis ticks). Recordings were done in 'continuous' (continuous line) or 'triggered' (dashed lines) mode. All the birds are recorded during day time of the photoperiod from 06:00 to 20:00 hrs, except for birds y001 on 244 dph and ZF8 on 326 dph, when the photoperiod maintained was from 08:00 to 22:00 hrs as part of a different experiment.

of age). The time of recording varied across sessions for the same bird and these durations may or may not have overlaps (Fig 3.3). All the recordings were done during the daytime of the photoperiod (14h light/10 h dark) that was maintained in the soundbox.

3.2.2.3. *Picking bouts on different days with a common bout criteria*

The recordings from different days in a bird and across birds were either in ‘triggered’ or ‘continuous’ mode (Fig. 3.4). Out of 72 sessions recorded, 56 were ‘triggered’ and 16 were ‘continuous’ mode. As mentioned earlier, the ‘triggered’ recordings in few cases deviated from the 2 seconds silence criteria in order to pick up enough bouts for analysis. It was observed that changing the bout criteria changed the number of INs, more so for interval criteria that were less than 2 seconds. However, the difference in average IN number between days was maintained for the same bout interval criteria (Fig 3.5 sample bird). To avoid the influence of bout criteria on the change in mean number of INs between days, the minimum bout criteria across different days of recordings was selected and consistently used to pick bouts on all the days recorded in that bird.

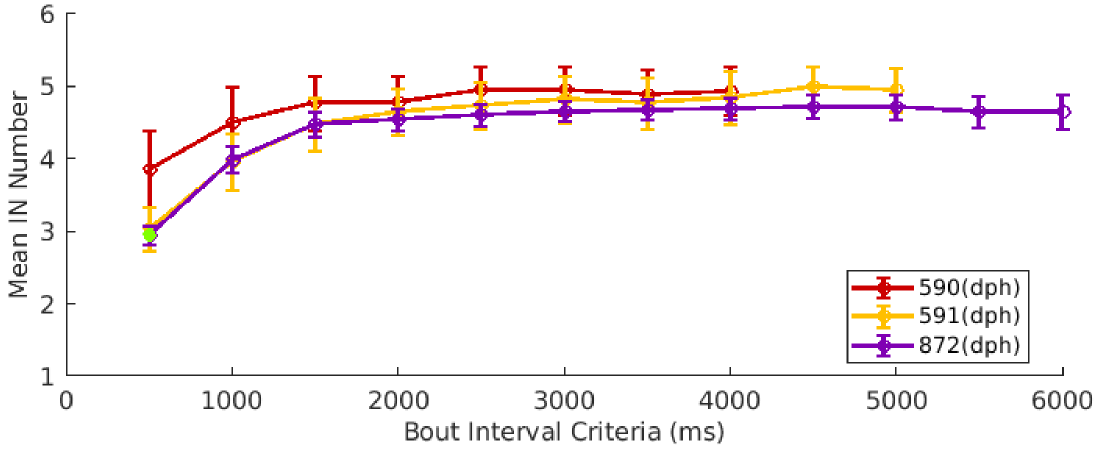

Fig. 3.5 Selecting bout interval criteria for sample bird recorded on multiple days
Mean number of INs across bouts (y-axis) for the bouts selected by using different criteria of silent periods between vocalizations (x-axis). Colored lines represent different days of recording. Circles and whiskers represent mean and the 95% confidence intervals respectively. Number of INs in this bird is reduced for bout criteria <2000 ms on all the days recorded.

3.2.2.4. *Analyzing the properties of IN and song within a session*

For a given session, different variable properties of INs were measured on every trial: (1) the number of INs, (2) the first IN interval which is the interval between the first two INs, (3) the last IN interval which is the interval between the last IN and the first song syllable, (4) the ratio of intervals between successive INs, (5) the acoustic distance of first and (6) last IN relative to distribution of last INs, and (7) the ratio of acoustic distance of successive INs. The measurement of these properties have been described in Chapter 2 – Methods. Eight acoustic features were used to calculate the acoustic distance of INs in this Chapter, instead of four features used in Chapter 2. Two other temporal properties related to IN sequence were measured, namely time-to-song and temporal similarity. The time-to-song is the duration of IN sequence from the onset of the first IN to the onset of the first song syllable. The measurement of temporal similarity is explained under Section 3.2.2.8. The mean values for all the above properties were calculated for every recording session. The variability of number, intervals and acoustic properties of a syllable was calculated by the Coefficient of Variation (CV = standard deviation/mean) or standard deviation as specified.

In order to correlate interval and syllable properties of INs with song, the interval between first two song syllables and the acoustic properties of the first syllable in the motif were measured for the corresponding trials in each session.

3.2.2.5. *Acoustic properties of syllables*

Eight acoustic properties for IN and song syllables were measured as described by Sound Analysis Pro 2011 (http://soundanalysispro.com/ using SAP MATLAB code, (Derégnaucourt et al., 2005)). These basic features reduce dimensionality of the complex sound spectrogram for analysis and make it easier to understand the changes in different aspects of sound. An intuitive understanding of the features is described below along with equations for calculation. Note that all the equations below have been reproduced from the SAP manual (http://soundanalysispro.com/). All measurements use zebra finch frequency range of 0.3-8KHz.

- Log Amplitude: It measures the loudness of a syllable. The intensity or power in audio signal is measured relative to an arbitrary baseline for silence. It is reported in log scale units or dB.

$$10\log_{10}\sum_{f} P_f - baseline$$

P_f is the power at any one frequency f and baseline is set to 70 dB as default.

- Duration: Time from the onset to the offset of the syllable. Onset and offset are determined based on an amplitude threshold that distinguishes sound from silence in the audio signal. It is measured in units of seconds or milliseconds.
- Mean Frequency: Audio signal can be decomposed into different frequencies present in the signal. The mean frequency is a pitch measure that assesses the center of distribution of intensity across different frequencies. It is measured in Hz.

$$\hat{f}=\frac{\sum_{f} f\left(\partial P_f^2 dt+\partial P_f^2 df\right)}{\sum_{f}\left(\partial P_f^2 dt+\partial P_f^2 df\right)}$$

$\partial P_f dt$ is the time derivative and $\partial P_f df$ is the frequency derivative of frequency f. Since mean frequency depends on intensity, it is amplitude-related property.

- Pitch goodness: It measures the periodicity of harmonic pitch (frequency stacks) in the syllable. Syllables with harmonics or ‘frequency stacks’ (as observed in a spectrogram) have higher value, and syllables that are noisy or pure tone have lower value. It is measured as the peak of the derivative-cepstrum (cepstrum is spectrum of log spectrum) calculated for harmonic pitch. The units are comparable to amplitude and can be converted to dB.
- Weiner entropy or entropy: It measures the noisiness in the syllable. A noisy syllable appears as broadband or white noise in the spectrogram with intensity spread equally across all frequencies. White noise has a value of 1 and pure tone (sound intensity concentrated at one frequency) has a value of 0. However, this is converted to log-scale of 0 to minus infinity. More noisy syllables will have entropy 0. It is measured as ratio of geometric mean to arithmetic mean of the spectrum and does not have units.

$$W=\log\left(\frac{\exp[\int dfLog(S(f))]}{\int dfS(f)}\right)$$

S(f) represents power spectrum at frequency f. Since entropy depends on intensity, it is amplitude-related property.

- Frequency Modulation: It is an estimate of slope of frequency trace on the spectrogram. Steeper the slope higher the modulation. It is the angular component of squared time and frequency derivatives and is measured in degrees.

$$FM=\arctan\left(\frac{Max_f \partial P_f^2 dt}{Max_f \partial P_f^2 df}\right)$$

$\partial P_f dt$ is the time derivative and $\partial P_f df$ is the frequency derivative of frequency f.

- Amplitude Modulation: It is an estimate of changes in amplitude envelope or loudness of sound during the syllable. It is positive in the beginning and negative at the end of each sound. It is measured as the overall time derivative across all frequencies within range. Units of AM are 1/t (1/milliseconds) and time units are defined by the ‘advance window’ parameter.

$$\frac{\sum_{f}\partial P_f dt}{\sum_{f} P_f}$$

P_f is the power at any one frequency f and $\partial P_f dt$ is the time derivative.

- Entropy Variance: It is a measure of variance of Weiner Entropy calculated for the duration of a syllable. Like entropy, this measure has no units.

3.2.2.6. Song properties associated with age-related changes

Parameters of song that have earlier been reported to change with age were measured for every session. These include spectral and temporal similarity (3.2.2.8), motif duration, number of motifs in the bout, song sequence consistency and entropy (3.2.2.9). Motif duration was measured as the time from the onset of the first motif syllable to the offset of the last motif syllable.

3.2.2.7. *Analyzing the change in properties of IN and song between two days*

Change in properties between two days was measured by subtracting the mean and variability calculated for an earlier session from those measured at a later session. All motif related parameters were restricted to the first motif of the bout. This was because I wanted to further analyze the relationship between INs in the beginning of the bout and its following song. The age-related changes were first measured individually for all IN and song properties in order to identify properties that prominently changed with age. The change in IN and song properties were correlated with each other to identify related properties.

3.2.2.8. *Similarity measures for song and IN sequences*

A different algorithm was employed for calculating spectral and temporal similarity of motifs compared to that used in Chapter 2.

- Similarity or Spectral Similarity

  Similarity Index algorithm as described in (Mandelblat-Cerf and Fee, 2014) was used to measure similarity between motifs. It follows a similar procedure as SAP algorithm but has a MATLAB script which made the analysis more efficient. Within-day similarity was calculated by selecting 10 motifs and comparing each of these motifs with the other 9 motifs. This was followed by taking an average of all the comparisons. The within-day similarity measured how similar songs on a given day were to each other. This was measured for motifs within a bout to reproduce an earlier result, followed by analysis of only the first motifs. The difference in within-day similarity values between two days measured the change in the stereotypy of songs between days. Across-day similarity between songs was measured by selecting 10 motifs on two days and comparing each of the motifs on one day with all the motifs on another day. The average similarity across all comparisons was measured as across-day similarity for the pairs of days and gave an estimate of changes in spectral structure of the song between two days.

- Temporal Similarity

  Temporal pattern of motif sequence was obtained by taking the amplitude profile and replacing this with '1s' for the duration of syllables and '0s' for the duration of intervals between syllables. Two patterns were compared up to the duration of the shorter sequence. For calculating temporal similarity between two patterns, the cross correlation method mentioned in Chapter 2 – Methods was used. Within-day temporal similarity was calculated by selecting 10 motifs and comparing each of these motifs with the other 9 motifs. This was followed by taking an average of all the comparisons. This within-day temporal similarity measured how similar the temporal structure of songs on a given day were to each other. The difference in within-day temporal similarity values between two days measured change in stereotypy of the temporal pattern of songs between days. Across-day similarity between songs was measured by selecting 10 motifs on two days and comparing each of the motifs on one day with all the motifs on another day. The average temporal similarity across all comparisons was measured as across-day temporal similarity for the pairs of days and gave an estimate of changes in temporal pattern of song between two days.

  Similar to comparison of song sequences, the temporal similarity was measured for IN sequences starting from onset of the first IN to the offset of the last IN. Two IN sequences were aligned at the offset of the last IN. The duration of IN sequences varied across trials. To make the durations equal for the two patterns being compared, '0s' were added as silent periods before the first IN of the shorter sequence. No stretching was employed, similar to Chapter 2) due to high variability across patterns of IN sequences. To maintain consistency, no stretching was employed for the temporal similarity calculations of song sequences as well.

3.2.2.9. Song sequencing properties

Sequence Consistency and Entropy were measured as described in James and Sakata 2019 (James and Sakata, 2019). The analysis was restricted to the first motif sequence in the bout. For each motif sequence one additional syllable at both ends was included for calculating sequence consistency and entropy.

- Sequence Consistency (SC): This measures the consistent transitions between syllables of the motif. It is calculated as the ratio of most typical transitions divided by total transitions (also (Scharff and Nottebohm, 1991)).
- Sequence Entropy(SE): This measures the variability of syllable transitions. It sums the transition probabilities of all transitions (p_i) from a syllable using the formula $\Sigma\text{-}p_i \times \log_2(p_i)$. Higher the value more variable the transition.

3.2.2.10. Repeat syllable analysis

The number, interval and acoustic progression for all types of repeat syllables was analyzed using a similar procedure as that for repeats of INs described above (section 3.2.2.4). As other repeats may not be followed by a syllable like INs, the last interval was taken as the interval between last two repeat syllable for the analysis of temporal progression. Only 4 acoustic properties were used to calculate acoustic distance. These were duration, mean frequency, log amplitude and entropy (described in section 3.2.2.5).

3.2.2.11. *Statistics*

Sample sizes

As recording sessions with a minimum of 15 song bouts were used for analysis (as described under section 3.2.2.1), the sample sizes for analysis of IN and song properties within a session were maintained above 15. This data was used for correlating IN and song properties within a session and for calculating the mean and variability (standard deviation or coefficient of variation) on a given day (for properties described in section 3.2.2.4). Among all the IN and song properties measured (Section 3.2.2.4 – 3.2.2.9), only the number of INs and the number of motifs in a bout are discrete variables. The change in mean and variability of IN and song properties was measured between pairs of days (Section 3.2.2.7). The sample sizes of number of birds for change between paired days in the different age groups were small but comparable (10-12) and non-parametric tests were employed for comparing age-related changes. These data were also used for correlating change in IN and song properties with age. The analysis of repeat syllables calculated mean properties across at least 10 instances. The analysis was done on birds with repeating syllables in their motifs. Thus number of data points in motif and IN groups were high, but the numbers in the call groups were few.

Statistical tests employed

Paired comparisons:

Wilcoxon signed-rank test was used for paired comparisons of intervals/ acoustic properties between INs and songs for one session each across birds (Fig. 3.6). As INs and song properties may have variable distributions across birds, non-parametric test was employed. This tested the null hypothesis that medians are equal between IN and song property measured across birds. A significance level of p-value ≤ 0.05 was used to reject the null hypothesis. Similarly, Wilcoxon signed-rank test was employed to compare average song syllable properties between higher and lower values of INs (Fig. 3.24 C, D, F). This test was also used to compare properties of first and last repeat position – mean and CV of intervals and acoustic distances for INs, motifs and call syllables (Fig. 3.25 D, E, G). Call syllables repeats were low in number and hence to maintain uniformity the non-parametric test was employed to test the null hypothesis that first and last positions are same. A significance level of p-value ≤ 0.05 was used to reject the null hypothesis.

Correlations:

Pearson's correlation coefficient was used to measure all correlations – Correlation between IN and song properties within a session (Fig. 3.7 and 3.8, Fig. 3.24 B, E, G) and correlation between age-related changes of IN and song properties (Fig. 3.15, 3.21, 3.22, 3.23, 3.24 B, E, G) . Prior to measuring correlations, the outliers were removed for continuous variables using 3 median absolute deviations, and for discrete variables using 3 standard deviations from the mean. Correlations were considered significant if p-value was ≤ 0.05. Qualitatively, the strength of significant correlations were interpreted based on correlation coefficients as negligible <0.1, weak 0.1-0.39, moderate 0.4-0.59, strong 0.6 to 0.79, very strong >0.8. Correspondingly the R-squared values were used to interpret correlation strength as negligible <0.01, weak 0.01-0.15, moderate 0.16-0.35, strong 0.36 to 0.63, very strong >0.64.

Multiple group comparisons:

Kruskal-Wallis test was used for multiple group comparisons (more than two groups) as the data across groups was less, or unequal or the underlying distribution was unknown. This was used to compare number of INs across days within a bird (discrete distribution and unequal sample sizes; Fig. 3.9A), song similarity between age groups (unknown underlying distribution; Fig. 3.16 B, C), change in mean IN or song properties between age groups (small sample sizes; Fig. 3.10- 3.14, 3.17-3.20), comparing motif duration corresponding to different IN numbers across birds (unequal sample sizes with some sizes being small; Fig. 3.24 A), and comparing average repeat properties of IN, call and songs (Fig. B, C, F, H). Kruskal-Wallis test tested the null hypothesis that all groups have the same mean ranks and a significance level of p-value ≤ 0.05 was used to reject the null hypothesis. If the null hypothesis was rejected, Dunn's post hoc test was employed to measure which pairs of groups were significantly different (code from: Cardillo G. (2006)). Dunn's test accounted for (FWER) in multiple comparisons and generated a critical Q-value for the studentized difference in mean ranks. Pairs with Q-value above the critical value were considered statistically different.

Decision rule for significance and interpretation of results

Results are reported as statistically significant when the significance level were met for the different statistical tests described above. However, attention was also given to magnitude of change (qualitative) to consider trends in data and to speculate the biological significance of the results.

3.3. RESULTS

To understand the relationship between INs and the upcoming song, I asked two specific questions: (1) whether variability in IN properties explain variability in the upcoming song across trials. (2) whether changes in IN properties explain changes in song properties with age. Similar to Chapter 2, all the analyses were done on self-initiated or undirected songs.

Three properties that define IN variability are the number, the intervals and the acoustic properties of INs (Fig. 3.6 C-K). The number of INs, at the beginning of the bout is variable compared to the number of syllables in the song which is fixed (Fig. 3.6 B – sample bird). To address whether IN variability explains song variability across trials, I asked whether interval and acoustic properties were correlated between INs and songs.

3.3.1. Variability of IN intervals is not correlated to variability of upcoming-song intervals.

IN sequences have variable interval patterns before the song (Fig. 3.6 B – sample bird). The variability of intervals reduces as INs progress to the song (Morris, 1954; Immelmann, 1969; Zann, 1996). The first interval between the first two INs (referred to as first interval) is more variable than the interval between the last IN and the first song syllable (referred to as last interval) (TABLE 3.1). Due to the difference in variability, the intervals at positions immediately following the first and last INs were separately compared with the intervals between first two syllables of the upcoming song, across trials. The interval between the first two song syllables was used as a measure of song interval in order to maintain consistency with the IN intervals measured between two specific syllables and, also to maintain consistency across birds that may vary in song sequence length and the intervals within them. Across bouts, the IN intervals at both positions were not significantly correlated to the upcoming song interval in a majority of the birds (Fig. 3.7 A-C – sample bird, TABLE 3.2 data across birds). Among a small fraction of birds that showed significant correlations (1/20 with first interval and 6/20 with last interval), the correlation strength (R-squared) was only weak to moderate , suggesting IN interval variability does not sufficiently explain interval variability in the upcoming song. Further, a greater fraction of birds showed significant correlation between two intervals within song (6/15), suggesting intervals between INs are not as related to the song as the intervals within the song.

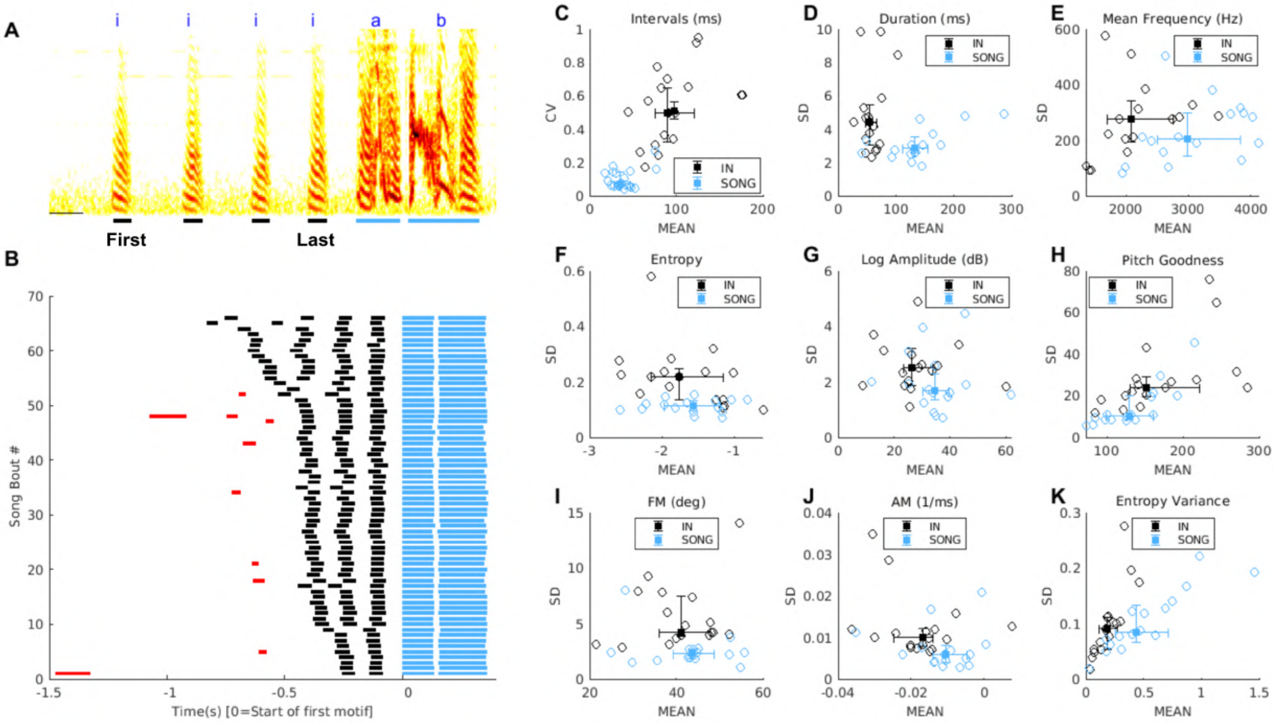

Fig. 3.6 Introductory Notes (INs) are variable relative to the song. (A) Example spectrogram of beginning of the bout showing introductory notes (labelled as 'i') followed by the first song/motif (syllables labelled as 'a' and 'b'). The bars below syllables and in (B) represent duration of syllables separated by silence. Black: INs, Blue: Song and Red: Calls (B) Multiple bouts beginning with INs before first song, all belonging to one session of recording, for the same bird shown in (A). Each row on y-axis represents syllables from one bout with durations aligned to onset of motif as 0 seconds on x-axis. Bouts are sorted in descending order of number of INs. (C-K) Mean (x-axis) and variability (y-axis) of IN (black circle) and song (blue) properties. Each circle is one session from one bird. Variability is measured as (C) CV of intervals and (D-K) SD of acoustic properties. Note that (C) CV for intervals between pairs of INs is greater ($p<0.05$, Wilcoxon signed-rank test) compared to CV of intervals between first two syllables of motif (D-K) SD for acoustic properties measured for IN syllable is higher than that of first motif syllable for most properties (for D, F-J, $p<0.05$ Wilcoxon signed-rank test). Squares and whiskers in (C-K) show median and interquartile range along both axes. Scalebar on spectrogram:100ms. Units: Intervals, duration – ms, mean frequency – Hz, log amplitude – dB, frequency modulation (FM) – degree, amplitude modulation (AM) – 1/ms, no units for entropy, pitch goodness and entropy variance.

IN Property	Mean (MEAN ± S.E.M)	Standard Deviation (MEAN ± S.E.M)
Number	3.20± 0.15	0.85 ± 0.06
First Interval	110.85 ± 10.96	59.82 ± 8.55
Last Interval	52.58 ± 5.49	7.66 ± 1.11
Ratio of intervals	0.93± 0.14	
Acoustic distance of first IN	8.32± 1.75	
Acoustic distance of last IN	3.25± 0.23	
Ratio of acoustic distance	0.814± 0.03	

TABLE 3.1 Variable properties of INs – average properties across birds

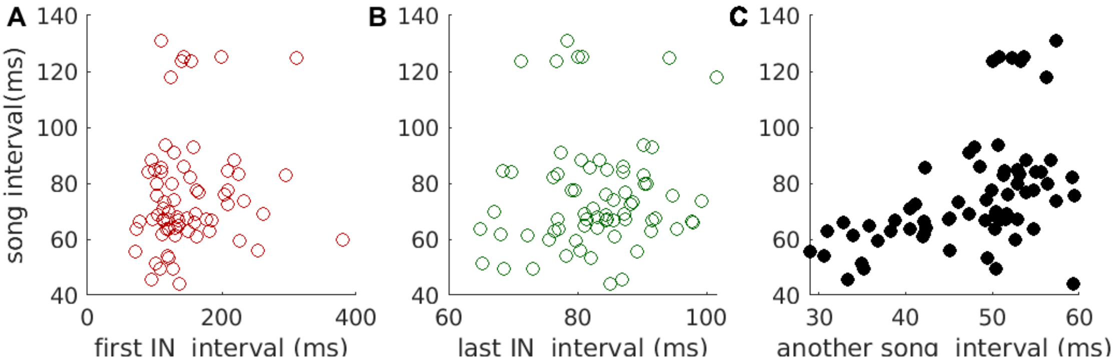

Fig. 3.7 Trial-to-trial variation of IN intervals is not correlated to variation of upcoming song intervals.

Example from one session of one bird that showed no correlation of song interval with either first IN interval or last IN interval. Circles represents trial or bout with interval between first two syllables of the motif on y-axis and (A) interval between first two INs (red) or (B) interval between last IN and the first song syllable (green) or (C) interval between 2^{nd} and 3^{rd} song syllable on the x-axis. Filled circles represent significant correlation. p-value <0.05, Pearson's correlation coeffiecient.

Interval	Proportion of birds significant	R-squared range for significant birds
First IN-song	1/20	0.04
Last IN-song	6/20	0.08-0.3
song-song	6/15	0.04-0.6

TABLE 3.2 Correlation betweeen IN and song intervals across trials

Pearson's correlation coefficient with p-value ≤0.05 is considered significant

3.3.2. Variability of acoustic properties of INs is correlated to variability of upcoming-song syllables

The variable acoustic properties of INs were compared to that of the upcoming song. Eight acoustic features (see METHODS 3.2.2.5) of INs at the first and last IN positions were separately compared to the first syllable in the song across trials (Fig. 3.8 A-H – sample bird). The first and last positions were considered separately as the acoustic properties of INs change with position (Rajan and Doupe, 2013). Only the first syllable in the song was considered for analysis in each bird in order to maintain consistency with IN measure for one specific syllable and, also to maintain consistency across birds with different number of syllables in the song sequence. The acoustic properties namely mean frequency, entropy and amplitude modulation of INs at both first and last positions were correlated to the first song syllable in a majority of the birds (TABLE 3.3, data across birds). As controls, the same properties were compared between two song syllables. The proportion of birds with significant correlations were similar in all three conditions. This shows that the correlation between IN and song are similar at different IN positions and at par with correlations among syllables of the song. The strength of the correlations varied across birds.

The correlation in amplitude-related properties between INs and the upcoming song may arise due to the variability in position of the bird from the suspended microphone across trials (see METHODS 3.2.1.2). This factor was ruled out as similar range of correlations in acoustic properties of INs and song were also present in birds recorded with head-fixed microphone (Fig 3.3 I-P, TABLE 3.3). Further, the proportion of birds with significant correlations were also similar to that observed with suspended microphone (TABLE 3.3). Additionally, the log amplitude of IN and song syllables was also correlated between syllables when recorded with the head-fixed microphone (TABLE 3.3). Together, the correlation results demonstrate a significant correlation between INs and song syllables

Overall, the trial-to-trial correlations between the individual elements of INs show that the variability in acoustic properties of INs, but not the intervals between INs explain variation of the upcoming song across trials.

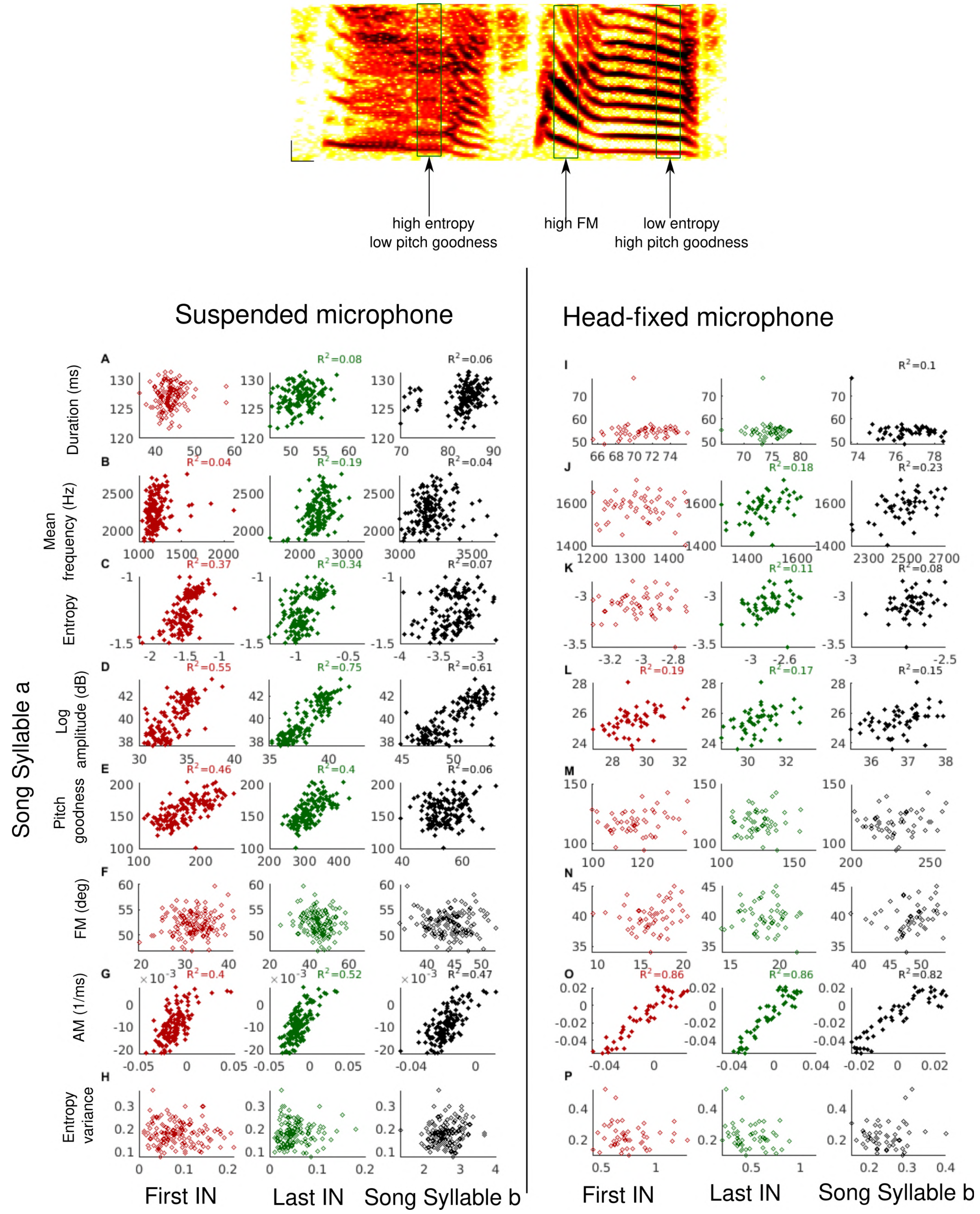
high entropy
low pitch goodness
high FM
low entropy
high pitch goodness
Suspended microphone
Head-fixed microphone
Song Syllable a
Duration (ms)
Mean frequency (Hz)
Entropy
Log amplitude (dB)
Pitch goodness
FM (deg)
AM (1/ms)
Entropy variance
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
R²=0.08
R²=0.06
R²=0.04
R²=0.19
R²=0.04
R²=0.37
R²=0.34
R²=0.07
R²=0.55
R²=0.75
R²=0.61
R²=0.46
R²=0.4
R²=0.06
R²=0.4
R²=0.52
R²=0.47
R²=0.1
R²=0.18
R²=0.23
R²=0.11
R²=0.08
R²=0.19
R²=0.17
R²=0.15
R²=0.86
R²=0.86
R²=0.82
×10⁻³
First IN
Last IN
Song Syllable b
First IN
Last IN
Song Syllable b

Fig. 3.8 Trial-to-trial variation of IN acoustic properties was correlated to variation in acoustic properties of the upcoming song syllables
TOP: Spectrogram showing 2 different syllables. The green rectangles mark parts to indicate acoustic properties. Left rectangle shows portions with white noise- high entropy and low pitvh goodness. Middle rectangle has a sweeping frequency pattern indicating high frequency modulation (FM), right rectangles shows harmonic stacks associated with high pitch goodness and low entropy. Scale bar: 20 ms (horizontal) and 1KHz (vertical). (A-H) Examples from one bird comparing 8 different acoustic properties between IN syllables and the first song syllable. Each circle represent acoustic property of first motif syllable of a bout on y-axis against first IN syllable (red) or last IN syllable (green) or second motif syllable (black) in the same bout. Filled circles represent dataset with significant correlations. (I-P) Same as (A-H) for a different bird with head-fixed microphone. Pearson's correlation coefficients with p-value ≤0.05 is considered as significant. Units: Duration – ms, mean frequency – Hz, log amplitude – dB, frequency modulation (FM) – degree, amplitude modulation (AM) – 1/ms, no units for entropy, pitch goodness and entropy variance.

Property	Suspended microphone (*n*=16)			Head-fixed Microphone (*n*=4)		
	First IN syllable	Last IN syllable	2nd motif syllable	First IN syllable	Last IN syllable	2nd motif syllable
Duration	2/16 [0.05-0.16]	5/16 [0.07-0.90]	8/16 [0.05-0.60]	1/4 [0.10]	0/4	1/4 [0.41]
Mean Frequency	10/16 [0.07-0.8]	10/16 [0.09-0.53]	10/16 [0.09-0.61]	0/4	2/4 [0.25-0.26]	3/4 [0.18-0.26]
Log Amplitude	6/16 [0.04-0.43]	6/16 [0.09-0.4]	6/16 [0.03-0.40]	0/4	2/4 [0.21-0.23]	2/4 [0.13-0.37]
Entropy	14/16 [0.10-0.72]	16/16 [0.09-0.84]	16/16 [0.10-0.90]	1/4 [0.26]	3/4 [0.09-0.29]	3/4 [0.10-0.23]
Pitch Goodness	4/16 [0.09-0.46]	7/16 [0.08-0.40]	6/16 [0.04-0.29]	1/4 [0.55]	0/4	0/4
Frequency modulation	2/16 [0.04-0.08	3/16 [0.05-0.07]	2/16 [0.13-0.15]	0/4	0/4	1/4 [0.14]
Amplitude modulation	12/16 [0.04-0.62]	12/16 [0.16-0.72]	13/16 [0.09-0.63]	3/4 [0.67-0.86]	4/4 [0.67-0.86]	4/4 [0.71-0.84]
Entropy Variance	1/16 [0.13]	3/16 [0.04-0.18]	4/16 [0.04-0.12]	1/4 [0.09]	0/4	0/4

TABLE 3.3.3 Correlation between IN and song syllables
Each cell indicates the proportion of birds with significant correlations along with the range of R-squared values for significant correlations, within parenthesis. Highlighted cells indicate properties for which significant correlation coefficients were observed in at least half of the birds. Pearson's correlation coefficient with p-value ≤0.05 is considered significant. Note that significant correlations between syllables were observed for mean frequency, entropy, amplitude and log amplitude (in the case of head-fixed microphone birds), all of which are amplitude-related acoustic properties

3.3.3. Age related changes to IN sequences in adults show increase in IN number and duration along with changes in temporal patterning and syllable acoustics

Another method to understand the relationship between IN and song is to compare the known changes in song properties with changes in IN properties. Songs of adult birds show changes in the first year post-hatch as increases in the tempo, the spectral and temporal stereotypy, and the regularity of the syllable sequence. I asked whether variable IN properties change in a correlated manner with the known age-related changes in song. As the changes in IN properties with age have

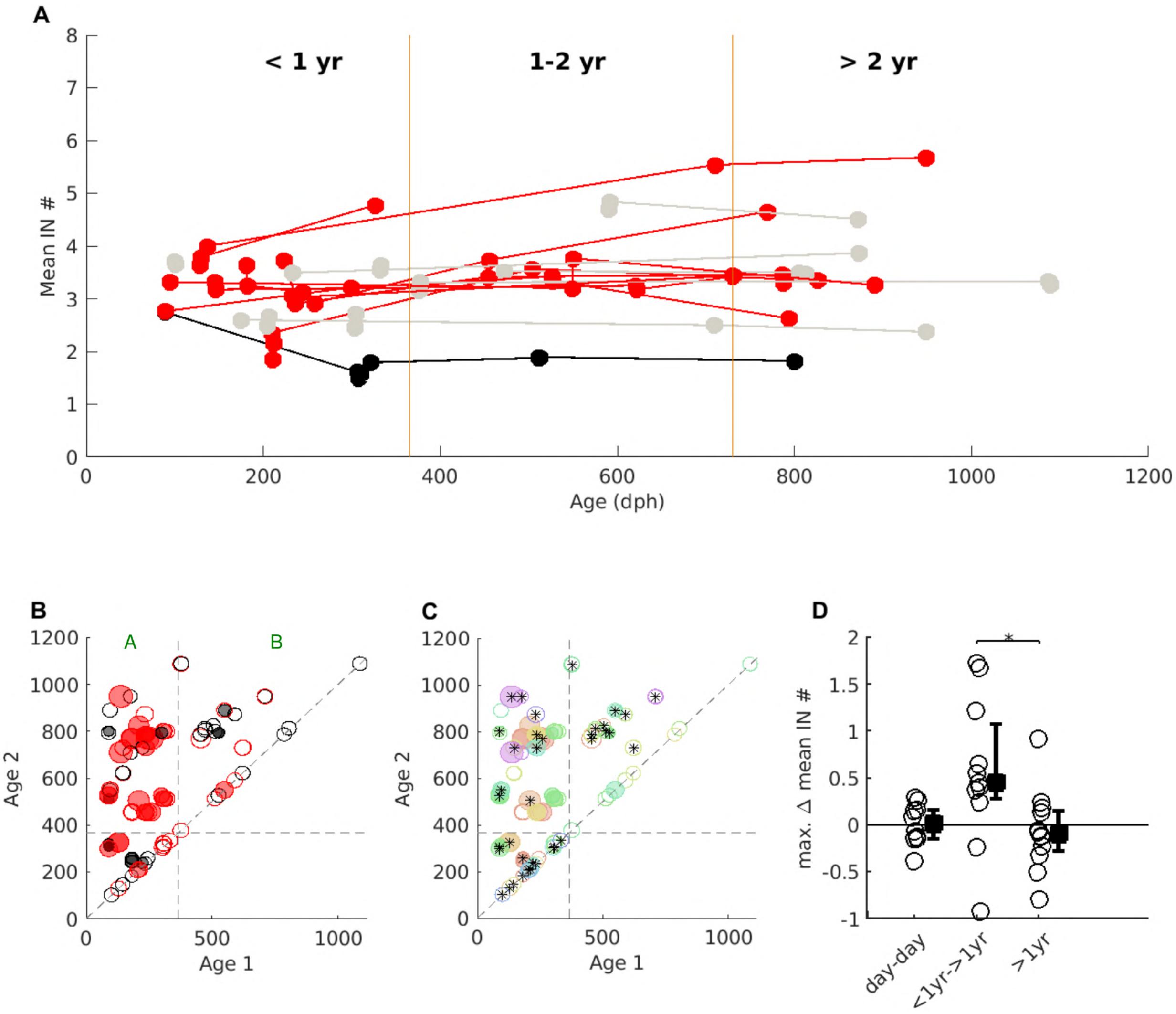

Fig. 3.9 The mean IN number increases in the first year of age
(A) Mean number (#) of INs recorded at different ages across birds. Each circle is mean for one session and each line joins data from one bird. Red lines indicate significant increase, black lines indicate significant decrease and grey lines indicate no change with age for a bird. (B) and (C) show change between mean IN number (size of circle) recorded on one day earlier in age (Age1 on x-axis) vs another day later in age (Age2 on y-axis). The vertical and horizontal dashed grey lines represent 1yr for x and y-axis respectively. The quadrant marked A represents Age1<1yr and Age2>1yr, the quadrant marked B represents Age1 and Age2 both >1yr. The diagonal dashed grey line represents Age1=Age2. Filled circles represent significant change. (B) red circles represent positive change and black circles represent negative change. (C) Different colors represent different birds. Points marked with star are selected change in IN number for (D). (D) Maximum change in IN number between pairs of days belonging to 3 groups: day-day (<5 day difference b/w ways), change b/w first day <1yr and second day > 1yr, change between two days >1yr. Each circle represents data from one bird. Square and whiskers represent median and interquartile range for the groups. p≤0.05 for Kruskal-Wallis test was followed by Dunn's test for individual birds in (A) and (B), and for group comparisons in (D). **Qvalue*>2.39 for (D)

not been characterized, an extensive analysis of the extent of change in IN properties within a bird was first carried out.

3.3.3.1. Age-related changes in mean IN number

Adult birds (*n*=20) were recorded on multiple days spanning a wide range of ages (89-1087 dph). The mean IN number across bouts was first calculated for every session. Within individual birds, the mean IN number was significantly different across days for 11/20 birds and remained the same for others (Fig 3.9 A red and black lines represent birds that showed significant difference between days, Kruskal-Wallis test, $p<0.05$). Out of the birds that showed significant difference, the mean IN number showed an overall increase (10/11) or decrease (1/11) with age. The age of recording was variable across birds and could be the reason for the difference in significant changes for individual birds. To assess this, the change in mean IN number was measured between all pairs of days recorded for a bird and observed for age-related effects (Fig 3.9 B-C). In all the cases, I subtracted the mean IN number of an earlier day (Age1) in age from a later day in age, (Age2). As reported in Chapter 2 (same as (Rao et al., 2019)), the change between nearby days (<5 days apart) at all ages were predominantly small, with no directional preference and no significant difference (Fig 3.9 B circles closer to the central diagonal line are mostly small, colored in both red and black, and unfilled indicating no significant differences). Contrary to this, comparisons between days <1 yr and days >1 yr had more instances with positive, significant changes (Fig. 3.9 B showing bigger, more number of filled red circles for Age1<1yr and Age2>1yr in the quadrant marked A). However, the change between any two days > 1yr were small, with no directional preference and no significant difference (Fig 3.9 B small, unfilled, both red and black circles for Age1 and Age 2>1 yr in the quadrant marked B). Thus, the change in mean number of INs between two nearby days show that short-term changes are not age dependent. However, long-term changes depended on the age of the days being compared i.e. mean IN number increased significantly in the 1st year and remained consistent after that.

To quantify the extent of change in a bird with age, the average change was measured across birds in three groups: (1) short-term day-to-day change (between two days <5 days apart) for all days <1 yr. (2) long-term change between a day <1yr to a day >1yr and (3) long-term change between both days >1 yr. Owing to the nature of recording of a bird on multiple days, the same bird may have more than one pair of days in a group (Fig. 3.9 C). In such cases, the pair of days with the maximum magnitude of change was selected as a measure of the extent to which the mean IN number can change in that bird during that period (Fig. 3.9 C - * shows selected days for different birds). The average extent of change was centered around 0 for day-day group and > 1yr group (Fig 3.9 D, day-day 0.01±-0.02, >1yr -0.05±-0.03). However, the average extent of change showed an increase by half from <1yr to > 1yr (0.56±-0.07). This was significantly different from change >1 yr (Kruskal-Wallis test, $p<0.05$). Together, an average increase in the group <1 yr to >1 yr and no further change in the group > 1yr (near 0) show that the changes occurred in the first year. Overall, these results show that mean IN number is maintained in the short-term, increases by a small amount during the first year of age and then remains consistent up to 3 years of age.

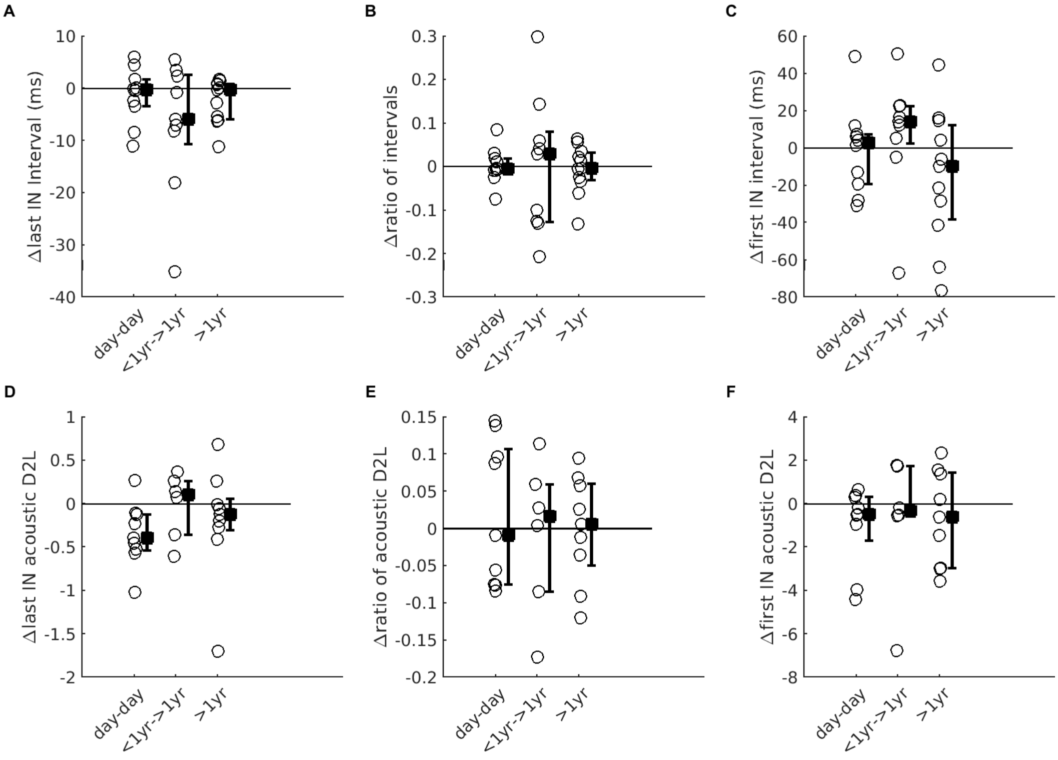

Fig. 3.10 Change in mean properties of IN progression with age
(A)-(C) represent change in properties associated with temporal progression of INs: (A) last interval is interval between last IN and first motif syllable in milliseconds (ms), (B) ratio of successive intervals between INs and (C) first interval between first two INs (ms). (D)-(F) represent change in properties associated with acoustic progression of INs. The acoustic distance to last (D2L) measures similarity to last IN. The change between pairs of days for this similarity at first (D) and last (F) IN position, and the ratio of similarity of successive INs (E) represent change in acoustic progression of INs. All pairs of days in (A)-(F) are the same pairs of days selected for maximum change in IN number in Fig. 3.9. (A)-(F) Each circle represents data from one bird. Square and whiskers represent median and interquartile range for the groups. $p > 0.05$, Kruskal-Wallis test.

3.3.3.2. *Age-related changes in mean and variability of temporal and acoustic progression*

To test if other features of INs also changed with age I considered the pairs of days in the three age groups showing maximum extent of change in mean IN number, and analyzed changes in mean for all other properties related to IN progression (Fig. 3.10 A-F). The change was not significantly different across age groups ($p>0.05$, Kruskal-Wallis test). This shows that the progression of IN intervals given by the first interval (Fig 3.10 C), the ratio of intervals (see Chapter 2 – Methods) (Fig 3.10 B) and the last interval before the song (Fig 3.10A) does not change significantly with age, across birds. Further, the acoustic similarity of INs at all positions to the last IN (Fig 3.10 D, F) and the progression of this similarity (ratio of acoustic distance, see Chapter 2 – Methods) to last IN (Fig. 3.10 E) did not change significantly with age, across birds. The variability of IN progression arises from the number and intervals being variable. The change in variability (CV or coefficient of variation) in number of INs (Fig. 3.11 A), the first interval (Fig. 3.11 B) and last interval (Fig. 3.11 C) was not significantly different across age groups (p>0.05, Kruskal-Wallis test). Taken together, the results show that the mean and variability of IN progression does not change significantly with age.

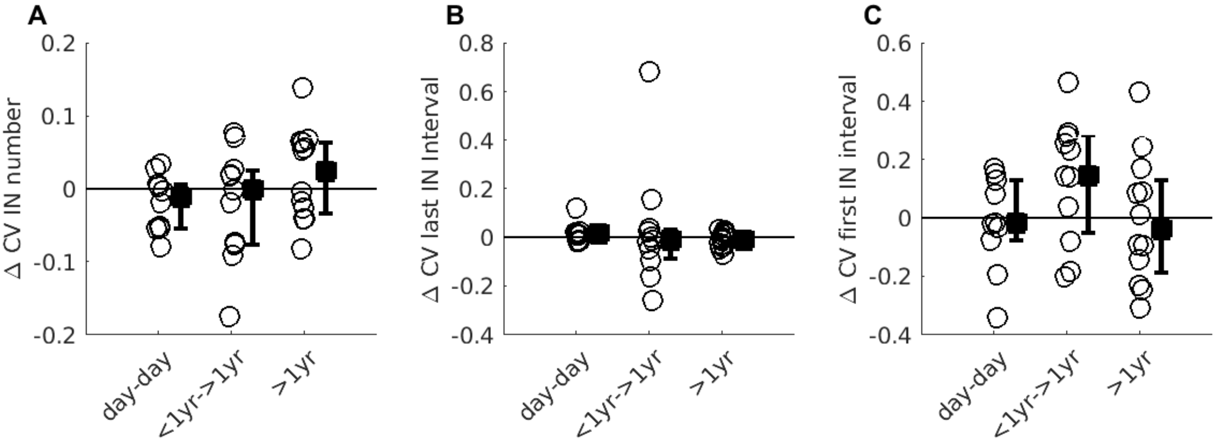

Fig. 3.11 Change in variability of IN progression with age
Change in variability measured by CV across bouts for (A) number of INs, (B) last interval (ms) and (C) first interval (ms) of INs. All pairs of days in (A)- (C) are the same pairs of days selected for maximum change in IN number in Fig. 3.9 D. Each circle represents data from one bird. Square and whiskers represent median and interquartile range for the groups. p >0.05, Kruskal-Wallis test.

3.3.3.3. Age-related changes in IN syllable acoustic properties

The acoustic properties of IN syllables were measured by mean (Fig. 3.12 A-H) and standard deviation (Fig 3.13 A-H) of eight features: duration, mean frequency, entropy, log amplitude, pitch goodness, frequency modulation (FM), amplitude modulation (AM) and entropy variance (see METHODS 3.2.2.5). Among all these properties, significant changes in the first year was observed only for mean amplitude modulation that showed a small increase in magnitude (Fig 3.12 G, p<0.05, Kruskal-Wallis test). Although not significant, trends were observed for changes in the first year for other properties like an increase in entropy (Fig. 3.12 C) and decrease in log amplitude (Fig. 3.12 D). The change in variability was not significantly different across age groups for all properties. Although not significant, trends were observed for decrease in variability of duration (Fig. 3.13 A), mean frequency (Fig. 3.13 B), frequency modulation (Fig. 3.13 F) and increase in variability of amplitude modulation (Fig. 3.13H) (Kruskal-Wallis test) in the first year.

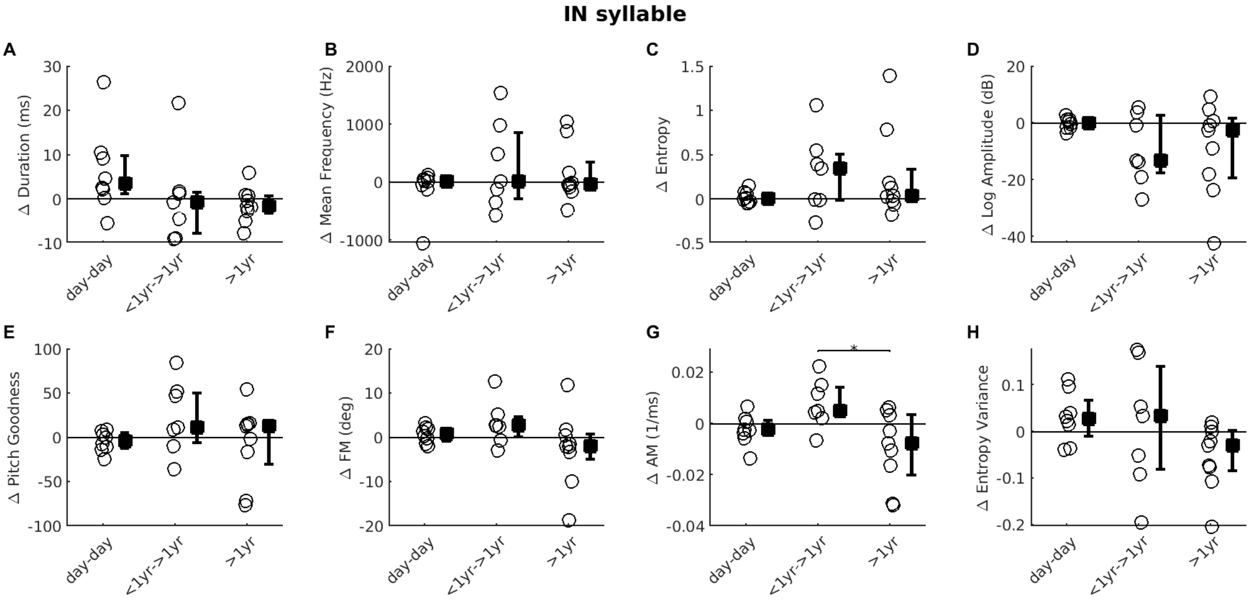

Fig. 3.12 Change in mean acoustic properties of IN syllable with age
Change in mean of 8 acoustic properties for IN syllable between pairs of days (A-H). All pairs of days in (A)- (H) are the same pairs of days selected for maximum change in IN number in Fig. 3.9. Each circle represents data from one bird. Square and whiskers represent median and interquartile range for the groups. $p \leq 0.05$ for Kruskal-Wallis test was followed by Dunn's test. **Qvalue*>2.39. Units: Duration – ms, mean frequency – Hz, log amplitude – dB, frequency modulation (FM) – degree, amplitude modulation (AM) – 1/ms, no units for entropy, pitch goodness and entropy variance.

3.3.3.4. Age-related changes in temporal sequence of INs

The temporal sequence of INs in each song bout is measured by the overall duration it takes to start the song, also called the time to song (see METHODS 3.2.2.4), and the pattern within this sequence consisting of IN syllables and silent interval between INs. The average time to song showed a significant increase from <1 yr to > 1yr compared to other groups of paired days (Fig 3.14 A) as expected by the increase in mean number of INs (Fig. 3.14 D, p<0.05, Kruskal-Wallis test). However, the variability was not significantly different (Fig 3.14 B). Further, the similarity of temporal pattern of INs across trials of a session did not show a change across age groups, indicating that the variability of IN sequence pattern in a given session is maintained with age (Fig 3.14 C, p>0.05, Kruskal-Wallis test). However, the temporal similarity between patterns from <1yr to >1yr although not significantly different, had a reducing trend (p=0.06, Kruskal-Wallis test) implying a change in patterning of IN syllables and silent intervals in the first year of age (Fig 3.14 D).

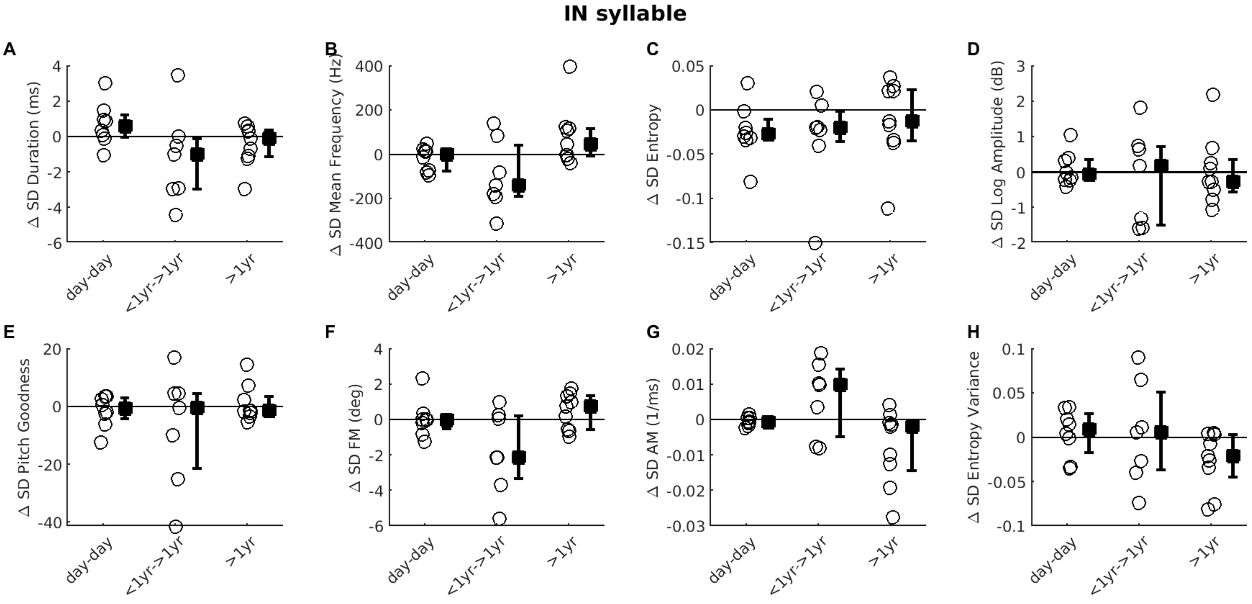

Fig. 3.13 Change in variability of acoustic properties of IN syllable with age
Change in variability of 8 acoustic properties for IN syllable between pairs of days, given by standard deviation (A-H). All pairs of days in (A)- (H) are the same pairs of days selected for maximum change in IN number in Fig.3.9. Each circle represents data from one bird. Square and whiskers represent median and interquartile range for the groups. p >0.05, Kruskal-Wallis test. Units: Duration – ms, mean frequency – Hz, log amplitude – dB, frequency modulation (FM) – degree, amplitude modulation (AM) – 1/ms, no units for entropy, pitch goodness and entropy variance.

Overall, the analysis on IN properties showed age-related changes in the first year. Specifically, the number of INs increase and the total duration of IN sequences get longer. Further, changes occur in the temporal pattern of INs. A subset of the acoustic properties of INs also change in mean and variability. The rest of the properties of INs that do not change significantly with age, did show significant weak or strong correlations with changes in properties that changed significantly with age (Fig 3.15 – property names in bold significantly changed with age). These correlations imply an interdependence of IN properties on each other, suggesting that most properties of INs are affected by age, even if not significant. Together, the results suggest age-related changes to the overall control of the IN sequence.

3.3.4. Age related changes to song in adults show improved spectral stereotypy, speed and sequence regularity.

Having established the changes in IN properties in the first year post-hatching, I next quantified the song changes with age in the same birds. Earlier studies have shown age-dependent changes in a number of song features (song spectral stereotypy, speed, number of motifs in the bout, sequence

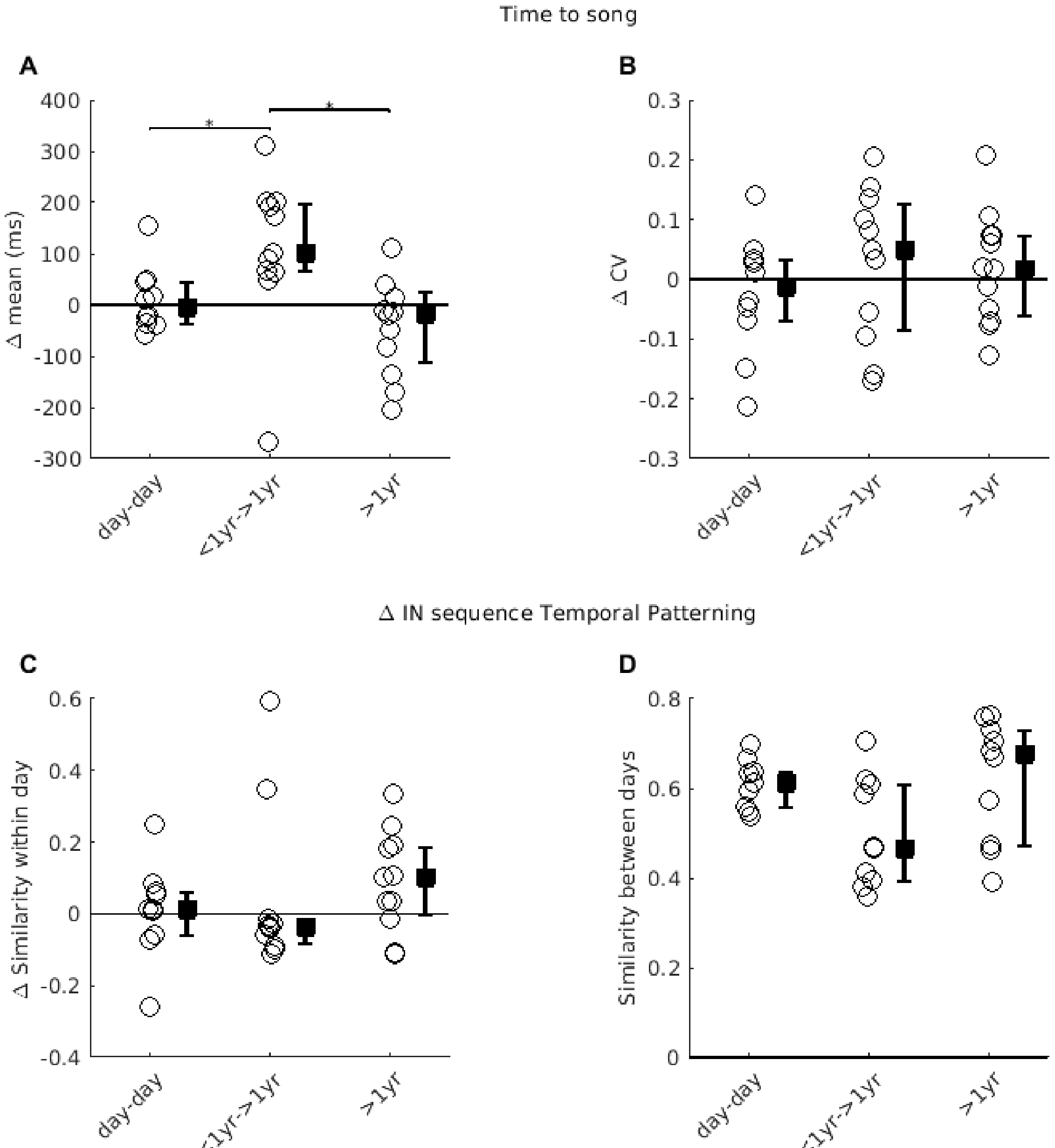

Fig. 3.14 Change in mean and variability of IN temporal sequence with age

Temporal sequence of IN can be represented by time to song (time from onset of first IN to onset of first song syllable measured in milliseconds) and temporal similarity of pattern of intervals across trials. (A) Change in mean time to song. (B) Change in CV of time to song. (C) Temporal similarity of IN pattern across trials was measured and the change was calculated between days. (D) Temporal similarity between trials from pairs of days. All pairs of days in (A)- (D) are the same pairs of days selected for maximum change in IN number in Fig 3.9. Each circle represents data from one bird. Square and whiskers represent median and interquartile range for the groups. $p \leq 0.05$ for Kruskal-Wallis test was followed by Dunn's test. *$Qvalue > 2.39$.

consistency and sequence entropy (Brainard and Doupe, 2001; Pytte et al., 2007; Glaze and Troyer, 2013; James and Sakata, 2019) and I verified these results for my dataset.

3.3.4.1. *Age-related changes* in *song stereotypy was more pronounced for subsequent motifs rather than first motif of the bout*

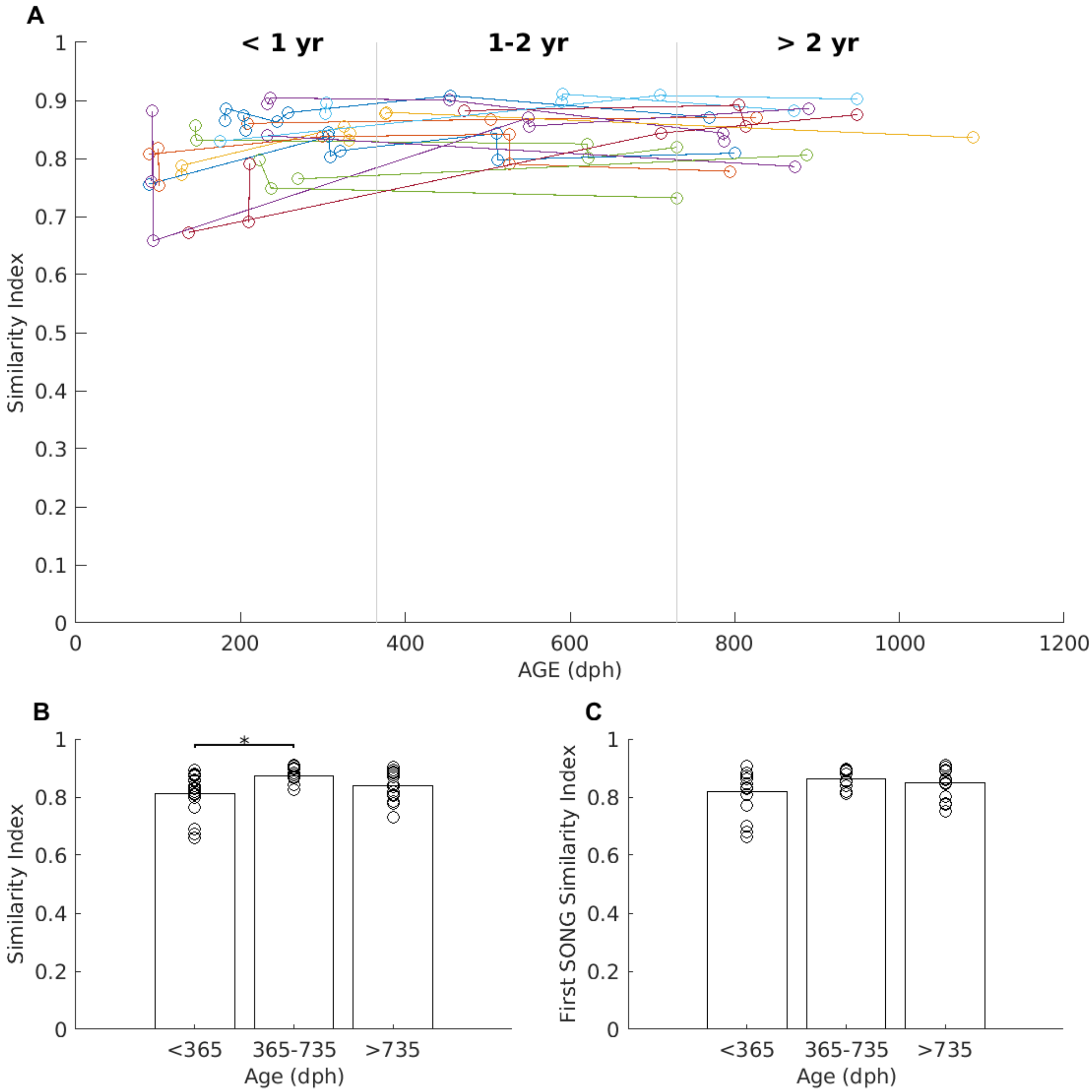

Fig. 3.16 Change in song stereotypy with age
(A) Similarity Index measuring similarity of within-bout motifs in a session vs. age of recording. Each circle represents similarity index from one session. Circles are connected by colored lines representing different birds. (B) Similarity Index of within bout motifs in (A) split into three age groups. All sessions from all birds were included. (C) Similarity index taking only the first motifs (SONG) in the bout split into age groups like in (B). $p \leq 0.05$ for Kruskal-Wallis test was followed by Dunn's test. **Qvalue*>2.39.

Song stereotypy is defined as the degree of similarity between songs across renditions. Using the songs recorded for each day, I calculated the similarity index as a measure of song stereotypy for each day recorded in each bird (see METHODS 3.2.2.8: within-day similarity). Previous results have shown that within-day song similarity increases in the first year, but the data used only the motifs within a bout (Pytte et al., 2007). I obtained similar results for similarity index calculated using within-bout motifs ($p<0.05$, Kruskal-Wallis test). The similarity index was high throughout for all the ages recorded across birds, showing that adult birds have highly stereotyped songs (Fig 3.16 A). Similar to previous studies, the average similarity index significantly increased from <1yr to 1-2yr (Fig. 3.16 B: <1yr 0.81±-0.004, 1-2yr 0.87±-0.0024). Thus, song stereotypy improved in the first year, consistent with earlier studies (Pytte et al., 2007).

As I wanted to compare IN properties in the beginning of the bout to the upcoming song, the similarity measurements were repeated using only the first motifs in the bout. Although the song stereotypy of first motifs showed similar trends of improvement for the first year, the difference was not statistically significant (Fig. 3.16C: <1yr 0.82±-0.004 , >1yr 0.86±-0.0026, p-value=0.22). The results suggest that age-related improvement in song stereotypy occur more prominently for within-bout motifs than first motifs in the bout.

3.3.4.2. Age-related changes to first motif in the bout

The changes in song properties with age are reported for first motif in the bout henceforth. Similar to analysis on IN changes and to ensure that one bird contributes to only one data point in an age group, the change in song property was measured for the same pairs of days as those analyzed for extent of change in IN number. Consistent with the above results, the change in within-day song similarity or song stereotypy of the first motif across bouts did not show significant changes with age (Fig. 3.17 A). Although, there is no change in song stereotypy, the song itself may show changes. Consistent with this, the across-day similarity (see METHODS 3.2.2.8) between pairs of days was significantly lower between <1yr and >1yr (Fig. 3.17 B). Other than spectral stereotypy measured by the similarity index, motifs also show temporal similarity defined as the stereotypy in the temporal pattern of syllables and silence within the sequence. The degree of temporal similarity of songs across bouts within a session did not show significant age-related changes (Fig. 3.17 C). However, the across-day temporal similarity did reduce significantly for pairs of days <1 yr to >1yr (Fig 3.17 D, $p<0.05$, Kruskal-Wallis test). The results show that changes to temporal pattern of the first motif occurred in the first year, although within-day stereotypy across bouts was independent of age.

Along with spectral and temporal changes the duration of song is known to get faster and less variable with age. The duration of first motif in the bout got faster in the first year (Fig. 3.17 F), however no significant age-related changes were observed in song duration variability (Fig. 3.17 E).

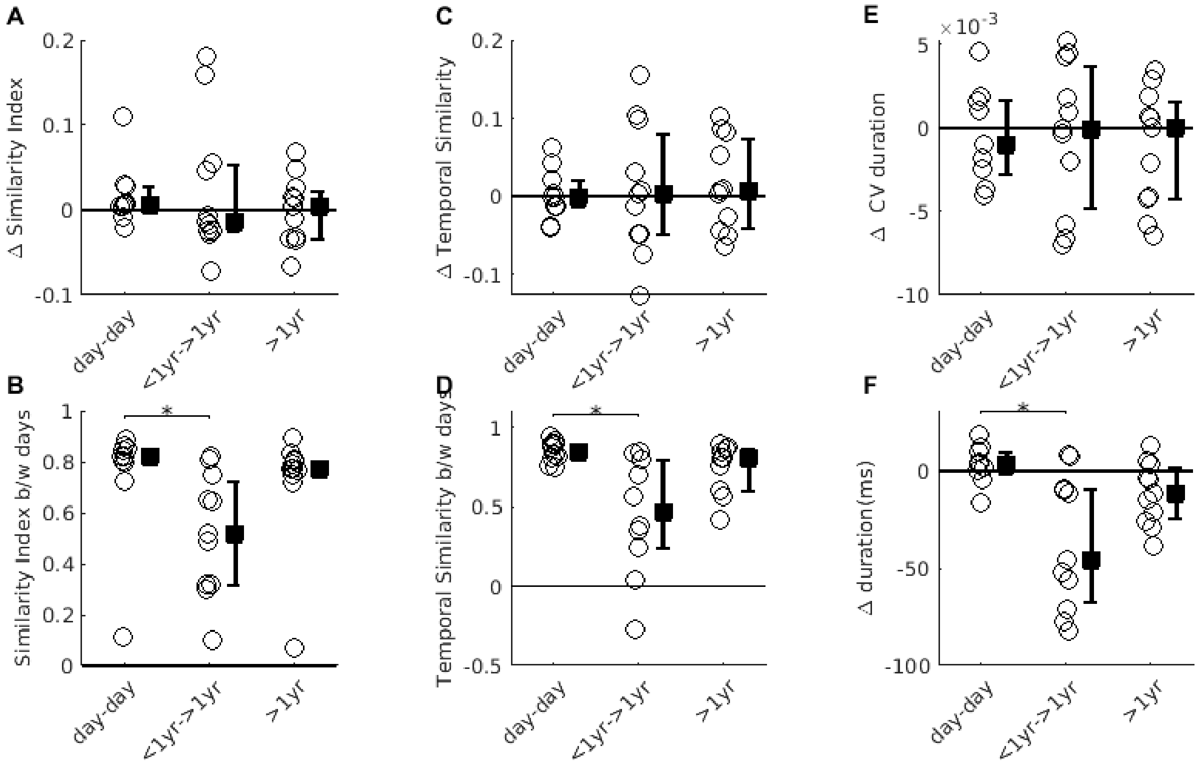

Fig. 3.17 Change in motif properties with age
(A)-(F) represent motif properties reported to change in the first year of age. Change in song variability between days represented by change in within-day similarity index (A) and similarity of motif temporal pattern or temporal similarity (C). Change in song between days is represented by similarity index (B) and temporal similarity (D) measured between songs from pairs of days. Change in mean (F) and variability (F) of motif duration (measured in milliseconds) between days. All pairs of days in (A)- (F) are the same pairs of days selected for maximum change in IN number in Fig. 3.9. Each circle represents data from one bird. Square and whiskers represent median and interquartile range for the groups. $p \leq 0.05$ for Kruskal-Wallis test was followed by Dunn's test. **Qvalue*>2.39.

3.3.4.3. Age-related changes following first motif of the bout

Other than changes within the motif, other reported changes in the first year of age include an increase in number of motifs in the bout and improved sequence regularity. I observed similar trends for the number of motifs, although the increase was not statistically significant (Fig. 3.18 A, p=0.08, Kruskal-Wallis test). The variability in the number of motifs did not show age-related changes (Fig. 3.18 B). I took only the first motif for measuring sequence consistency and sequence entropy. The sequence consistency did not show any age-related effects showing that the first motif sequence is largely consistent throughout adulthood (Fig. 3.18 C). However, the sequence entropy did show a decrease from <1yr to >1yr and was significantly different from other groups (Fig 3.18 D). Taken together, this shows that the transition from the end of first motif to the next syllable did become more regular with age, similar to earlier reports.

3.3.4.4.

3.3.4.5 Age-related changes in song syllable acoustics

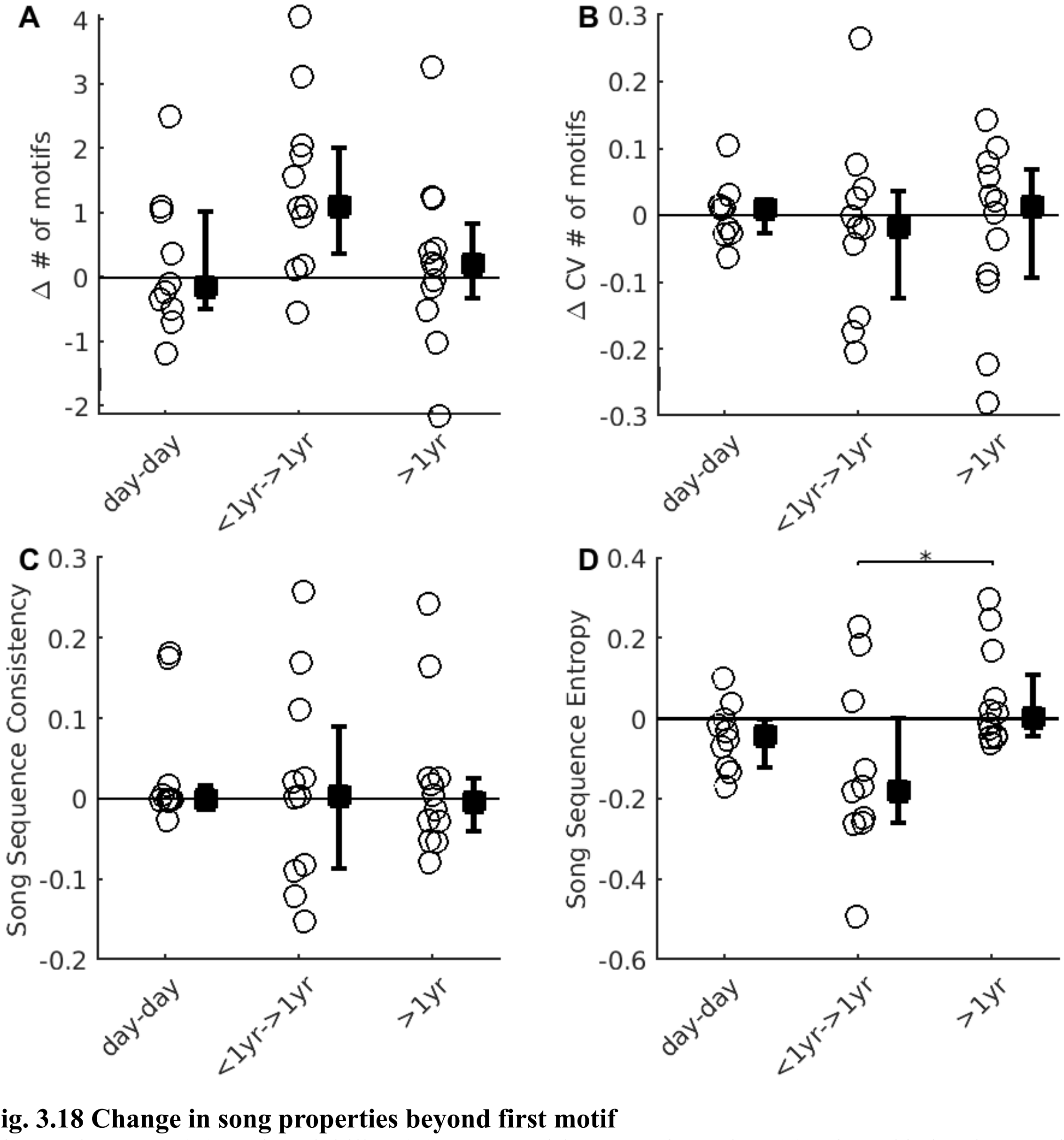

Fig. 3.18 Change in song properties beyond first motif
Change in mean (A) and variability (B) measured by CV of number (#) of motifs in about. Change in sequence consistency (C) and sequence entropy (D) for sequence consisting of first motif syllables and one syllable after first motif. All pairs of days in (A)- (D) are the same pairs of days selected for maximum change in IN number in Fig. 3.9. Each circle represents data from one bird. Square and whiskers represent median and interquartile range for the groups. $p \leq 0.05$ for Kruskal-Wallis test was followed by Dunn's test. **Qvalue*>2.39.

Similar to comparisons within a day for elements of IN sequences and song, I analyzed age-related changes in song elements, specifically associated with the mean (Fig 3.19 A-I) and variability (Fig 3.20 A-I) of the interval between first and second syllable in the motif and for the first syllable of the first motif. The change in mean and variability of song interval was not significantly different across groups (Fig. 3.19 A and Fig. 3.20 A). Among all the acoustic properties, mean pitch goodness (Fig 3.19 F) of the first syllable increased significantly in the first year.

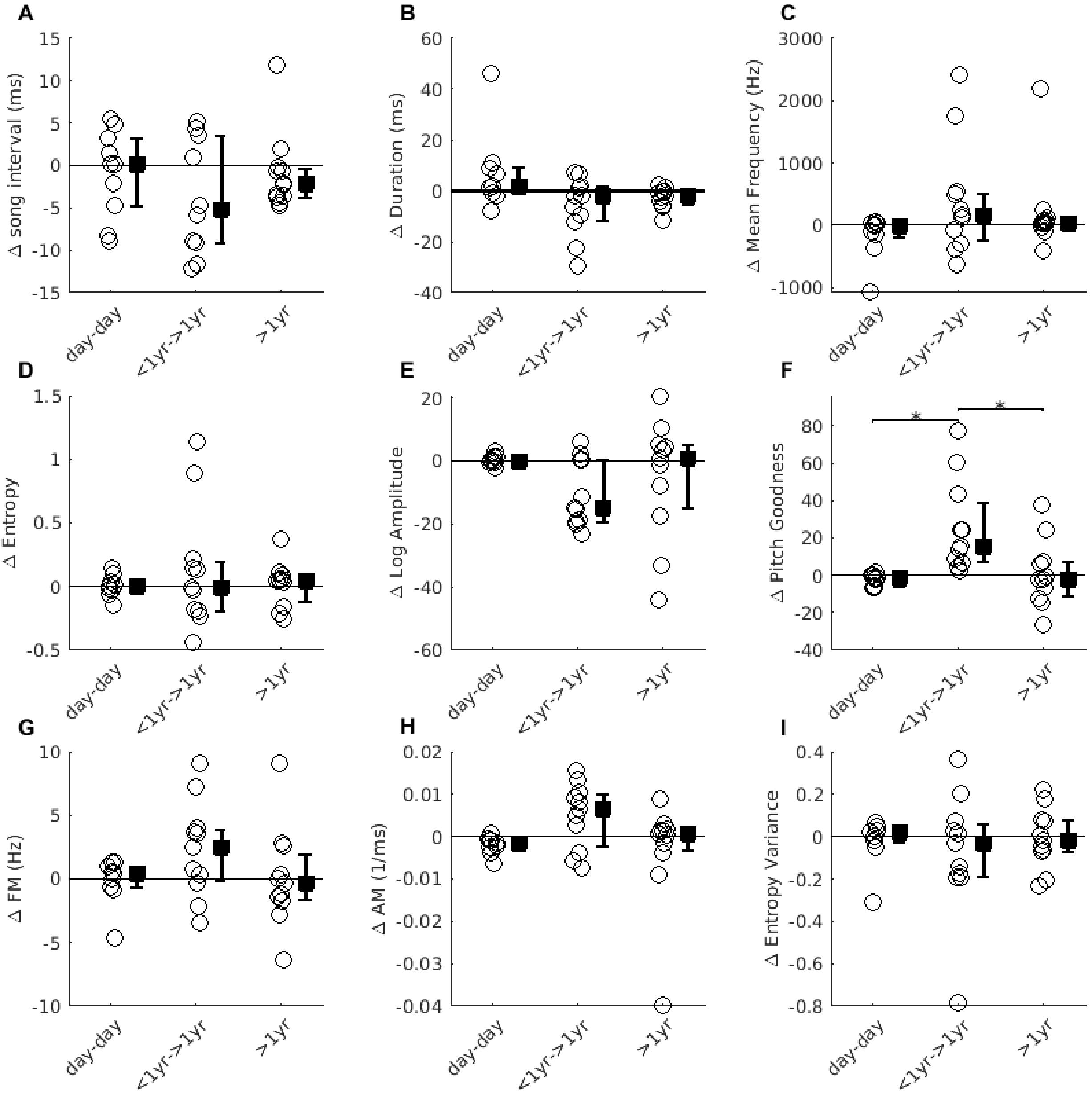

Fig. 3.19 Change in mean acoustic properties of song syllable with age
(A)Change in mean of 8 acoustic properties for song syllable between pairs of days (A-H). All pairs of days in (A)- (H) are the same pairs of days selected for maximum change in IN number in Fig 3.9. Each circle represents data from one bird. Square and whiskers represent median and interquartile range for the groups. $p \leq 0.05$ for Kruskal-Wallis test was followed by Dunn's test. **Qvalue*>2.39. Units: Duration – ms, mean frequency – Hz, log amplitude – dB, frequency modulation (FM) – degree, amplitude modulation (AM) – 1/ms, no units for entropy, pitch goodness and entropy variance.

In addition, the variability of frequency modulation decreased (Fig 3.20 G) and the variability of amplitude modulation increased (Fig 3.20 H) significantly in the first year ($p<0.05$, Kruskal-Wallis test).

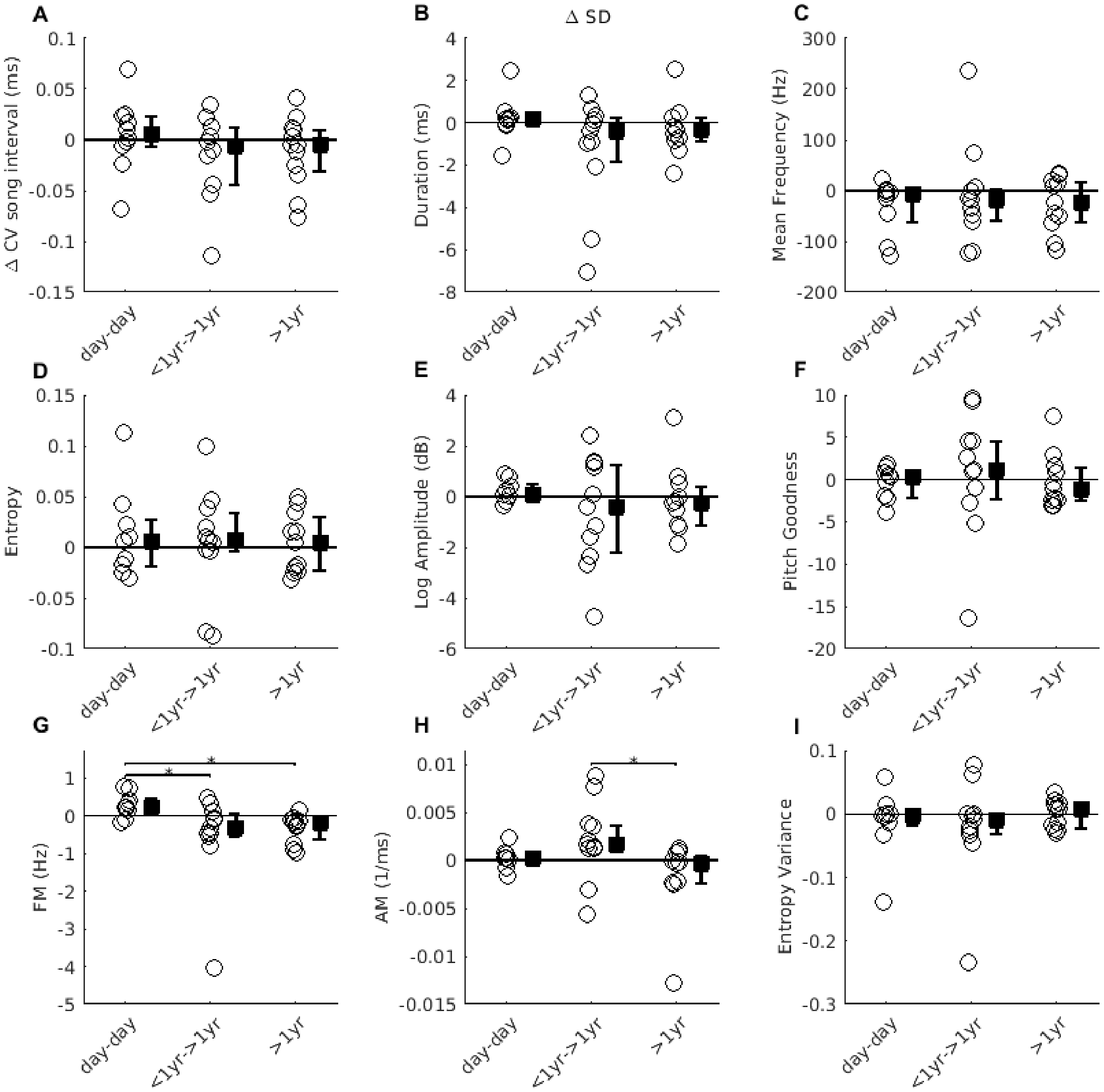

Fig. 3.20 Change in variability of acoustic properties of song syllable with age
(A)-(H) Change in variability of 8 acoustic properties for song syllables between pairs of days, given by CV (A-H). All pairs of days in (A)-(H) are the same pairs of days selected for maximum change in IN number in Fig3.3.4. Each circle represents data from one bird. Square and whiskers represent median and interquartile range for the groups. $p \leq 0.05$ for Kruskal-Wallis test was followed by Dunn's test. **Qvalue*>2.39. Units: Duration – ms, mean frequency – Hz, log amplitude – dB, frequency modulation (FM) – degree, amplitude modulation (AM) – 1/ms, no units for entropy, pitch goodness and entropy variance.

Although all acoustic properties of first syllable did not change significantly in the first year, the change between properties are correlated to each other, ranging from weak to moderate correlations (Fig 3.21: weak correlations: blue squares, moderate correlations: green). These also include correlations with motif and other higher order sequence properties that changed significantly with age, suggesting interdependence of song properties. For instance change in song duration was correlated to change in first syllable duration and interval, even though individually these did not significantly change with age.

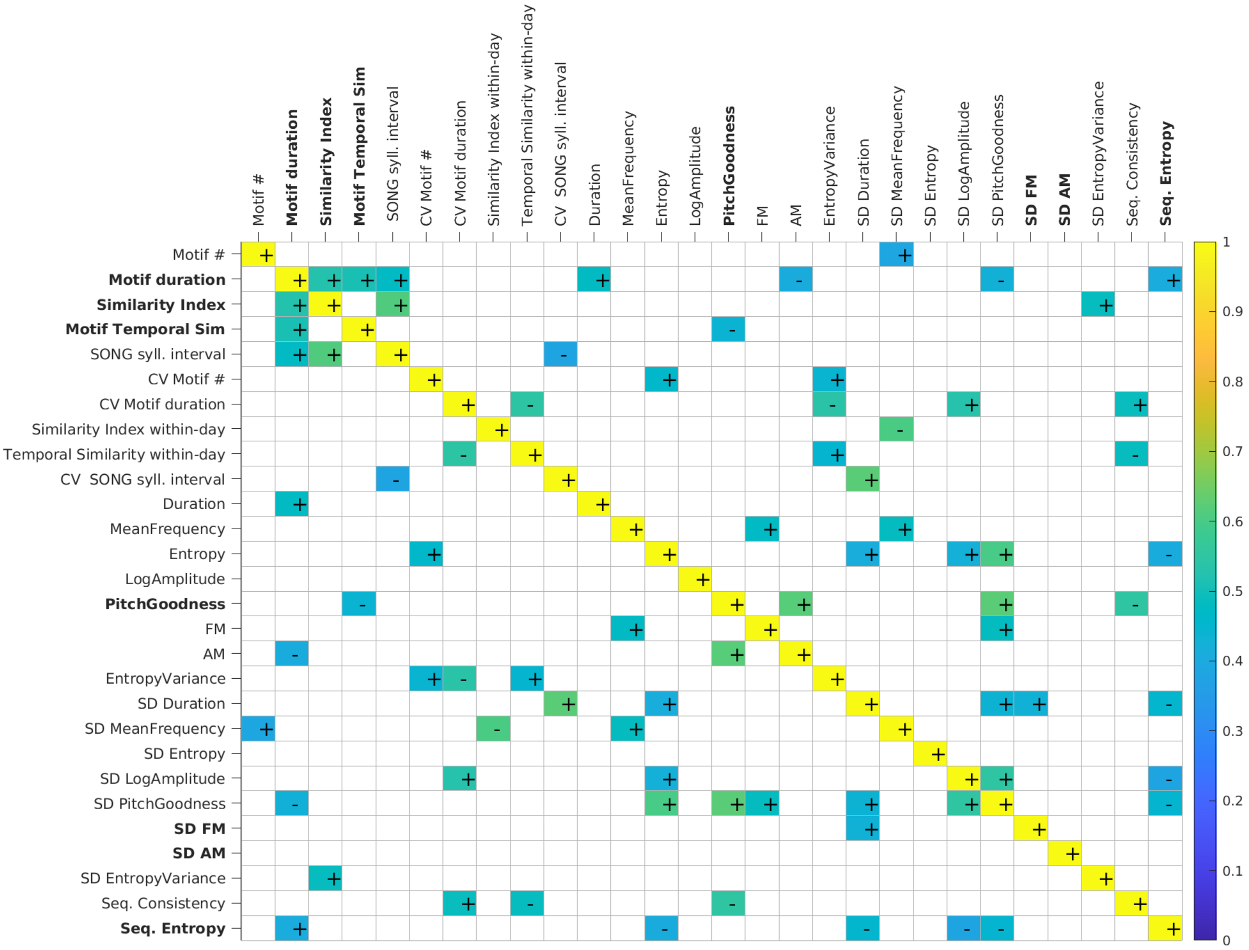

Fig. 3.21 Correlation matrix for change in song properties with age
Correlation matrix representing change in all song properties with one another. The colored cells represent significant correlations. The color scale represents the magnitude and the sign inside the cells mark the direction of correlation. The pairs of days to calculate change for all properties are the same pairs of days selected for maximum change in IN number in Fig. 3.9. Colored cells $p \leq 0.05$, Pearson's correlation coefficient.

3.3.5. Age related changes in INs and song properties are correlated

Finally, to understand the relationship between IN and song properties, I compared the age-related changes in IN properties to age-related changes in song (Fig 3.22) between the same pairs of days selected earlier. A range of correlations were obtained across properties, given the high interdependence of song and IN properties within themselves (colored squares in Fig. 3.22). These correlations further suggest relationship between IN and song properties that may occur indirectly, as age independently influences control of IN and song. Alternatively, INs have been hypothesized to prepare the zebra finch brain for producing song. This hypothesis would predict correlations between INs and song in an age-independent manner. To distinguish between these two possibilities I picked the song and IN properties that changed significantly with age (Fig. 3.23) and analyzed whether these properties are linked on a given day, independent of age.

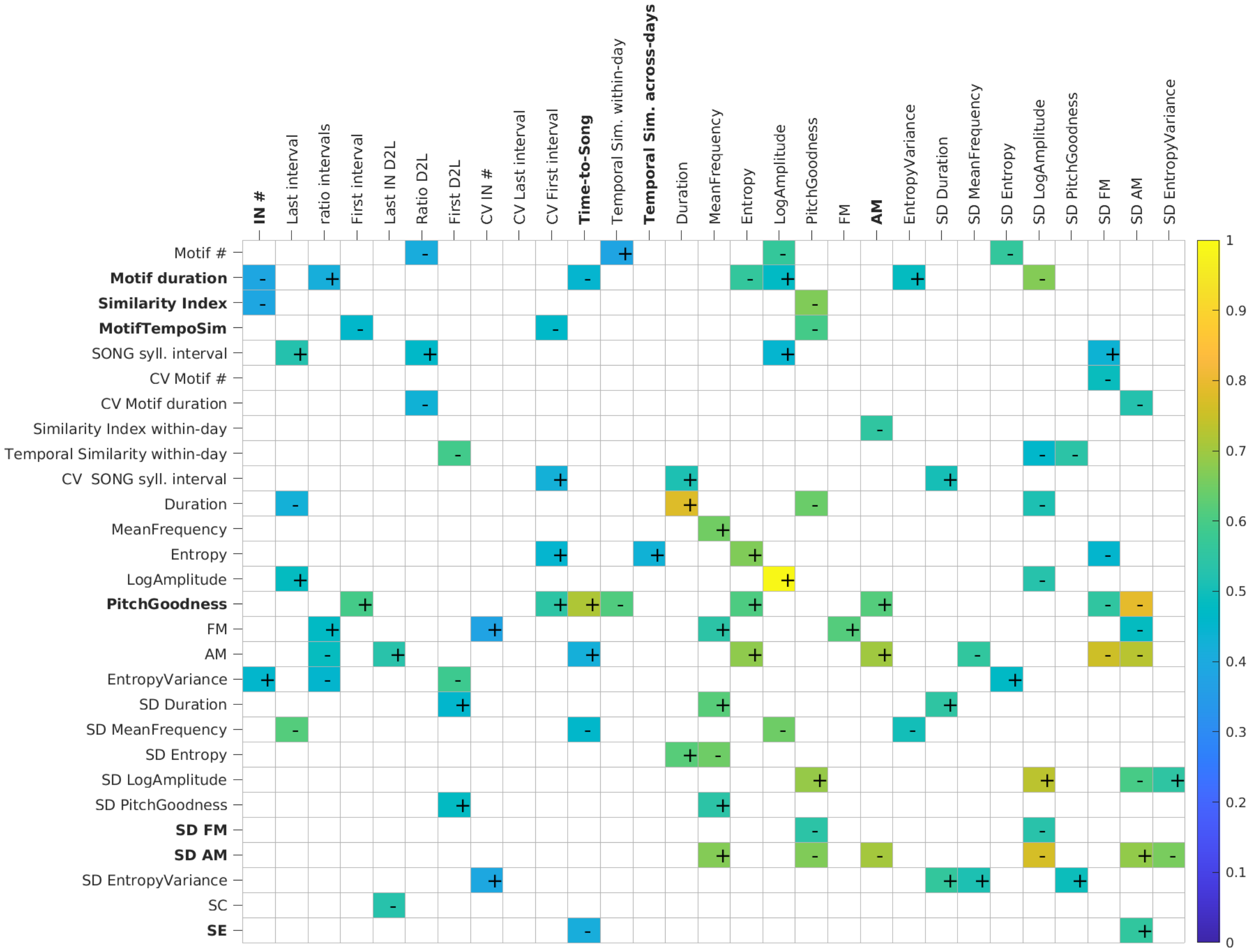

Fig. 3.22 Correlation matrix for change in song vs change in IN with age
Correlation matrix representing change in all song properties with change in all IN properties. The highlighted properties are the ones that significantly change with age. The colored cells represent significant correlations. The color scale represents the magnitude and the sign inside the cells mark the direction of correlation. The pairs of days to calculate change for all properties are the same pairs of days selected for maximum change in IN number in Fig. 3.9. Colored cells $p<\leq 0.05$, Pearson's correlation coefficient. SD – standard deviation, CV – coefficient of variation, FM – frequency modulation, AM – amplitude modulation, SC – sequence consistency, SE – sequence entropy.

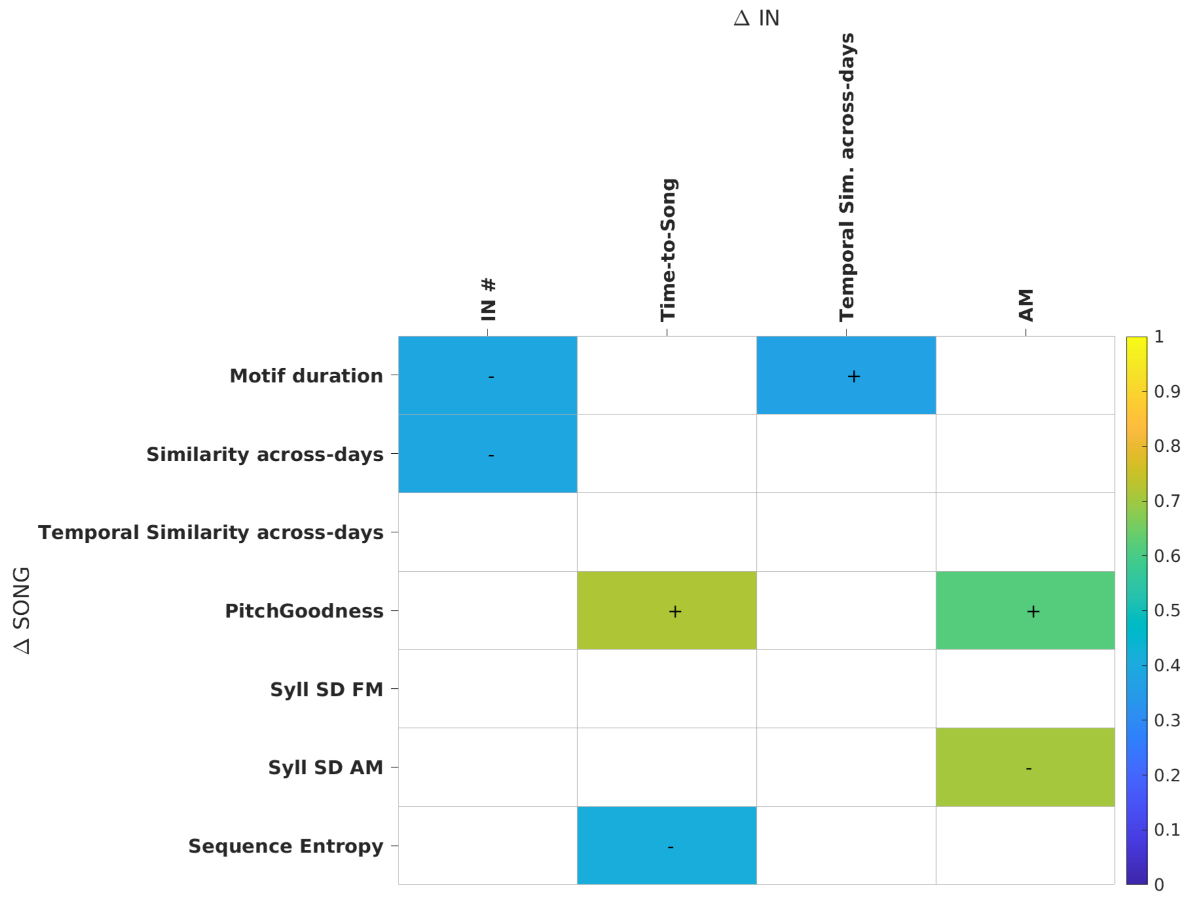

Fig. 3.23 Correlation matrix for change in song vs change in IN with age
Correlation matrix representing song properties with IN properties that significantly change with age. The colored cells represent significant correlations. The color scale represents the magnitude and the sign inside the cells mark the direction of correlation. The pairs of days to calculate change for all properties are the same pairs of days selected for maximum change in IN number in Fig. 3.9. Colored cells $p<0.05$, Pearson's correlation coefficient. SD – standard deviation, FM – frequency modulation, AM – amplitude modulation.

3.3.6. Temporal properties of INs are not linked to song timing, acoustics or sequencing.

The significant age-related changes that were correlated between IN and song included temporal, acoustic and sequence related properties. Each of these relationships were analyzed for predictive trends within a session. For example, the increase in number of INs with age was correlated to a decrease in song duration and a change in temporal pattern of song. This correlation could be an age-dependent change in both features. Alternatively, this could simply be a result of a correlation between IN number and song duration in individual sessions also; i.e. higher IN numbers would occur before shorter songs and lower IN numbers would occur before longer songs in individual sessions too. I compared whether more number of INs in a given session was related to shorter songs and whether songs with larger differences in IN numbers had lower similarity in temporal pattern (Fig. 3.24 A, B). Motif durations were similar across different number of INs on all days in a given bird (Fig. 3.24 A, p>0.05 Kruskal-Wallis test). This trend was consistent across birds. Similarly, the difference in number of INs was not correlated with similarity of temporal patterns between songs in majority of the birds (Fig 3.24 B – sample bird one session, significant correlations in n=2/19 birds, R^2 range [0.09-0.13]). These results show that the number of INs is independent of the temporal properties of the upcoming song.

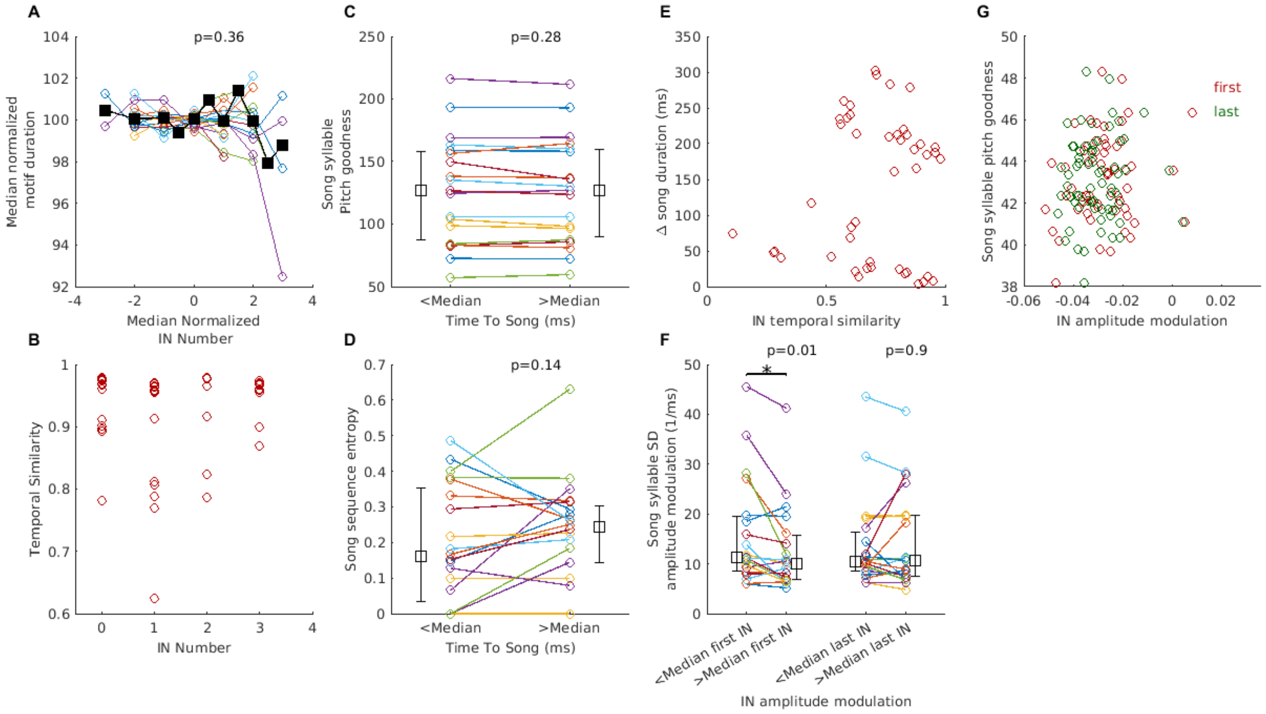

Fig. 3.24 Correlated IN and song changes with age are not related on given day
(A) Motif duration (normalized by median) on x-axis vs. IN number (median subtracted) on y-axis. Different colors represent individual bids and the mean across birds is represented by black squares. Each circle is mean of normalized motif durations across days for the corresponding IN number. Black squares represent median across birds. (B) Sample bird one session showing similarity between pairs of songs vs difference in IN number. (C)-(D) Mean pitch goodness and sequence entropy corresponding to bouts with shorter (<Median) and longer (>Median) time to song respectively. (E) Sample bird one session showing change in IN song duration vs. temporal similarity of IN between pairs of sequences. (F) Standard deviation of amplitude modulation corresponding to bouts with lower (< median) and higher (>median) amplitude modulation of first IN syllable or last IN syllable. (G) example bird, one session showing pitch goodness of first song syllable vs amplitude modulation of first (red) or last (green) INs. (C), (D), (F) Circles joined by lines are from one session of recording in a bird and the squares and whiskers represents the median and interquartile range across birds. (A) p-value >0.05 Kruskal-Wallis test. (B), (E), (G) Pearson's correlation coefficient was measured. (C), (D), (F) $*p \leq 0.05$, Wilcoxon signed-rank test.

The increase in time-to-song with age was related to increased pitch goodness of first song syllable and reduced sequence entropy of the first motif. I compared whether songs with shorter (less than median) vs., longer (greater than median) durations of time to song were associated with difference in pitch (Fig. 3.24 C) or sequence entropy (Fig. 3.24 D). Taking one session across birds, these properties were not significantly different for the two types of time to song (C, D, $p>0.05$ Wilcoxon signed-rank test). These results show that the duration of IN sequence does not influence the acoustics or the sequencing of the upcoming song.

The reduced similarity between temporal pattern of INs with age was correlated to faster song durations. In individual sessions, I compared whether lower temporal similarity of INs between pairs of bouts was related to larger difference between upcoming song durations. Comparing one session each across birds, the correlations were not significant in a majority of the birds (Fig 3.24 E – sample bird one session, significant correlations in $n=8/19$ birds, R^2 range [0.09-0.93]). The results show temporal patterning of INs is independent of upcoming song durations.

In addition to correlation in corresponding amplitude related properties, correlation between age-related changes in IN and song syllable was observed across different acoustic properties. The increase in amplitude modulation of INs was related to decrease in variability of amplitude modulation and increase in mean pitch-goodness of the first song syllable. I compared separately for both first and last INs, whether lower (less than median) vs. higher (greater than median) amplitude modulations were related to difference in pitch goodness (Fig. 3.24 G) or variability of amplitude modulation of upcoming song syllable (Fig. 3.24 F). While amplitude modulation and pitch goodness were not related (Fig. 3.24 G – sample bird one session, significant correlations in $n=1/16$ birds, $R^2=0.038$]), higher first IN amplitude modulation was related to lower variability in amplitude modulation of first song syllable (Fig 3.24 F). Although this trend was not present for the last IN, it could be explained by relatively low variation in properties of last IN. These results suggest only corresponding acoustic properties related to amplitude are related between IN and song syllable across trials and may thus share common control on a trial to trial basis.

Together, these results suggest that the number and timing of INs do not predict temporal, acoustic or sequencing properties of the upcoming song. However, corresponding acoustic properties between IN and song related to amplitude may be dependent on each other. Thus, the control for timing of INs is distinct from that of the song, while syllable acoustics for INs and songs may share control.

3.3.7. The temporal progression of INs is unrelated to repetitions within and outside song

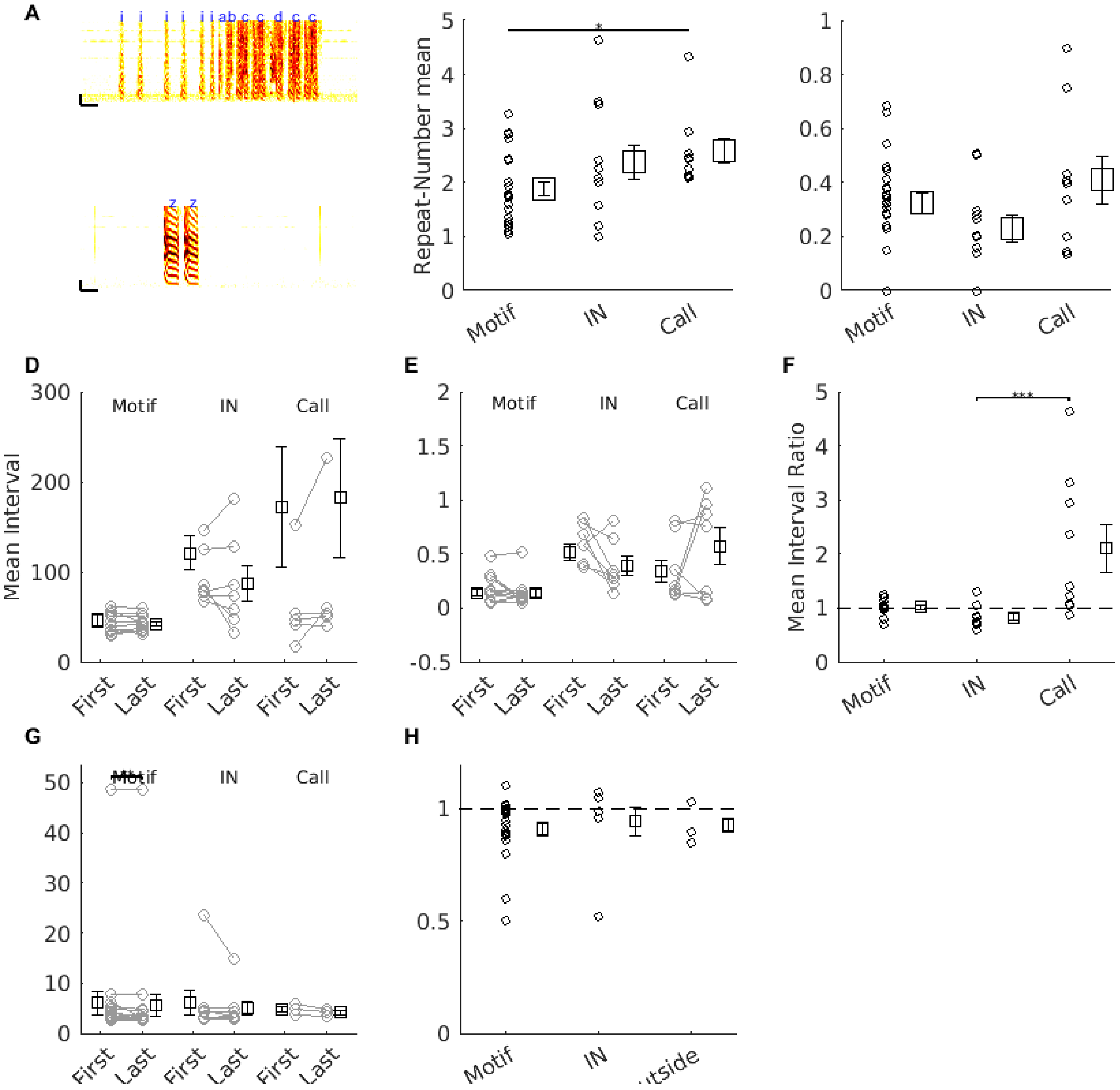

Fig. 3.25 Temporal progression of INs is not present in other repeats
(A) Top – Example spectrogram of a bird that produces repeating INs ‘i’ and repeating syllable ‘c’ within the motif. Bottom – example of a bird producing repeats of call syllables outside song. (B) and (C) Mean and CV of number of syllable repeats for motifs, INs and calls. (D) and (E) Mean and CV of first and last interval (measured in miliseconds) within the repeat sequence in case of motifs, INs, calls. (F) Ratio of intervals vs type of repeating syllable. (G) Acoustic distance with respect to the last syllable distribution for first and last position of repeats. (H) Ratio of acoustic distance for successive INs. (B), (C), (F) and (H) Circles represent individual birds and squares and whiskers represents the mean and sem across birds. ****p* value<0.005, Kruskal-Wallis followed by rank-based Tukey-Kramer test. (D), (E) and (G) each line joining circles represent one repeating syllable from a bird. Squares and whiskers represent the mean and SEM across birds, p-value >0.05 Wilcoxon signed-rank test.

An important caveat with the above comparisons of INs with song syllables is the fact that INs are repeating syllables. As song syllables in zebra finch are largely devoid of repetition, it is possible that timing of INs are a product of repetition. To test this possibility, I looked at songs of subset of birds in colony that repeated a syllable within song and asked whether repetition was similarly controlled between INs and songs. The properties of INs were compared to repeating song syllables (19 syllables in 17 birds) and also call-syllables that tend to repeat outside song (Fig. 3.25 A – example spectrograms). The number of repeats for syllables that occurred within song were significantly lower than that during calls (Fig. 3.25 B). Calls and INs produced similar number of repeats. The variability in number of repeats (Fig. 3.25 C) was not significantly different between syllable types. This suggests a similar control of number variability for all types of repetitions. The mean (Fig. 3.25 D) and variability (Fig. 3.25E) of intervals at first and last position within repeating syllables was not different for motif and call syllables. The ratio of intervals was above 1 for calls and was significantly different from INs. The ratio of intervals for repeating syllables within motif tended to be around 1 although not significantly different from INs (Fig. 3.25F). Together, this suggests that the temporal progression is present only in the case of IN repeats. However, similar to INs, the acoustic distance of repeating syllables improved significantly from first to last for motif syllable (Fig. 3.25 G) and also the ratio of acoustic distance was maintained at less than one (Fig. 3.25 H). Similar trends were observed for call repeats. These results suggest acoustic progression is present in all repeating syllables, i.e. the acoustic structure of the syllables changes with consecutive repeats independent of whether the syllable is an IN or a song syllable or a call. Taken together, the data support that number variability and acoustic progression may be controlled similarly during INs, songs and other repetitions outside song, however the timing between INs is a distinct property not correlated to repetition and possibly driven by a different control mechanism.

3.4. DISCUSSION

In this study I set out to test various hypotheses on the relationship between IN and song. INs may be produced in order to prepare for the upcoming song such that the properties between IN and song sequence are related to each other. Further, INs could be unrelated to song serving a distinct function, similar to an alerting function shown in other songbirds (Richards, Douglas, 1981). Alternatively, INs could be produced just as another syllable in the sequence such that INs are as related to the song as other syllables in the song sequence. I show that IN and song are correlated in their syllable acoustic properties. However, the variable number and timing of INs do not correlate to the properties of the upcoming song. The properties related to number, timing and acoustic properties of INs change with age in the first year, systematically correlated to changes in post-learning consolidation of song. This showed that age influences the control for both INs and song. Together, the results show that the number and timing of INs are unrelated to song and do not contribute to motor preparation. Further, the results also implicate a shared control of syllable acoustic production during INs and song, but an independent control of the timing of INs that speed up before song.

3.4.1. Shared control and predictive relationship between IN and song

The correlation between acoustic properties of INs and song syllables across trials suggest a relationship between IN and song. These acoustic properties included mean frequency, entropy, log amplitude and amplitude modulation all of which are amplitude-dependent properties and may be influenced by amplitude related artifacts in recordings. Similar correlations were observed for birds with head-fixed microphone, that controlled for amplitude-related artifacts. This verified the correlation in acoustic properties of INs and songs. As both first and the last INs show similar correlations to song syllables, it suggests that the predictive nature of what the song is going to be gets established at the start of the IN sequence. This is consistent with the idea that the acoustic progression of INs as it repeats is associated with motor preparation. However, it is also possible that IN syllable share properties to song syllables by virtue of being a part of the sequence. In support of this we observed that the relationship of INs and first song syllable was similar to relationship between first and second song syllable, suggesting all syllables have common sequence related properties. However, the current study implies correlations and lack evidence for a causal relationship between IN and song. Future experiments to causally test the relationship between acoustic properties of INs and songs may involve manipulating the song syllable amplitude by playback and analyzing changes in IN and song properties.

3.4.2. Differences in the control of INs and songs

The number and temporal properties of song were not correlated to properties related to mean or variable properties of the upcoming song. This goes against the motor preparation hypothesis that implicates that the control related to the upcoming movement gets established earlier on. The number and first interval between INs were shown to relate to the time for song initiation (Rajan and Doupe, 2013; Rao et al., 2019). Further, proximity of calls within <200 ms of INs shorten the first IN intervals and decrease the time for song initiation (Chapter 2 – same as (Rao et al., 2019). It is possible that these properties are involved in determining when the song should start independent of what the parameters of song are going to be. Together, the results suggest that INs (or any other syllables) may reflect some on-going internal process related to time.

3.4.3. Age related changes of INs

Properties of INs showed changes in the first year of age. Although not all properties showed a significant change in the <1yr to > 1yr category, there were trends in some other properties too and greater variability in the data for this group could underestimate changes. Another caveat that could affect the results was the error in age estimation of outside bought birds. As mentioned in the methods (3.2.1), the exact age of these birds were not known and assumed to be 60dph on the day of purchase. Among 6 birds used for age-related analysis only 3 birds contributed to <1yr to >1yr category (3/11). Further, excluding these birds from the estimation of change in IN number did not affect the results – the trends were the same (data not shown) verifying IN changes in the first year hold true for lab bred birds as well.

This is the first time properties of INs have been tracked long-term with age. On the one hand, the results show that the properties of INs are remarkably stable with very little variation. However, similar to changes in the timing of song, small but significant changes in INs occur in the first year of age. These changes may be a result of both central and peripheral changes with age. Song changes with age are associated with "improvement" in song, we do not know what the change in number and temporal pattern of INs mean for the bird, as these are not related to properties of the song (Brainard and Doupe, 2001; Pytte et al., 2007; Glaze and Troyer, 2013). Motor preparation and movements like song get faster with practice and age (Brainard and Doupe, 2001). However, I show that the number and time to start song increases making these sequences longer. The changes in first interval showed a trend, although not significant. It suggests that initial state of IN progression but not the progression itself changes with age. The initial state changes may be related to inherent circuit changes themselves that may change aspects of how songs are initiated.

3.4.4. Age related changes to first song of the bout

Our analysis of age-related song changes differed with previous studies. Earlier work compared songs within-bouts or irrespective of bout positions (Lombardino and Nottebohm, 2000; Pytte et al., 2007). However, in this study I focussed on just the first song in the bout as I wanted to understand the immediate relationship between INs in the beginning of the bout and the upcoming song. Our results showed smaller changes to within-day spectral and temporal similarity that did not significantly change with age. First motifs may become more regular early on compare to the rest of the bouts (Pytte et al., 2007) and could be one reason for the results shown here. Taking only first motifs in the bout allowed us to directly compare INs in the beginning of the bout with the immediate song that follows. The speed of the song still changes with age and similarity dropped between pairs of days in the first year indicating song spectral and temporal structure may themselves change with age.

Overall, this study showed that the temporal progression is unique to INs and that the timing of INs might be controlled differently from song. However, the acoustic properties of IN syllables might share control with the upcoming song.

CHAPTER 4

Neural Correlates of Introductory Note Progression to Song

4.1. INTRODUCTION

Singing in zebra finches begins with variable repeats of introductory notes (INs) before the stereotyped song (Price, 1979; Sossinka and Böhner, 1980). As INs repeat in the sequence, a progression in temporal and acoustic properties accompanies its transition towards the song (Rajan and Doupe, 2013). Chapter 3 highlighted that the progression in temporal properties is a unique feature of IN repetitions, hinting at its potential importance in the (unknown) functional role of INs. What then drives the progression of INs to the upcoming song? Chapter 2 (same as (Rao et al., 2019)) showed that the overall progression of INs to the song occurs independent of real-time external sensory feedback, further indicating that the progression of INs is an internal process driven entirely by the brain. Although previous studies have reported several brain areas relevant for song that also show activity related to INs (Williams and Vicario, 1993; Yu and Margoliash, 1996; Leonardo and Fee, 2005), it remains to be understood how the brain represents the systematic progression of INs to the song. As the production of repeats of INs is also a motor behavior like song, I studied the IN related neural activity specifically in the areas of the song motor pathway, in order to identify the neural correlates of IN progression.

HVC (used as proper name) and RA (robust nucleus of the arcopallium) are two important premotor areas (Chapter 1 – Fig. 1.1) for driving the production of the stereotyped song (Nottebohm and Arnold, 1976; Nottebohm et al., 1976; McCasland and Konishi, 1981; Hahnloser et al., 2002; Leonardo and Fee, 2005), and have also been reported to show IN related activity (Yu and Margoliash, 1996; Leonardo and Fee, 2005; Rajan and Doupe, 2013). During the production of both INs and songs, the firing patterns of individual neurons in HVC and RA changes from that during silence (McCasland, 1987; Yu and Margoliash, 1996). During the song, it is known that the activity of neurons in HVC and RA show stereotyped firing patterns time-locked to the song (McCasland and Konishi, 1981; Yu and Margoliash, 1996; Hahnloser et al., 2002, 2015; Kozhevnikov and Fee, 2007; Vallentin and Long, 2015; Lynch et al., 2016). Specifically, the projection neurons in HVC (HVC_X – projection neurons from HVC to Area X and HVC_{RA} – projection neurons from HVC to RA (Nottebohm et al., 1982)), show sparse bursts at specific times in the song that is predictive of song sequence timing (Hahnloser et al., 2002). The interneurons in HVC that receive strong inputs from HVC projection neurons increase their firing rates to a structured firing pattern during the song (Mooney and Prather, 2005; Kosche et al., 2015). Moreover, the neurons in RA that also receive inputs from HVC projection neurons (HVC_{RA}) show characteristic dense bursts, again in sync with the song (Yu and Margoliash, 1996; Leonardo and Fee, 2005). Together, the reliable activity of these neurons during song represent the precise sequence of songs produced across renditions. Similar to song, the activity of a subset of these neurons in HVC and RA also changes (relative to silence) during INs (Yu and Margoliash, 1996; Leonardo and Fee, 2005; Rajan and Doupe, 2013). As INs repeat variably and switch to song, activity during INs in these areas may represent the progression of the IN sequence towards song. Several lines of evidence supporting the presence of activity changes during INs, as described below, provide the potential to further investigate whether representation of IN progression is present in the areas of the song motor pathway.

The progression of INs may be represented by neurons of HVC. A majority of HVC_X neurons that are active during INs and the interneurons show a different pattern of activity for the last IN relative to the earlier INs (Rajan and Doupe, 2013) The difference in activity is indicative of the start of the song. Whether this pattern of activity represents the end of a systematic progression, as INs progress to song, remains unclear.

The progression of INs may also be represented by neurons of RA. A small number of RA projection neurons recorded during INs show irregular burst timing (Yu and Margoliash, 1996). This is unlike song where activity of RA bursts are very precise in accordance with the song (Yu and Margoliash, 1996; Hahnloser et al., 2002). Further, activity of RA neurons in bengalese finches, a related species of song birds, were shown to be correlated to trial-to-trial variation in acoustic properties of song syllables (Sober et al., 2008). Together, these studies suggest that RA neurons may be associated with variable activity during INs to represent the variable behavioral properties associated with INs as they progress to song. Whether such systematic changes in RA activity occur with IN progression, has not been studied.

How do the changes in systematic neural activity during INs relate with progression of INs to song? From chapter 3, we know that a characteristic feature of IN progression is the change from long to short interval between INs. A possible hypothesis would be that the activity related to the interval between INs determine the timing of the next syllable. For example, increased activity during intervals may relate to faster onset of next syllable and in-turn resulting in shorter intervals. Thus, the systematic changes in activity may represent intervals getting shorter. Alternatively it is possible that changes in activity are independent of interval durations and represent the position of INs in the repeat sequence or the state of progression towards song. These related hypothesis were analyzed for neurons recorded in HVC and RA in order to understand the neural representation of IN progression.

The results show that the interneurons in HVC have interval-related activity. However, the activity at a given position does not change based on duration of the interval. Instead, the activity pattern of interneurons combining both IN syllable and the following interval was specific to a position in the repeat. The relative pattern of activity differed between INs at first, middle or last IN. Three categories in the activity pattern of interneurons were identified which potentially indicated identity of previous or next syllable: (i) neurons that fire similarly for all but the last IN, indicating the next syllable is an IN; (ii) neurons that fire similarly for all but the first indicating the previous syllable was an IN; and (iii) that fire only for the last indicating the next syllable is start of the song. Together with the known activity of HVC_X neurons that bursts during different position of INs, the three categories show that HVC may represent IN progression by keeping track of past and future syllables. Similar analysis in multi-unit activity of neurons recorded in RA showed consistent activity across position of INs, suggesting the absence of representation of IN progression in RA. Together, the results implicate a representation of IN progression in the song motor pathway, specifically in HVC. Further, this representation of variable INs hold similarity to the representation of variable sequences in HVC of other songbirds with complex songs, suggesting shared mechanisms within.

4.2. METHODS

Experimental procedures performed at IISER Pune were approved by the Institute Animal Ethical Committee in accordance with guidelines of the Committee for the Purpose of Control and Supervision of Experiments on Animals (CPCSEA, New Delhi, India). Experimental procedures performed at UCSF (CA, USA) were approved by the UCSF Institutional Animal Care and Use Committee in accordance with NIH guidelines.

4.2.1. Birds

All the birds recorded in this study were adult male zebra finches (>90 dph) either purchased from a local vendor or bred at IISER Pune (recordings in RA) or UCSF (recordings in HVC). Neural activity of live singing birds were analyzed for 2 birds recorded in HVC and 3 birds recorded in RA. Overall, a total of 14 birds were used for anesthetized or chronic recording in RA. The birds were housed in individual acoustic chambers for the duration of the experiment under a 14h (light):10h (dark) cycle, and food and water were provided *ad libitum*.

4.2.2. Details of overlap with previous studies

Songs and neural data in HVC were recorded as part of 3 different studies characterizing neural activity during song, INs and pre-song activity (Rajan and Doupe, 2013; Woolley et al., 2014; Rajan, 2018). The previous study that looked at IN related activity focused on changes in patterns of activity before the last IN and showed the presence of distinct activity patterns for the last IN. However, changes in activity during the earlier INs and the associated progression was not analyzed in detail. Here, I used this data to analyze the activity during the interval between INs and for further analysis on progression of INs.

4.2.3. Assembly of microdrives and tetrodes for extracellular recording

Microdrives similar to those used in previous studies were assembled for chronic extracellular electrophysiological recordings in zebra finches (Fig. 4.2). The body of the light-weight microdrives and the shuttles were custom-designed and 3D printed to hold 2-4 high impedance tungsten microelectrodes (FHC 8-10 MΩ), The shuttles housed custom-machined screws (pitch 212 µm) that allowed for the vertical movement of the electrodes. A custom-made screw-driver allowed quarter turn movements of the screw at a time. Lateral movement was also achieved with a lateral screw that pushed the electrode bundle laterally as described earlier (Fee and Leonardo, 2001; Otchy and Ölveczky, 2012). Across microdrives implanted in different birds, the tip of electrodes in the assembly were positioned to compensate for curvature of the skull, at the location of the implant. The electrodes of the 2-electrode implants were positioned along the anterior-posterior axis of the bird. The 4-electrode implant had electrode tips arranged as vertices of a square. A difference of 0.5 mm depth from the posterior-most to the anterior-most electrode was maintained. The vertical movement screw allowed a generous movement of 2-4 mm. The lateral screw was placed such that the electrode could be moved about 1 mm medially post-surgery. All the electrodes were connected through copper wires to a connector for plugging-in during recordings.

Separately, tetrodes were also built by twisting nichrome wires (SANDVIK ~ 13 micron diameter) using standard procedures. The tips were gold-plated in order to make their impedances in the range of 100-200 kΩ. A single tetrode soldered to a connector was plugged-in to the INTAN-headstage during acute recordings in anesthetized birds.

4.2.4. Surgery and electrode implantation

Surgery was performed in each bird to implant electrodes in RA. The bird was weighed and administered with oral analgesic (meloxicam 0.25 mg/kg) about an hour before surgery. Prior to the surgery, the bird was deeply anesthetized by intramuscular injection of ketamine (30 mg/kg), xylazine (3 mg/kg) and diazepam (7 mg/kg). Absence of response to toe pinch was used to assess the depth of anesthesia. The anesthetized bird was prepared for surgery by securing its legs with surgical tape, draping a warm jacket and trimming feathers around the ear. The bird was secured on the stereotaxic apparatus (Narshige, custom-designed) and the head was fixed in place by fixing the ear bars and beak holder on the apparatus. The beak angle was maintained at 45°. The surgery was performed under heat lamp to assist the anesthetized bird in regulating its body temperature. The feathers were plucked on the head and local anesthesia was injected (xylocaine or lidocaine 2%) under the scalp before cutting the skin with the help of microscissors and exposing the skull for surgery. A tetrode mounted on the micromanipulator (Narshige, SM 11) was positioned above the co-ordinates for RA (0.25 mm posterior and 2.43 mm lateral to the divergence point of the central sinus or Y sinus). These co-ordinates for RA were as obtained from standardization experiments performed in a separate set of birds (described in the next section 4.2.5), and are within the range of RA co-ordinates reported in the stereotaxic zebra finch atlas (Nixdorf-Bergweiler and Bischof, 2007). Using fine forceps and scalpel, the two layers of skull were removed at the implant site, followed by careful incision and removal of the dura above the surface of the brain. The tetrode was gradually advanced into the brain using fine screws of the manipulator and activity was recorded at various depths. A silver wire inserted on the contralateral side between the skull and dura served as the electrical ground. The tetrode was connected to the 16 or 32 channel INTAN amplifier (RHD2132) for extracellular recording. Digitized signals from amplifier were acquired by the INTAN evaluation board (RHD2000) that connected to the computer. The INTAN software-interface was used for online viewing of signals and saving recorded files. The bird set-up was covered with a modified bird cage to serve as Faraday cage for acute recordings. The activity was recorded at different depths to locate positions with spontaneous regular/ tonic firing in the range of 10-30 Hz, characteristic of RA activity (Mooney, 1992; Hahnloser et al., 2002; Leonardo and Fee, 2005; Spiro et al., 2015). RA was successfully located at the standardized co-ordinates in 8/9 birds implanted with electrodes. The tetrode on the manipulator was then replaced with the microdrive and the posterior most electrode in the implant was lowered at the same co-ordinates as the tetrode. The electrodes were advanced until all electrodes were in the brain and the posterior most electrode was just above the depth where RA activity was identified with the tetrode (Depth RANGE 1.7-2.73, MEDIAN 2.33). The poly-imide tubing that holds the electrodes and the exposed brain at the implanted site were covered with mineral oil or bone-wax before applying dental cement. This was done to prevent access of dental cement to electrodes or the brain. The microdrive, along with the ground wire on the contralateral side, and the connector for recording were then fixed on the skull with dental cement. For further reinforcement of the dental cement on the skull, broken pieces of tungsten-electrode-material were inserted tangentially at 2-3 places through holes in the first layer of skull as described previously (Okubo et al., 2014). These holes were already made in the beginning of the surgery. In all birds, the implant was fixed on the left hemisphere. The bird was allowed to recover from anesthesia and monitored for signs of pain and discomfort.

Similar procedures were performed to implant electrodes in HVC, the details of which were described in previous studies that also analyzed this data (Rajan, 2018).

4.2.5. Standardizing stereotaxic co-ordinates for RA

Standardization procedures were performed to identify co-ordinates that can reliably locate RA across birds using the stereotaxic (Narshige) setup in the lab. Tetrodes were built as described previously in literature and acute recordings were carried out in anesthetized birds. Briefly, birds were anesthetized and prepared for surgery as described in the above section. Using the stereotaxic co-ordinates from the zebra finch atlas (Nixdorf-Bergweiler and Bischof, 2007), an area of brain above RA was exposed. Tetrodes were lowered and activity was recorded at multiple depths (RANGE 1-3.5 mm). Multiple penetrations were performed and activity was recorded for a range of co-ordinates across 5 birds and 7 hemispheres (AP [-0.25 to +0.4] and ML [2.25 to 2.52]). The beak angle was 45° for 2 birds and 3 hemispheres, and 80-90° for 3 birds and 4 hemispheres. Spontaneous, highly regular activity 20-30 Hz is characteristic of RA (Fig. 4.1 A). Such locations were identified, and then electrolytic lesions were performed (100 uA for 100 seconds) at 1-3 depth locations. The birds were perfused, and standard histological procedures were performed to verify if the lesions in the brain were within anatomically identified nucleus of RA (Fig 4.1 B). Co-ordinates of 0.25 mm posterior and 2.43 mm lateral to the Y sinus reliably identified RA in birds recorded at 45° beak angle. For all further surgeries of chronic recordings, one of the multiple electrodes of the microdrive was implanted at these co-ordinates (8/9 birds) after verifying “RA-like” spontaneous activity during surgery. As songs were damaged in the first two birds that were implanted, a different beak angle at 80-90° was attempted to locate RA, but these were unsuccessful in finding “RA-like” activity.

4.2.6. Song and neural recording

After surgery, the birds were housed in individually isolated acoustic chambers and allowed to recover until they began singing (3-10 days). They were regularly weighed and monitored for discomfort. 3/9 birds that were implanted, did not recover and died within 1-3 days post surgery. Birds that recovered and resumed singing were plugged-in for few hours initially and then continuously for many days. In 1/6 birds that survived surgery, RA was not located during surgery and characteristic RA activity was not obtained post-surgery as well. This bird was not recorded further. Out of 5 remaining birds, the songs post-surgery were compared to that pre-surgery to check for possible changes to the song post surgery. Such changes would indicate potential damage to song control areas (RA or HVC as electrodes pass through both). Two birds changed their songs and never recovered. One of these birds was recorded but excluded from analysis as the post-surgery song did not have singing-related activity. One bird changed its song initially but recovered within few days and recordings were carried out after recovery. Overall, 3 birds were successfully recorded in RA simultaneously while singing. Both directed and undirected songs were recorded, but only undirected songs are analyzed for this study. It took longer for birds to start singing undirected songs (20-30 days post surgery) and thus the data was limited by number of songs recorded from the birds.

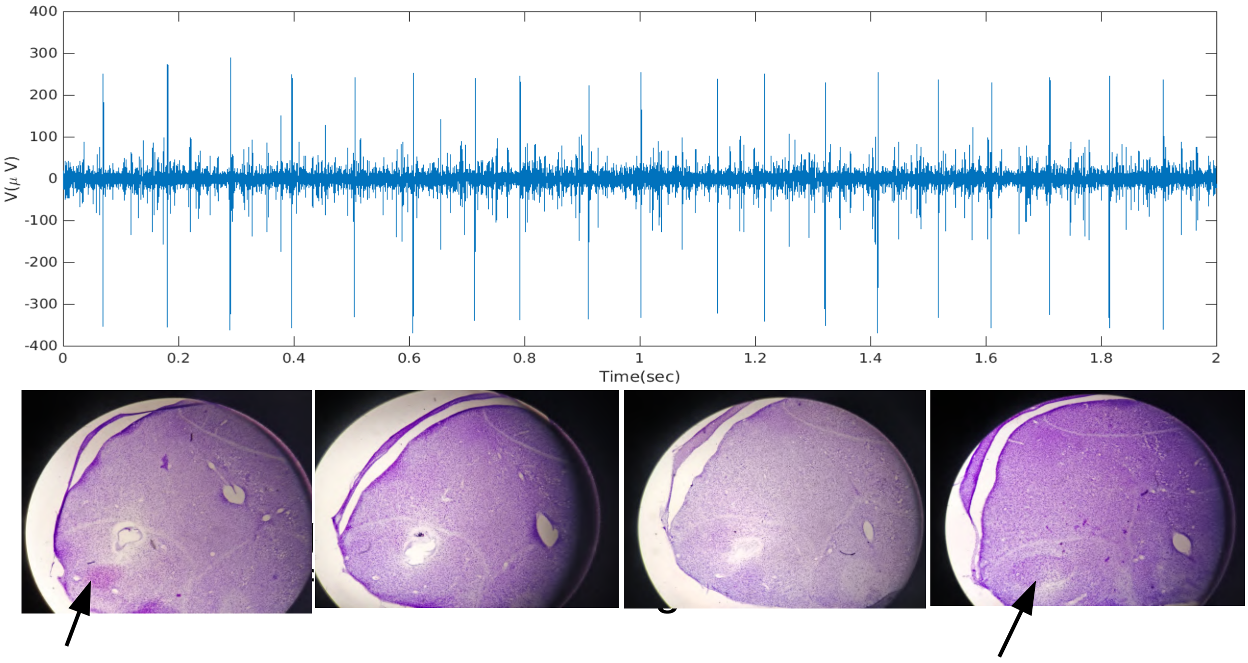

Fig. 4.1 Standardization experiments were done to locate RA
(A) Sample trace from one tetrode channel for RA activity during anesthtized recordings. Spontaneous RA activity showing regular firing. (B) Photographs of histology sections stained with cresyl violet for qualitative identification of RA lesions. The left and right panels show histology sections (40 μm depth) with intact RA (pointed by black arrows). The two sections in the center show successful RA lesions. The photographs were taken using phone camera that was aligned to the eyepiece of microscope, hence lacks scale bar.

In order to simultaneously record song and electrical activity another microphone (Adafruit MAX4466) that could be connected directly to the INTAN-acquisition board was used at sampling rate of 30 KHz. The implant was plugged-in to the 16/32 channel head-stage that connected to a commutator (Crist instrument) via flexible cables and in turn to the INTAN evaluation board. The commutator was fixed to the roof of the bird cage and prevented tangling of wires, allowing the bird to move freely inside the cage. In the initial days, birds were monitored continuously for discomfort and plugged-in for few hours. Eventually all the birds were continuously plugged-in for the entire period of recording.

During recording, the amplifier was set at sampling rate of 30 KHz and bandwidth of 500 Hz to 7500 Hz. Audio and electrical activity were recorded in the "continuous" mode. Similar procedures were performed to record song and neural activity in HVC, the details of which were described in previous studies (McCasland, 1987; Rajan and Doupe, 2013).

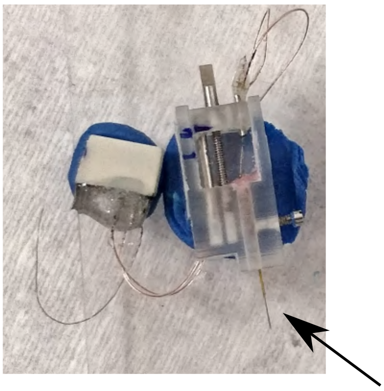

Fig. 4.2 Custom-designed implant
Picture of implant custom-built for chronic recordings. Arrowhead points towards the electrodes. The electrodes are placed inside a 3D printed microdrive

4.2.7. Data analysis

All data analyses were done using custom-written scripts in MATLAB.

4.2.7.1. Analyzing single unit activity in HVC

Single units recorded in HVC

The spike sorting for HVC interneurons and HVC-X neurons were carried out earlier as part of a previous study (Rajan, 2018). Out of 17 HVC-interneurons recorded from 4 birds, 13 neurons from 2 birds (7 and 6 neurons respectively in the two birds) had more than 10 bouts beginning with INs.

These were included for analysis in the current study. Out of 38 HVC-X neurons recorded from 7 birds, 12 neurons had more than 10 bouts beginning with INs. The number of spikes and the mean firing rate measured as number of spikes per second were analyzed during INs.

Activity related to IN syllable or interval or position

A 50 ms offset was considered while analyzing activity related to the behavioral output. This time offset accounts for the pre-motor latency used by earlier studies (Katz and Gurney, 1981; McCasland and Konishi, 1981; McCasland, 1987; Williams and Vicario, 1993). Activity related to IN syllable was measured by counting spike events between 50 ms before syllable onset to 50 ms before syllable offset. Similarly, the activity related to an interval was measured by counting spikes between 50 ms before onset of interval to 50 ms before offset of interval. For activity related to an IN position, the sum total of spikes related to IN syllable and the following interval were measured.

Onset and offset aligned activity

The activity related to syllable or interval were aligned in order to measure average activity across trials. The reference for alignment was either onset or offset of syllable or interval which is subtracted from the time of spike events, for alignment. As interval durations are variable across trials, the rasters and average activity for interval duration were obtained by aligning to onset and offset separately.

Pairwise correlations

To compare pattern of activity between pairs of trials, pairwise correlations were calculated similar to earlier studies (Kao et al., 2008; Rajan and Doupe, 2013). For each trial, the instantaneous firing rate (IFR) was obtained by binning spike events to 0.1 ms resolution and smoothing the spike train with a gaussian filter (width SD=5 ms; length of filter = 4SD). The correlation coefficient between two trials i and j was obtained by using the following formula:

$$Corr(i,j)=\frac{(IFR(i)-mean(IFR(i)))*(IFR(j)-mean(IFR(j)))}{norm(IFR(i)-mean(IFR(i)))*norm(IFR(j)-mean(IFR(j)))}$$

Comparing relative activity between first, middle and last IN

To compare activity patterns between first, middle and last positions of INs, the firing rate for the three positions were compared. As intervals may be of different lengths at the different positions of INs, the duration equal to the sum of syllable duration and shortest last position interval were considered for activity related to each position of INs. The proportional number of spikes between the first, middle and last positions were represented on the three axes of the tern-plot. The proportional activity was measured for 20 ms windows moved in steps of 5 ms in order to obtain trajectories of relative activity pattern. Based on the trajectories for a neuron on the tern plot, inferences were made about the difference in proportional activity of the neuron during three positions.

4.2.7.2. Analyzing multiunit activity in RA

Multiunits recorded in RA

Multiunit activity recorded in RA simultaneously with undirected songs were analyzed for 3 birds. 15 sites with >= 3 bouts of undirected songs were obtained (MEDIAN =12, RANGE [3 64]). 13/15 sites that had regular spiking in the non-singing condition and changed to bursts during singing were analyzed.

Analyzing multi-unit activity related to INs

The raw activity was rectified and smoothened by convolving with a 20 ms width gaussian filter (width SD=5ms; length of filter = 4SD) to obtain activity pattern for INs. Activity related to position of INs were analyzed using similar time windows as HVC analysis.

4.2.7.3. Statistics

Sample sizes:

All statistical analyses were done for parameters calculated for n=13 HVC interneurons and non-parametric tests were employed throughout.

Statistical tests:

Wilcoxon signed-rank test was employed for pairwise comparisons of neural activity across 13 neurons: interval vs syllable related activity (Fig. 3 B), first vs last IN interval (Fig. 4.3 C, D, Fig. 4.5 B), compare average correlated activity patterns when middle positions of INs were kept separate or combined (Fig. 4.8 B). The significance level was set at p-value ≤ 0.05 to reject the null hypothesis that the medians between the two groups are equal. Multiple group comparison of average spike counts at different IN positions (Fig. 4.7) was done using Kruskal-Wallis as there was variability in the number of neurons contributing to different positions. The significance level was set at $p \leq 0.05$ to reject the null hypothesis that all groups have equal mean ranks.

To compare activity across trials during IN intervals (Fig. 4.4) in individual neurons, the correlation was measured using Pearson's correlation coefficient with p-value ≤ 0.05 as the significance level.

4.3. RESULTS

The representation of IN progression were analyzed for extracellular recordings from HVC and RA. The data consisted of single-unit recordings of putative interneurons and putative HVC_X projection neurons, and multi-unit recordings of projection neurons in RA (13 sites). Single units of HVC interneurons and HVC_X neurons classified in the previous study and activity sorted for spikes were directly used in this study. Out of 17 identified putative interneurons, 13 had at least 10 bouts beginning with INs and only these were included for analysis. Similarly, 12/38 HVC_X units were used in analysis wherever comparison to results from interneuron activity was required. All of the analysis of neural activity during INs include recordings from awake singing birds under undirected condition.

4.3.1. HVC interneurons show interval related activity that is different at the start and the end of IN progression

Putative HVC interneurons show spiking activity throughout the duration of INs, including intervals between INs (Fig. 4.3 A – Example spectrogram and trace of one HVC interneuron showing INs before song and the corresponding syllable and interval related activity, shown by red and green color shading respectively). Across birds, the number of spikes during IN intervals were not significantly different from that during INs (Fig 4.3 B, p=0.19, Wilcoxon signed-rank test). Given the presence of activity related to IN intervals, I next asked whether activity was different between first interval (interval between first two INs) and last interval (interval between last IN and first song syllable) of IN progression. The mean (Fig. 4.3 A, $p=2.44 \times 10^{-4}$, Wilcoxon signed-rank test) and variability (Fig. 4.3 B $p=2.44\times10^{-4}$, Wilcoxon signed-rank test) of number of spikes were significantly more during the first interval (Fig. 4.3 A and B, $p<0.05$, Wilcpxon signed-rank test). The results suggest that HVC interneurons show significant activity during intervals and the interval related activity is different between start and end of IN progression.

4.3.2. Interneuron activity related to intervals does not indicate onset of next syllable

The difference in first and last interval related spiking suggests that the activity during the interval between INs may be important for IN progression. I next asked whether at a given interval position spiking activity explained variation in the interval across bouts. The firing rate related to first IN

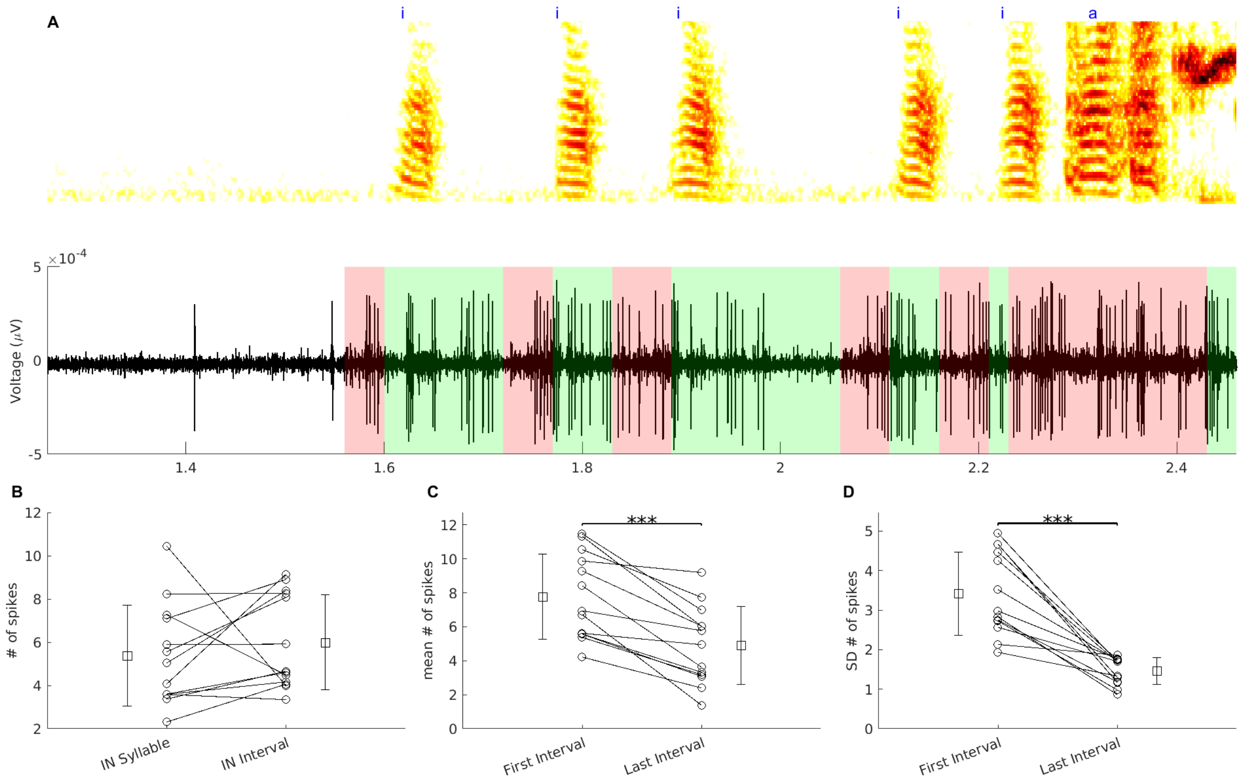

Fig. 4.3 Interneurons show interval related activity that is different between first and last position of INs (A) Example spectrogram (top) of one bout beginning with INs (labelled as ‘i’) before motif (first motif syllable labeled as ‘a’), along with activity trace for single-unit HVC interneuron (bottom). The red and the green shades mark the premotor period for syllables and intervals respectively. (B) Average number of spikes related to syllable and interval. (C) Mean number of spikes related to first and last interval. (D) Standard deviation of number of spikes related to first and last interval. Number of spikes were counted for the duration of syllable or interval during premotor period i.e from 50 ms before onset to 50 ms before offset. Circles are mean data from individual neurons. The square and whiskers represent median and interquartiles across neurons. (B-D) *** $p<0.001$, Wilcoxon signed-rank test.

interval (Fig. 4.4 A magenta ticks on raster) was negatively correlated for the majority of neurons (10/13) across the two birds (Fig. 4.4 B, C; RANGE [-0.71 0.25], median=-0.55, Pearson’s correlation coefficient, p-value<0.05). Thus, longer intervals were associated with slower activity. Given the variation across bouts in both interval and interval related activity, I next asked whether the activity during the intervals drives the progression of INs by indicating the timing for the next syllable or the end of the interval. If this is true, I hypothesized that the activity for the time during the interval that is common across bouts should be related to when the interval ends. Thus, I correlated the activity of neurons for the shared duration within this interval to the length of the interval across bouts. The shared duration is the premotor period related to the shortest interval.

During this common period across trials, number of spikes were counted and correlated to the interval length (Fig. 4.4 A – raster for example neuron with shaded area showing the common duration during which spikes were counted for comparison with corresponding interval. Fig. 4.4 B

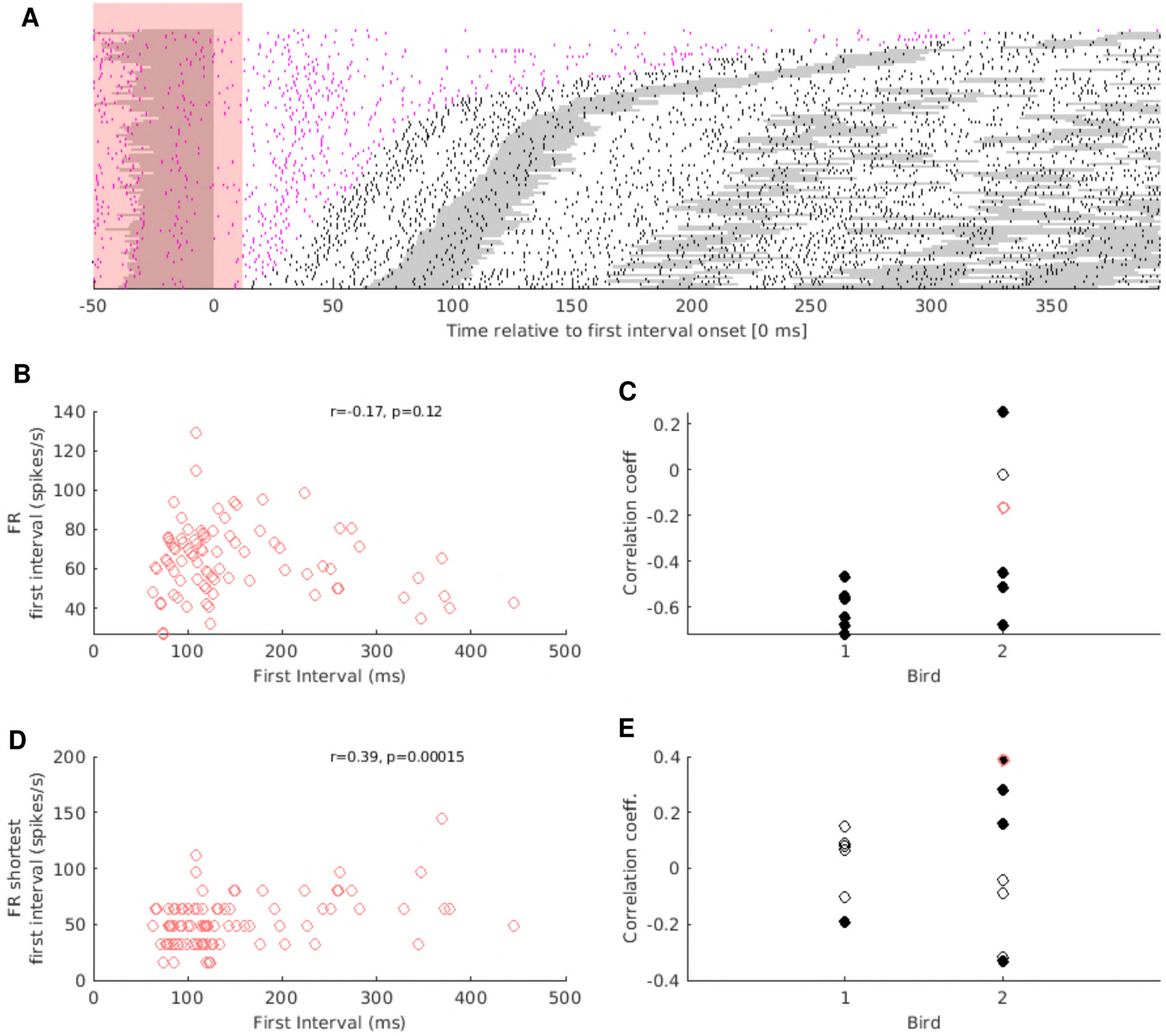

Fig. 4.4 Interval related activity is not related to duration of the interval
(A) Raster for sample neuron overlayed on IN sequences (grey bars represent duration of IN) across bouts. Bouts are sorted in descending order of duration of first interval and aligned to the onset of first interval between INs. The pink ticks represent activity premotor to the first interval and the red shaded area marks the period premotor to shortest first interval. (B) and (D) First interval vs firing rate of neuron shown in (A) for the entire interval duration (B) and for only the premotor period of shortest first interval (D). Each circle represents data from one bout. r and p show correlation coefficient and significance level. (C) and (E) Correlation coefficient for all neurons. Red circle represents the same neuron as in (A). (C), (E) Filled circles represent significant correlations, p-value≤0.05

scatter plot for example neuron). Majority of interneurons (8/13) did not show significant correlations, while weak correlations were present in 5/13 neurons (Fig. 4.4 C, RANGE [-0.33 0.36], median=0.06, Pearson's correlation coefficient, p-value≤0.05]). Similar to activity across trials at the first interval, I compared the average activity for a common duration between the first and last interval across neurons (Fig. 4.5 A-C). The average firing rates were not significantly different between the two groups. No consistent trends were observed between average firing rates of first and last position across neurons. These results suggest that interval related activity may not drive the interval length or alternatively the time for onset of the next syllable at all positions of INs. Thus, progression of INs may not be driven by interval related activity in HVC.

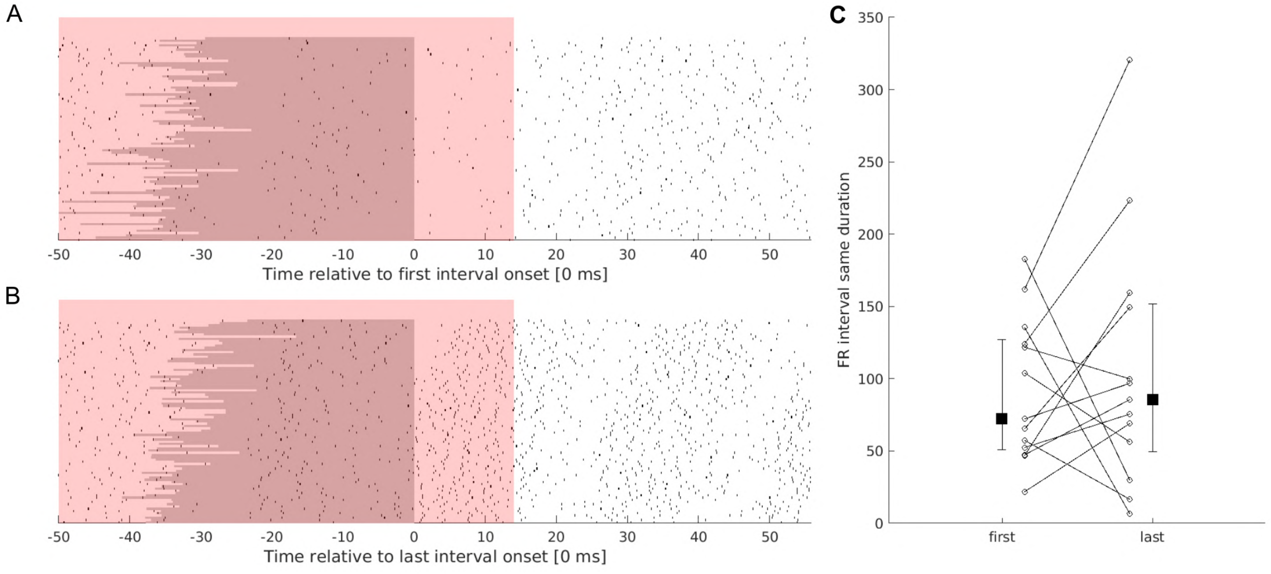

Fig. 4.5 Interval related activity for the same period of first and last interval is not consistent across neurons (A) and (B) Raster for first and last IN interval respectively for a sample neuron, similar to Fig 4.4 A. Firing rates were measured for a premotor period equal to the shortest last interval (red shaded area) in both A and B. (C) Firing rate (spikes/s) for first and last IN interval related activity across neurons (circles connected with lines). The square and whiskers represent median and interquartiles for the group data. Note that the firing rate differences between first and last positions do not show a consistent trend across neurons. p>0.05, Wilcoxon signed-rank test.

4.3.3. Interneuron activity related to intervals encodes position in the sequence

I next hypothesized that the progression of INs may be related to position of INs as they progress to the song. For this, I asked whether pattern of activity in an interval was consistent for a position. Pairwise correlations (see Methods section 4.2.7.1) of activity patterns across bouts for the first interval show that the patterns are correlated more than the chance level (Fig. 4.6 A and B raster for example neuron aligned to onset or offset of interval and the corresponding trace for correlations in Fig. 4.6 C and Fig. 4.6 D). An interval lies between two IN positions. I asked whether the position related activity was more consistent at one end of position than the other. Across neurons, the correlations were higher when the activity was aligned to onset rather than the offset (Fig. 4.6 E and F). Following the end of interval related activity, the correlations again increase during the period premotor to the onset of the next IN showing increased consistency of activity pattern for the next position (Fig. 4.6 F, red vertical line). Thus, the activity of the interval is associated with the position of IN preceding it rather than the position of IN ahead of it. Together, this analysis shows that the activity related to an IN syllable and its following interval represent a position in the IN sequence with consistent activity.

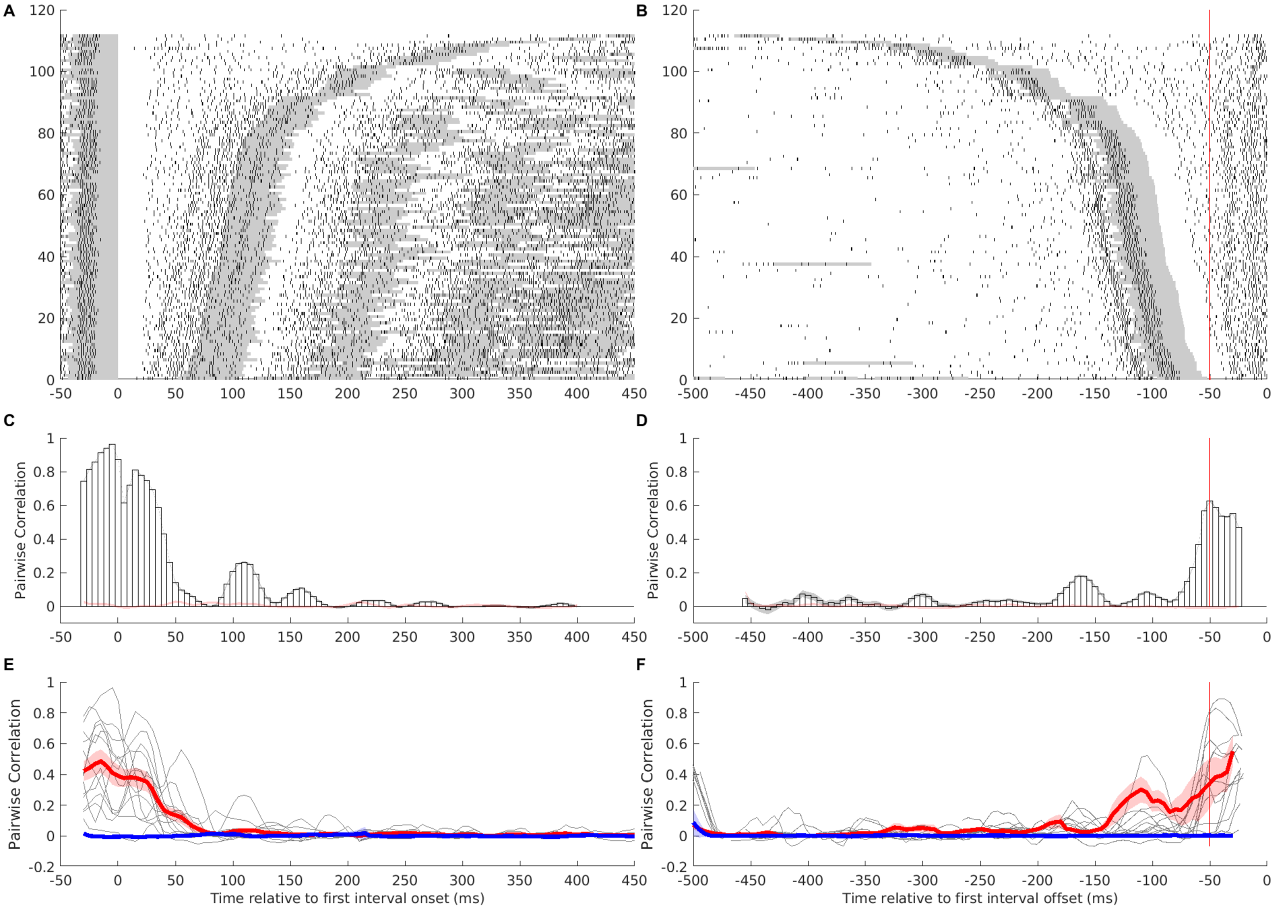

Fig. 4.6 Interval related activity is specific to the position of IN it follows
(A) Raster plot overlayed on color plot of IN sequence aligned to onset of first interval (B) Raster plot overlayed on color plot of IN sequence aligned to offset of first interval. Bouts in (A) and (B) are sorted based on duration of first interval between INs. (C) and (D) pairwise correlations between bouts for the sample neuron in (A) and (B) respectively. The red traces (near 0) show correlation related to chance level. Red vertical line in (D) marks the beginning of premotor period for next IN. (E) and (F) Pairwise correlation curves for all neurons (grey lines) and the mean ± SEM across neurons (red line and shaded area). The blue line and the blue shaded area mark the mean and 95 % confidence intervals for chance level (obtained by randomizing spike times)

4.3.4. Position related activity of INs represented by activity at first, middle and last positions

Based on the result that interneuron activity is specific to position of INs, I next hypothesized that IN progression may be represented by changes in activity related to IN position. The number of INs vary across bouts and hence the position of INs could be aligned to the first or the last position. With either of these alignments, the average number of spikes was not significantly different across IN positions(Fig. 4.7 A and B). However, there was a trend for gradual reduction in number of spikes for later INs, which was more prominent when positions were aligned to the last. I next asked whether pattern of activity was different for the different IN positions. Pattern of activity for each trial was compared to average pattern of activity at different IN positions (Fig. 4.8 A sample neuron, positions were aligned to last IN position). Given the variability in repeat-number of INs, the possible number of positions may differ for the activity of neuron across bouts and the first and last aligned positions may have a different subset of trials. To avoid this caveat, the positions may be reduced to three positions common to all trials (Fig. 4.10 A – illustration): 1) the first IN which is not preceded by any syllable and followed by an IN. 2) the middle IN which is preceded and followed by IN syllable and 3) the last IN which is preceded by an IN syllable and followed by the song syllable. I observed that the average correlation for a position remained high, whether distinct IN positions were considered or only "first-middle-last" (Fig. 4.8 B, C). Further analysis was done using these three positions.

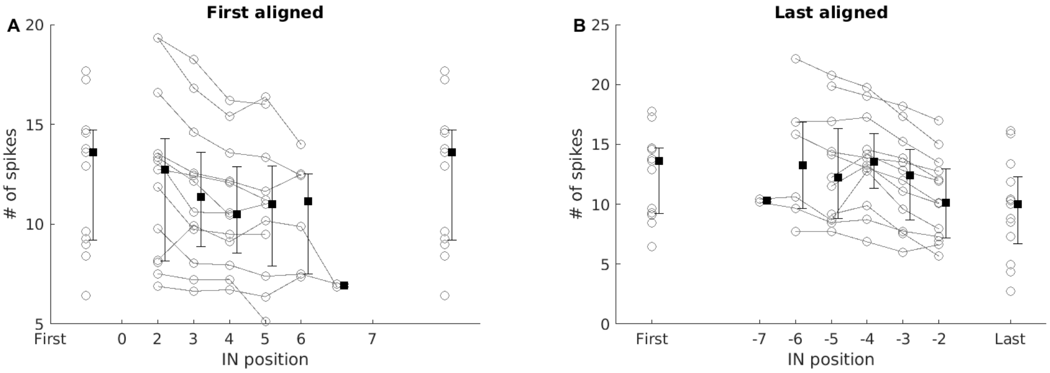

Fig. 4.7 Spike counts are consistent across IN positions Average number of spikes for each IN position from first to last. The positions of INs between first to last were either aligned to first (A) or last (B) IN. Circle connected by lines represent data from individual neurons. Square and whiskers represent median and interquartiles for the group. The groups were compared by Kruskal-Wallis test, $p>0.05$.

4.3.5. Activity of HVC interneurons at different positions of INs start the same but end differently

The pairwise correlations show position-based differences in pattern of activity across neurons (Fig. 4.9). Further, the position-based differences varied across neurons. For example, in a majority of neurons the activity pattern were similar between first and middle positions (Neuron # 1, 2, 3, 4, 5, 6, 7, 10 and 12). For another set of neurons, the activity pattern was correlated between middle and last position (Neuron # 3, 11, 12, 13). These differences suggest that the pattern of activity changes for IN position across neurons. Next, for each neuron, the average pattern of activity at first, middle and last positions were compared at different time points (Fig. 4.10 B, C – example neuron raster for the three positions and the pattern of activity). The start point was common for all these positions – start of premotor period for the IN syllable at that position. However, owing to the reduction in duration of intervals with IN progression, the end time of activity pattern was not consistent across all positions.

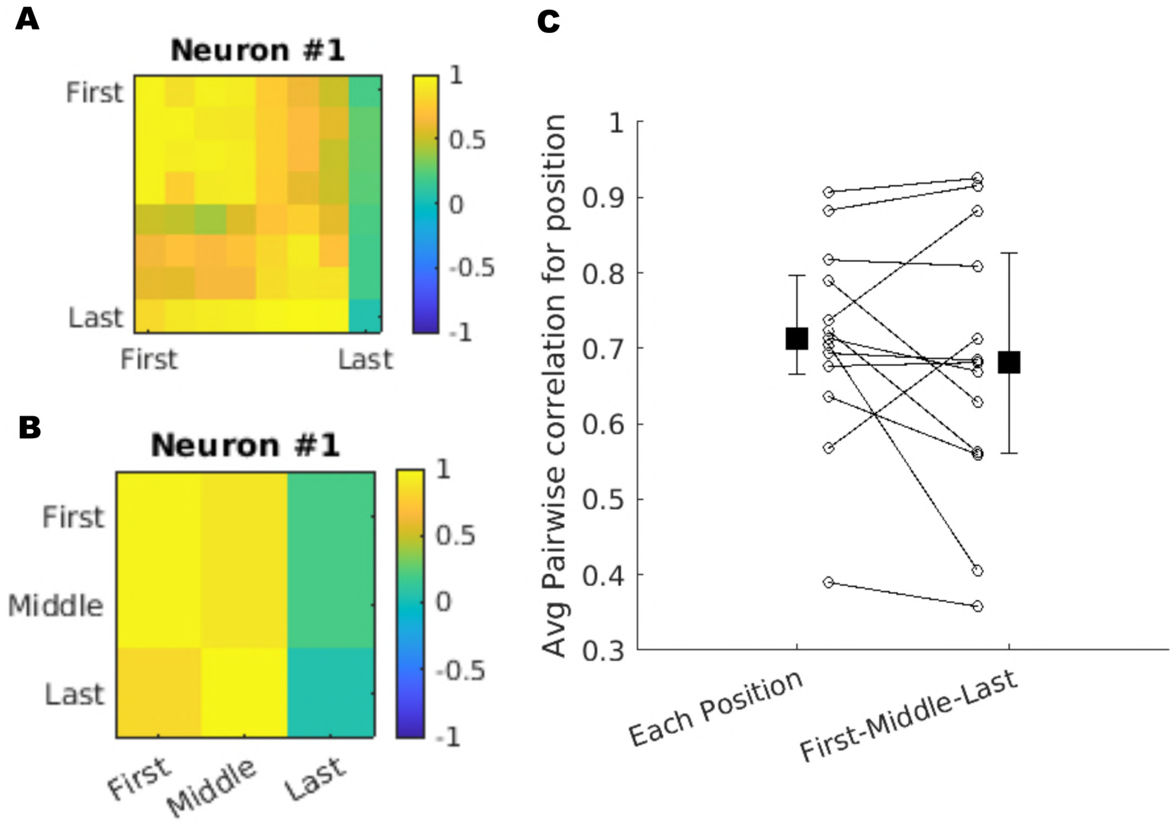

Fig. 4.8 The pattern of activity for an IN position is similarly correlated when all middle positions are combined (A) -(B) Correlation matrix for pairwise correlations of activity pattern at each position (vertical) with mean activity of a position (horizontal) for a sample neuron. The different IN positions were either aligned to last position (A) or combined to one middle position (B). The color represent the correlation value according to color scale. (C) Average correlations for a given position (diagonal of correlation matrix) compared for the two methods across neurons. The circles joined by lines represent indivudal neurons. The squares and whiskers represent medians and interquartiles for the groups.>0.05, Wilcoxon signed rank test.

The activity pattern was thus compared until the end of premotor period for the shortest last interval across renditions (Fig. 4.10 D, example relative pattern from start – circle to end - asterix). Across neurons, the start of neural activity pattern changed from start to end of the position (Fig. 4.11 A-B). The pattern at the position end differed considerably for majority of the neurons (Fig. 4.11 B). The neurons could be grouped into three categories based on the relationship between positions at the ending time point (Fig. 4.11 B): 1) Neurons that had similar activity for first and middle, but no activity at the last position. 2) Neurons that had similar activity for middle and last, but no activity at the first position. 3) Neurons that were active only for the last position. These position-based differences in the activity pattern of interneurons that evolve during the position, occur along with position-based bursts known to be present in HVC (HVC$_X$ activity reported in a previous study, for the same birds are overlayed). Together with activity in HVC_X neurons (4.11 A, B triangle points), these results suggest IN progression may be represented by population of HVC neurons that may change the combination of active neurons with each position of INs.

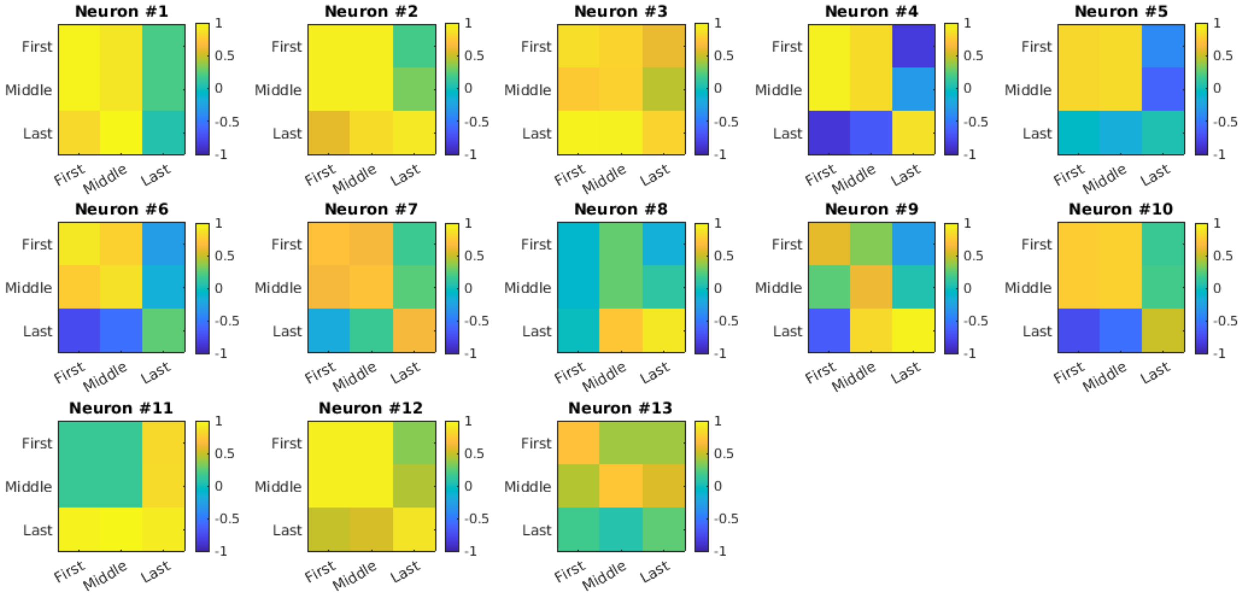

Fig. 4.9 The average pattern of activity varied between first, middle and last IN positions Correlation matrices for all interneurons comparing activity patterns at first, middle and last IN positions (average pairwise correlation for activity pattern at each position (vertical) with mean activity of a position (horizontal)). The color represent the correlation value according to color scale.

4.3.6. Multiunit activity of RA neurons is consistent across IN positions, unlike HVC

Given the results of HVC neurons having different position-related activity, I next asked whether these differences propagate from HVC to RA in the song motor pathway. I hypothesized that RA neurons may also represent progression of INs in a similar manner to HVC. Similar analysis as HVC neurons were repeated for 15 multiunit sites recorded in RA (Fig. 4.12 A, B– example bird and trace recorded for RA). However the pattern of activity at three positions of first, middle and last were consistent and did not show trends like in HVC (Fig. 4.13 A, B). The consistent activity of RA neurons across IN positions suggest IN-related activity in RA does not represent progression of INs similar to HVC.

4.4. DISCUSSION

The results of this study show that the activity of HVC interneurons changes with the position of INs, as INs repeat and progress towards the song. Different combinations of neurons are active at

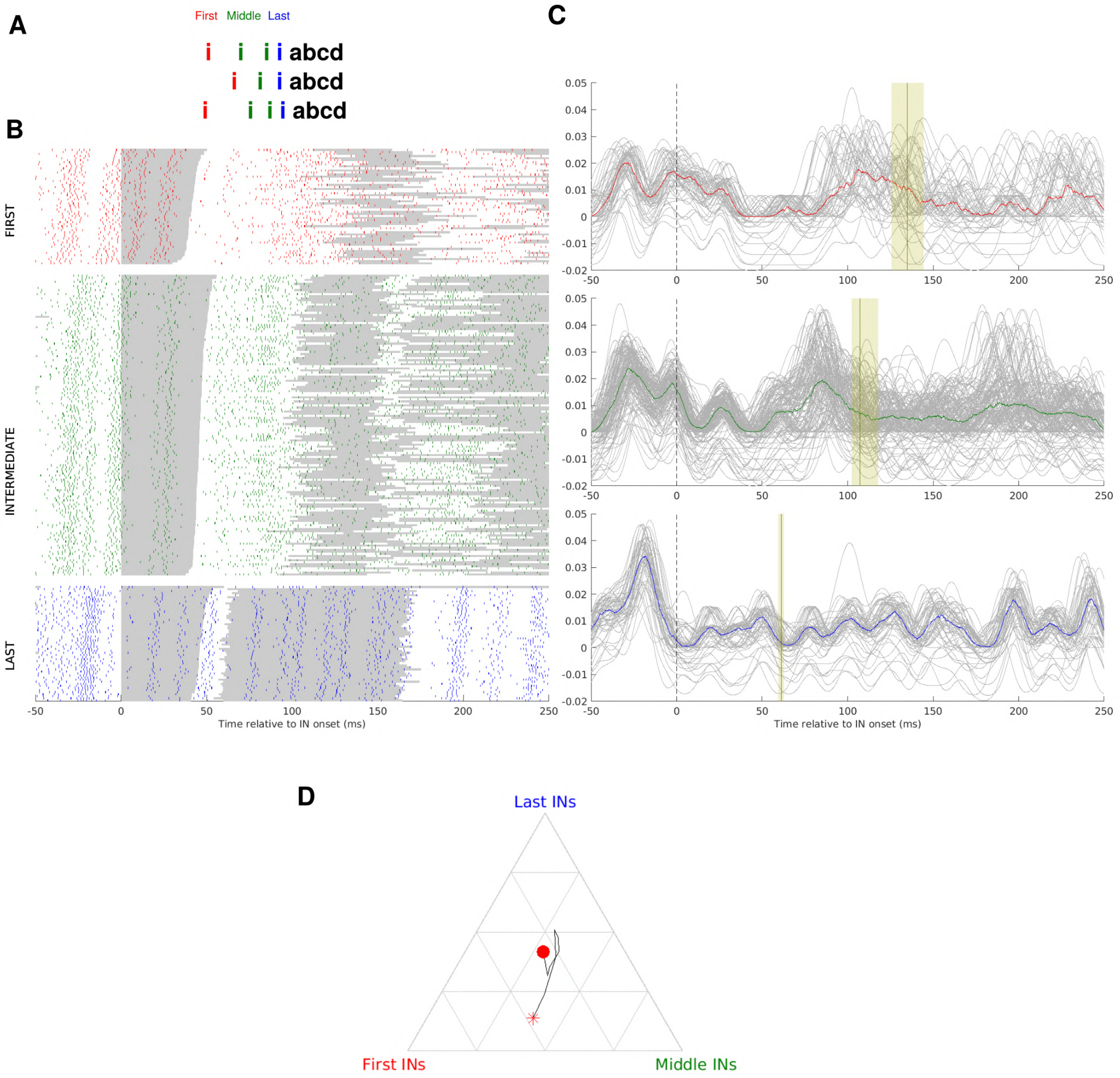

Fig. 4.10 Relative activity of a neuron at first, middle and last position
(A) Schematic showing positions of INs across trials as first, middle and last positions (B) Raster of neuron separated for first, middle and last position of IN. The grey bar represent duration of IN syllables. (C) Normalized activity trace for each trial (grey) and the mean (colored) for the three positions. The vertical line and yellow patch represent median and interquartile range for duration from onset of IN to offset of following interval. (D) Relative activity between the three positions from start (circle) to end (*). Activity was binned at 20 ms windows moved in steps of 5 ms. The triangle plot represents proportional activity between three positions, each vertex and opposite base representing axis for one position. Start point (circle) is start of position at 50 ms before IN onset and end point (*) is end of position at 50 ms before shortest last interval. The triangle plot represents relative proportions between 3 axes- value at base is 0 and at the vertex is 1 for first, middle and last. The centroid represents equal proportions.

the first, middle and the last position of INs. This adds to the already known position related IN activity of HVC-X neurons, a subset of which are differently active for the last IN. The presence of

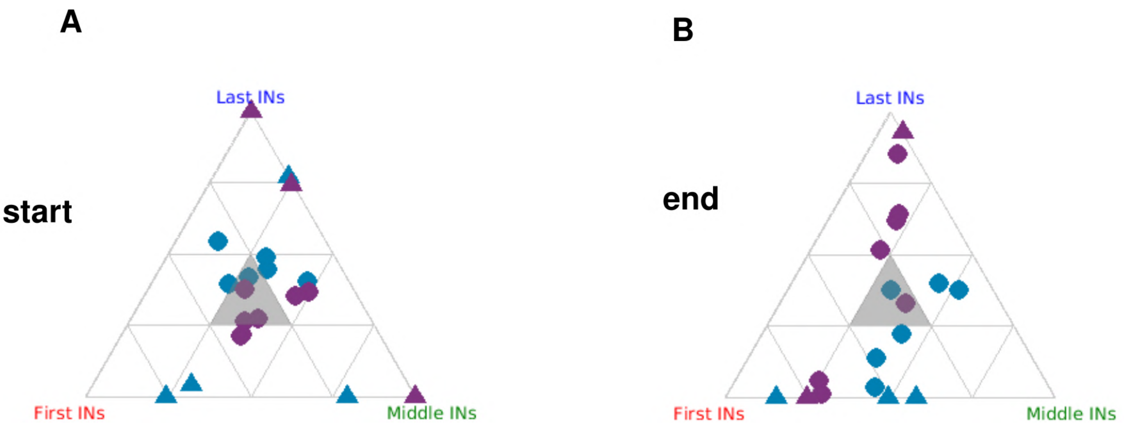

Fig. 4.11 Position related activity of HVC interneurons end differently for first, middle and last INs Relative activity between first, middle and last positions for all neurons at the start (A) and end (B) of position. Each circle is data from one interneuron. The triangles represent data for HVCx neurons. The two colors represent neurons from two different birds. The triangle plot represents relative proportions between 3 axes- value at base is 0 and at the vertex is 1 for first, middle and last. The centroid represents equal proportions.

the three types of HVC interneurons described above may together represent the progression of INs to song (Fig. 4.14). A possible explanation for the differential pattern of activity at first, middle and last positions is that the activity may be representing the identity of the previous or next syllable. Corresponding to the three neuron types: 1) Neurons active for first and middle, but not the last may represent that the next syllable is an IN. 2) Neurons active for middle and last, but not the first may represent that the previous syllable was an IN and 3) Neurons active only for the last may represent that the next syllable is song syllable, marking the end of progression. The analysis was done across two birds and more recordings are necessary for stronger conclusions. However, given these data, the results suggest a possible representation of IN progression by combination of neurons in HVC that keep track of past and future syllables. Contrary to neurons in HVC, RA multiunit activity was consistent across IN positions suggesting that at the level of groups of neurons in RA, there is no representation of IN progression.

4.4.1. Representation of IN progression by HVC_{RA} neurons

This study did not directly analyze activity of HVC_{RA} neurons that are directly involved in the production of vocalizations in the song motor pathway. HVC_{RA} neurons do not have spontaneous activity, are smaller in size, and only half of them are involved in singing (Hahnloser et al., 2002). Further, only a fraction of those active for song were shown to be active during INs (Kozhevnikov and Fee, 2007). These reasons make recording of HVC_{RA} neurons difficult for IN related activity. However, interneurons are readouts of multiple HVC_{RA} neurons and HVC_X neurons show bursts similar to HVC$_{RA}$ neurons (Kozhevnikov and Fee, 2007). Thus results for representation of IN progressions from interneurons and HVC-X neurons may be extrapolated for HVC-$_{RA}$ neurons as well. As the pattern of interneuron activity repeats for some successive positions, this would mean that HVC$_{RA}$ neurons may also repeat and be active for more than one IN in the sequence. This is unlike the reported activity of HVC$_{RA}$ neurons for song where they burst only once during the song sequence. Thus, the results suggest different activity of HVC$_{RA}$ neurons than song, in order to represent IN progression.

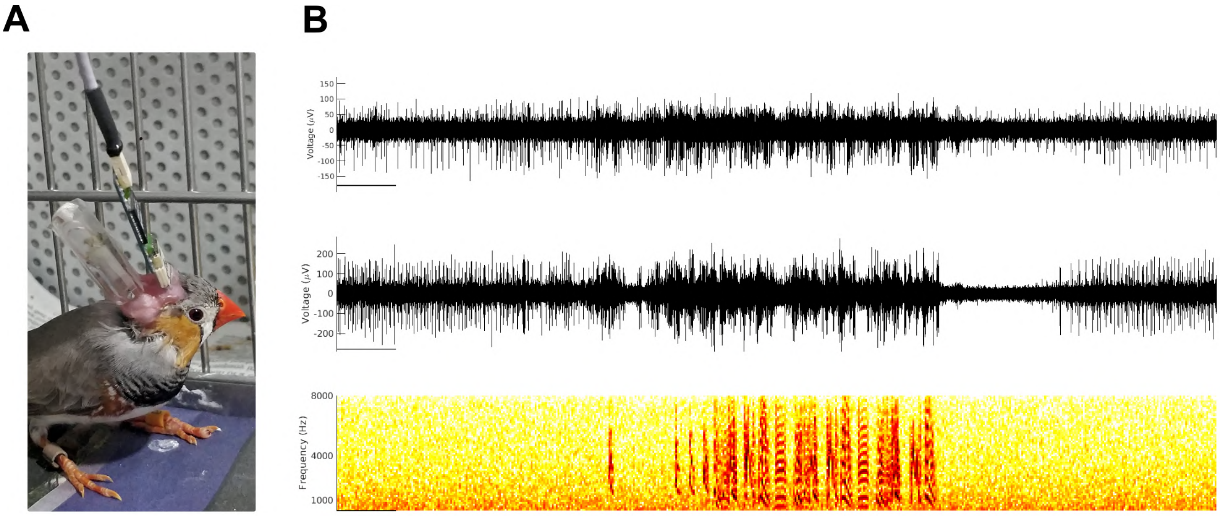

Fig. 4.12 Extracellular recording in RA in an awake singing bird
(A) Picture of a bird implanted with electrodes and plugged in to the amplifier. Song and electrical activity were recorded simultaneously in undirected condition
(B) Example recording trace from two simultaneously recorded electrode channels aligned to the spectrogram of undirected song (B). Recorded activity represents multiunit sites. RA activity switches from regular tonic spiking during silence to dense bursts during song. A characteristic brief silence follows at the end of the bout. Scalebar represents 500 m

Fig. 4.13 Activity across multiunit sites in RA start the same and end the same for all positions of INs
Relative activity between first, middle and last positions for all neurons at the start (A) and end (B) of position related activity. Each circle is data from one multiunit site. The three colors represent sites from three different birds. The triangle plot represents relative proportions between 3 axes- value at base is 0 and at the vertex is 1 for first, middle and last. The centroid represents equal proportions.

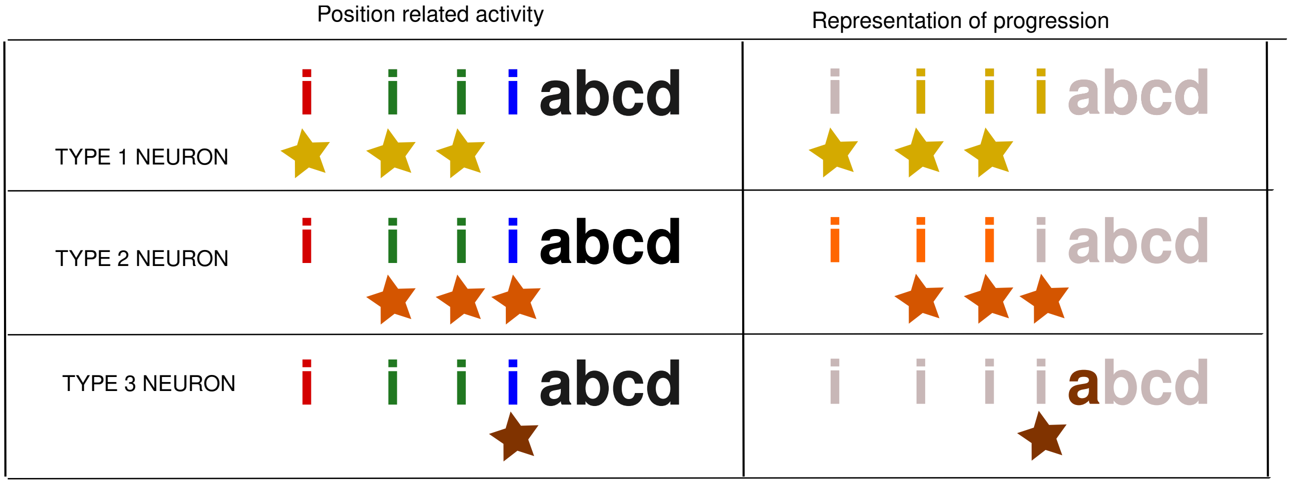

Fig 4.14 Position related activity indicates representation of progression by representing past and future syllables
Left column shows position related activity for three types of interneurons identified in HVC. Right column is the interpretation of the position related activity in the three types of neurons as an indicator of past or future syllable. The activity of type 1 neuron is consistent with identity of next syllable as IN, the activity of type 2 neuron is consistent with identity of previous syllable as IN and the activity of type 3 neuron is consistent with identity of next syllable as the first song syllable. Together combination of the three types of neurons could represent progression of INs as they repeat and switch to the song.

4.4.2. HVC representing past and future syllables

The HVC neurons in zebra finches are known to represent a moment in song and do not represent past or future syllables (Hahnloser et al., 2002; Kozhevnikov and Fee, 2007). The analysis of IN related activity draws parallels to HVC activity during calls and in canaries (Benichov et al., 2015; Cohen et al., 2020). Recent studies on call-timing during social interactions suggested that HVC plays a role in predicting and witholding responses to calls. The precise timing was abolished with inactivation in HVC. This evidence supports that HVC can anticipate future vocalizations. The HVC of canaries are known to encode past syllables related to transition history that occurred a few phrases earlier – canary songs are 20-40 seconds long consisting of phrases, each phrase is a sequence containing syllable repetitions. This evidence suggests HVC in the canaries is capable of keeping track of past syllables. Further, representation of sequences containing variable transitions have also been shown for bengalese finches, a species that produces relatively more variable songs (Fujimoto et al., 2011; Warren et al., 2012; Benichov et al., 2015). As IN repetitions are also a version of variable sequences, it is possible that HVC behaves similar to other species with variable songs, during INs. Such variable sequencing have been explained by higher order Markov models (Kershenbaum et al., 2014). The representation of INs as analyzed in this chapter are indicative of hidden states – i.e syllable transitions depend on history of previous syllables and that instances of same syllables may be represented by different combination of neurons. The similarities of INs with variable sequences in other birds, suggest a conserved function of HVC across songbird species.

4.4.3. Interval related activity in HVC does not arise from within HVC

I showed that HVC interneurons have interval related activity, but only a fraction of these neurons show correlations with the duration of intervals. Thus, HVC activity does not represent how long the next IN will take, suggesting that the timing of INs may be arising from outside HVC. The neurons with significant correlations are weak suggesting the relationship may be present at the level of a group of neurons. This analysis was done for the first intervals which are relatively longer in duration. Given the position-based representation of IN progression, it is possible that a change in state results in the control of timing for later INs during shorter intervals within HVC. Chapter 2 showed that presence of calls <200 ms before the first IN, facilitated IN progression. It suggests that controlling the interval length may influence activity in HVC. Studies on heating and cooling of HVC suggest 70 % stretching of syllables, but intervals were not stretched uniformly (Long and Fee, 2008). Together, with our current data, it can be speculated that the control for interval timing arises from outside HVC and is received by HVC. Two major inputs to HVC are Nif (nucleus Interface) and Uva (uvaeformis) (Williams and Vicario, 1993; Vyssotski et al., 2016). These show IN related activity and thus may be recorded to understand interval related activity during INs. Suggested future experiments to test their role include stimulating these areas during first interval and later intervals of INs.

4.4.4. Contribution of single neurons in RA to IN progression

Multiunit activity was not different across IN positions. As RA neurons receive input from HVC_{RA} neurons (Leonardo and Fee, 2005), this result is surprising. Previous recordings in RA at the level of single neurons have reported variable burst timing during INs (Yu and Margoliash, 1996). Further, activity in RA of bengalese finch show relationship with acoustic variability of syllables they encode (Sober et al., 2008). A possible explanation for this is that one RA neuron receives inputs from multiple HVC neurons (Hahnloser et al., 2002). It is possible that neurons receive inputs from all combinations of neurons in HVC, such that they cancel out and do not reflect any difference in RA activity across positions. Alternatively, it is possible that multiunit recordings of nearby neurons cancel out the different inputs that they receive from HVC. Earlier studies have recorded single units in RA and reported variable activity during INs. Thus, it is possible that these cancel out for multiunit activity. Another possibility is that the dorsal and ventral RA behave differently during INs. Dorsal RA has more projections to respiratory pathway compared to ventral

RA neurons that project to syrinx (Roberts et al., 2008). All my recordings in RA were at a consistent location across birds and thus cannot indicate such differences in RA activity. Further, Chapter 2 results show that removal of syringeal inputs did not affect IN progression. Future experiments may involve targeted recordings in dorsal and ventral RA to test if activity related to IN progression is represented in RA.

4.4.5. Propagation of IN progression through song motor pathway

RA neurons directly connect to downstream motor neurons that produce vocalizations (McCasland, 1987; Williams and Vicario, 1993). Given the current results that multiunit or groups of RA neurons do not show differences with IN position, how may IN progression be reflected as changes in temporal and acoustic properties in behavior? An explanation for this could be that changes in acoustic properties of INs are small and current multiunit recordings do not have the resolution to make the distinction based on activity. Alternatively, it is possible that downstream motor areas sending projection to syrinx and respiration receive independent inputs other than RA and thus show IN progression in behavior. This hypothesis may be tested in future by simultaneous recording of RA neurons along with EMG of syringeal muscles or respiratory pressure.

4.4.6. Representation of IN progression and sequence representation in HVC

The sequence of song is suggested to be represented in HVC by chains of HVC_{RA} neurons (Long and Fee, 2008; Long et al., 2010). These chains do not repeat for the duration of the song, as every neuron fires only once in the song. Given our interpretation of HVC_{RA} neurons firing more than once owing to the same pattern of interneuron activity repeating for successive INs, same chains may be getting active at each IN. At the start of a position, more interneurons showed similar activity, but this changed as the position moved to the interval. These may reflect HVC_{RA} chains starting with similar neurons but changing by the time the chain propagates. Finally, the chains propagate to the song. Although this hypothesis does not fit the traditional synch-fire chain models (Long et al., 2010), the newer polychronous chains (Egger et al., 2019) may be tested to understand the transition of IN progression to the song sequence.

Together, the results of this study have provided new insights for representation of IN progression and sequences in HVC. The study suggests that the temporal representation of IN progression may be arising from outside HVC. Further, it shows that HVC may represent progression of INs by a combination of neurons representing past and future syllables.

CHAPTER 5

Discussion

Using the zebra finch song, an ethologically relevant behavior, this thesis provides insight on how complex natural movements are initiated. The zebra finch begins its song with repetitions of introductory notes (INs) that are followed by one or more renditions of a stereotyped song sequence (Price, 1979; Sossinka and Böhner, 1980). The INs systematically change in their temporal and acoustic properties as they progress towards the song. Due to the progression of INs towards the song, a plausible role suggested for INs is preparation for song (Rajan and Doupe, 2013). To further understand the nature of IN properties and the contribution of INs towards a preparatory role, I addressed a specific question in this thesis of how INs progress to the song. A combination of behavioral analysis, surgical manipulation techniques and neurophysiology were used to test a number of related behavioral and neural hypotheses about the progression of INs to song. These have led to interesting findings that contribute towards understanding the mechanisms of song initiation, preparatory function of INs and the initiation of complex natural movements.

5.1. Summary of results

The results show that the progression of INs is independent of peripheral sensory feedback and this rules out a possible role of INs as calibration of the motor periphery before song (Rao et al., 2019). Together with the behavioral evidence that proximity to calls facilitate initiation of IN progression, these results suggest that the progression of INs represent changes in internal neural processes. As both INs and songs are vocalizations produced by the brain, I then compared properties of INs and songs to understand the similarities and differences between the two. INs and songs were correlated across trials and for the changes in their properties due to age. I show that the number of INs and the interval between INs are not related to the upcoming song for a given bird. However, the acoustic properties predominantly related to amplitude were correlated between syllables of INs and songs. The variability of number of INs were also similar to all other types of syllables that the birds repeat. Together, these data show shared relationships for syllable production between INs and songs, further suggesting shared control mechanisms. On the contrary, the temporal progression was unique to INs, suggesting distinct mechanisms control the timing of INs. To understand the neural correlates of IN progression, I analyzed neural activity during INs in two areas of the song motor pathway. I show that the representation of IN progression in the brain was present in the population of neurons of the premotor nucleus HVC. The representation at the level of single neurons was not related to intervals, but to the position of INs in the repeat sequence as first, middle or last INs. Further, the data suggests that a combination of neurons keeping track of the identity of past or future syllables form representations of IN progression in HVC. These representations were not present in multi-unit recordings of RA – the premotor nucleus downstream of HVC. This suggests that the representation of IN progression observed in HVC does not propagate downstream of motor pathway. In conclusion, the progression of INs is a process independent of sensory feedback, consisting of unique temporal properties, and is associated with a gradual reconfiguration of neurons in HVC as INs repeat and switch to the song.

5.2. Mechanism for song initiation

The work from this thesis report several findings about how zebra finches initiate their songs. Song initiation was unaffected after removal of two forms of sensory feedback – auditory feedback from the ear and proprioceptive feedback from the syringeal muscles. This ruled out that an active calibration of motor periphery, utilizing real-time sensory feedback from the vocal organ, precedes song initiation. Instead, it suggested that the time required for song initiation is potentially limited by internal changes in the brain. This was further supported by the fact that history of vocalizations influenced song initiation. The presence of calls in proximity to INs reduced the number of INs and resulted in a faster initiation of song. As neural activity related to calls have been reported in areas of the song motor pathway, this suggested that transient activity changes due to call production facilitated initiation for song. An explanation for this is that a possible change in the state of song related areas due to immediately preceding vocalizations allow change in the initial state of INs towards faster song initiation. Consistent with this idea, I also observed that shorter intervals between first two INs were related to a shorter time for song initiation. Thus, the time for song initiation depends on an initial state. In support for this, it was also shown that pre-bout preparatory activity in HVC preceding INs that successfully transition to song were different from that preceding solitary INs (Rajan, 2018). Together, all of these studies indicate the importance of initial state for determining the time for song initiation. Causal relationships between initial state and time for song initiation may be tested by disrupting the first interval by playback of noise interruptions or by stimulating in HVC to verify if song initiation is delayed and whether state of HVC is important for determining the start of song.

Following the determination of initial state, analysis of HVC activity during INs showed that song initiation involved repetitions of INs represented by specific combination of neurons encoding first, middle and last INs before switching to the song. Future experiments will involve understanding how HVC decides the sequence during INs, i.e whether next syllable after an IN should be an IN or a song syllable. These sequencing strategies may be tested in network models of HVC.

5.3. Preparatory role of INs

The progression of INs to song from an initial variable state to a final consistent state suggested a motor preparatory role for INs (Rajan and Doupe, 2013). Motor preparation has been described as structuring of motor commands before the execution of the actual movement. Consistent with this, it predicts relationship between preparatory activity and the properties of the upcoming movement. In the case of the zebra finch song, comparison of properties between IN and song showed no correlation between variable properties of INs – the number and the time interval between INs to any aspects related to song stereotypy or song. This showed that the number and temporal progression of INs is not a reflection of song related motor commands getting ready for song. Although the acoustic properties of INs were related between IN and song, the relationship was as similar as that between pairs of song syllables. Thus, the acoustic properties of INs may either play a preparatory role or may hold similarity to song as they are also syllables part of the same sequence. Separately correlating properties of first and last position to song of head-fixed mic birds did show more birds correlated at the last position for amplitude related properties. However, the number of birds were few to infer conclusive results. Recently it was shown that the number and structure of INs are learned from the tutor (Kalra et al., 2020). This evidence argues against the idea of acoustic properties of INs contributing towards motor preparation. Together, the data show that the properties of INs may not be related to motor preparation of song properties.

Another possible "preparatory" role of INs could be to synchronize the respiratory and vocal motor outputs before song initiation. Song is produced by coordination of muscles of syrinx with respiration (Wild et al., 1998; Franz and Goller, 2003; Cooper and Goller, 2006). Syllables are produced in synch with expirations and intervals between syllables occur during inspirations. Song is associated with high amplitude and high frequency breathing cycles compared to non-singing periods where breathing cycles are low amplitude and low frequency. A previous study suggests changes in breathing cycles start before the first IN (Daliparthi et al., 2018) but whether they transition gradually during INs has not been studied. A possible function of INs could be to establish this synchronization. This hypothesis suggests that respiratory pressure for initial INs may not be in sync initially but gets corrected towards the last IN before the song. Hypothesis related to this function of INs may be tested by carrying out respiratory pressure recordings along with song. A caveat with this hypothesis is the finding in Chapter 4 that showed that multiunit activity in RA were consistent across IN positions. RA sends projections to both respiratory motor nuclei and syringeal motor nuclei simultaneuously. However, it may be argued that although RA sends consistent inputs to respiratory nuclei, the output from respiratory nuclei lag behind and then catch up after multiple inputs from RA for successive INs. Respiratory pressure recordings may be carried out along with RA recordings to distinguish between possibilities. Another possibility is that dorsal and ventral RA encode differently during INs. Anatomical projections form dorsal RA are proportionally high to the respiratory motor areas vs. ventral RA that send more projections to the motor nucleus (nXIIts) for syringeal control (Roberts et al., 2008).

5.4. Other functions of INs

As properties of IN progression were not related to motor preparation of song, it is possible that these properties in the beginning of the song contribute towards other functions.

Songs have a communicatory role in songbirds. INs are present across many songbird species and have been implicated to play an alerting role before singing, as a marker of species identity (Price, 1979; Richards, Douglas, 1981). The similarity in acoustic properties between INs and songs support that INs and songs are vocalizations as part of the same sequence. Further, recent work on learning of song initiation showed that the number and structure of INs is learned from the tutor (Kalra et al., 2020). This goes against the idea of preparation which is an innate property, and supports a function which allows learning. Communication signals are thus a possible role for IN syllables. It is possible that acoustic and temporal properties of INs are distinct and serve different functions for preparation and communication respectively.

Support for a role in communication also arises from the fact that more number of INs are produced when zebra finches sing to the female (directed) vs alone (undirected). Thus, INs may play a role in seeking female’s attention before song. Hypothesis related to this may be further tested using behavioral preference assays. The increased number of INs during directed song vs undirected song deviates from results of neural preparatory activity that takes longer for self-timed movements vs those in response to stimulus. However, it is difficult to extrapolate the findings of simple movement to complex movement as (i) self-timed movements are not true self-initiated movements and (ii) directed song is more complex and may take longer to initiate due to involvement of more brain areas or may be encoded as a movement with a different final state compared to undirected song. Future experiments to test this will involve comparing activity in different areas during INs of directed and undirected songs.

5.5. Comparing IN progression to preparatory activity

The INs are proposed to be behavioral reflections of preparatory activity. Like neural preparatory activity is related to the aspects of the movement, a subset of IN properties were related in a similar manner to aspects of the upcoming song. The first interval between INs reflects the time for song initiation similar to initial state of neural preparatory activity determining the reaction times. Further experiments disrupting the first interval will help understand if the initial state of IN is causally related to start of the song, analogous to neural preparatory activity.

Other properties of INs did not show activity similar to neural preparatory activity. Neural preparatory activity at the time of movement initiation is related to properties of the upcoming movement. The properties of INs were not related to the song that followed. The evolution of neural preparatory activity has been attributed to optimal feedback control. However, the progresssion of INs was independent of peripheral sensory feedback.

5.6. Comparing representation of IN progression to preparatory activity

The progression of INs was represented in HVC in the form of population of neurons keeping track of past and future syllables. Here activity in HVC neurons was present for repeats of multiple positions of INs. Although the neural preparatory activity from other studies is devoid of such patterns, it has been related to short-term-memory that links past and future events and is encoded by persistent activity. The ramp-to-threshold model were associated with a combination of neurons with transient and persisitent activity involved in preparation. Firing of HVC neurons is different for projection neurons and interneurons and these are densely connected. Thus, similar mechanisms may be involved in how the types of HVC neurons come together to keep track of past and future syllables. These need further investigation.

5.7. Preparatory-movements before self-initiated natural movements

The hypothesis that introductory notes represent preparation suggest behavioral output associated with motor preparation. Can motor behaviors be preparatory? Traditionally, preparatory activity has been defined as activity not associated with behavioral output. However, in the instruction-based paradigms, subjects have to compulsorily fixate and not move before trials begin, and small movements below a threshold are ignored. Thus, we don't know from these experiments if simpler movements can contribute to reaction times or movement preparation. Moreover, complex motor skills have examples of repetitive movements preceding performance. Repetitive bouncing of the ball regularly precedes a tennis serve or a free throw in basket ball. Several other sports report pre-performance routines that potentially help players build focus before movements (Cotterill, 2010). Other animals have also been reported to show repetitive actions before natural movements. The anoliz lizard does repetitive pushups before its head-bobbing display, that adds an alerting component to its communication signal (Ord and Stamps, 2008). Among songbirds, white-crowned sparrows are known to learn better from songs preceded by introductory whistle, suggesting their importance in species identity (Soha and Marler, 2000). In all these behaviors, the introductory actions preceding complex movements enhance the goal of performing the behavior. Thus like INs, these repetitive introductory behaviors may serve a role in preparation of the upcoming movements.

5.8. Initiation of complex natural movements

This thesis uses a natural movement to understand the nature of movement initiation. Earlier studies on externally triggered simple movements show the existence of preparatory neural activity that evolves to reach a consistent state before the movement is executed. The final state reached is within an optimal state space and is related to the variation in the upcoming movement. The evolution through state space during preparation has been attributed to optimal feedback control. The progression of INs to song and their representation in the premotor area HVC make many interpretations about initiation of natural movements. Some are similar and others are different. Similar to simple movements, INs progress to reach a final state. The representation of INs changes as they progress to the song, although it does not follow similar trajectories like simple movements. Unlike simple movements, there is no involvement of feedback during INs and IN production are overt movements before the song. It is possible that these repetitive movements form an additional preparatory step to the preparation that happens before the first IN (Rajan, 2018). It is possible that the changes in activity that occur well before first IN may be similar to preparatory activity for simple movements. Following the first phase of preparatory activity, the repeats of INs occur accompanied by temporal progression to reach the final state before the start of the song. Extrapolating this to all natural movements, it suggests that initiation may involve a period where the brain gets ready by keeping track of past and future neural events before the movement is initiated. Although this activity in the second phase do not represent parameters of the movement, these may be involved in synchronizinging with other areas important for the movement such as respiratory areas in case of song.

5.9. Future Implications

The work from this thesis have made several predictions about the mechanisms that control INs and song initiation. It has shown that INs have a unique way of getting to the song by speeding up their timing. It remains to be understood what drives this timing and open new directions for understanding song initiation. Further, the activity of HVC during INs show that zebra finches although known to encode stereotyped sequences, can represent variable sequences in a manner similar to other birds that produce variable song sequences. Thus, further work in this direction may make zebra finches as models to study both stereotyped and variable sequences in the same animal.

REFERENCES

Afshar A, Santhanam G, Yu BM, Ryu SI, Sahani M, Shenoy K V. (2011) Single-trial neural correlates of arm movement preparation. Neuron 71:555–564.

Amador A, Perl YS, Mindlin GB, Margoliash D (2013) Elemental gesture dynamics are encoded by song premotor cortical neurons. Nature:1–11.

Benichov JI, Benezra SE, Vallentin D, Globerson E, Long M a., Tchernichovski O (2015) The Forebrain Song System Mediates Predictive Call Timing in Female and Male Zebra Finches. Curr Biol 26:309–318.

Betta E, Turatto M (2006) Are you ready? I can tell by looking at your microsaccades. Neuroreport 17:1001–1004.

Bottjer SW, Arnold AP (1984) Role of feedback from the vocal organ I. Maintenance of Steretypcal Vocalisations by Adult Zebra Finches. J Neurosci 4:2387–2396.

Bottjer SW, Miesner E a, Arnold a P (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. Science 224:901–903.

Brainard MS, Doupe AJ (2001) Postlearning Consolidation of Birdsong : Stabilizing Effects of Age and Anterior Forebrain Lesions. 21:2501–2517.

Brainard MS, Doupe AJ (2013) Translating Birdsong : Songbirds as a Model for Basic and Applied Medical Research.

Brenowitz E a (1982) Long-Range Communication in the Red-Winged Blackbird. Behav Ecol Sociobiol 10:29–38.

Cardillo G. (2006). Dunn's Test: a procedure for multiple, not parametric, comparisons. http://www.mathworks.com/matlabcentral/fileexchange/12827

Carlson C, Devinsky O (2009) The excitable cerebral cortex. Fritsch G, Hitzig E. Über die elektrische Erregbarkeit des Grosshirns. Arch Anat Physiol Wissen 1870;37:300-32. Epilepsy Behav 15:131–132.

Chen TW, Li N, Daie K, Svoboda K (2017a) A Map of Anticipatory Activity in Mouse Motor Cortex. Neuron 94:866-879.e4.

Chen Y, Clark O, Woolley SC (2017b) Courtship song preferences in female zebra finches are shaped by developmental auditory experience. Proc R Soc B Biol Sci 284:18–20.

Churchland MM, Afshar A, Shenoy K V. (2006b) A Central Source of Movement Variability. Neuron 52:1085–1096.

Churchland MM, Cunningham JP, Kaufman MT, Ryu SI, Shenoy K V (2010) Cortical preparatory activity: representation of movement or first cog in a dynamical machine? Neuron 68:387–400.

Churchland MM, Santhanam G, Shenoy K V (2006c) Preparatory activity in premotor and motor cortex reflects the speed of the upcoming reach. J Neurophysiol 96:3130–3146.

Churchland MM, Shenoy K V. (2007) Delay of movement caused by disruption of cortical preparatory activity. J Neurophysiol 97:348–359.

Churchland MM, Yu BM, Ryu SI, Santhanam G, Shenoy K V. (2006a) Neural Variability in Premotor Cortex Provides a Signature of Motor Preparation. J Neurosci 26:3697–3712.

Cohen RG, Rosenbaum DA (2007) Directional bias of limb tremor prior to voluntary movement: Research report. Psychol Sci 18:8–12.

Cohen Y, Shen J, Semu D, Leman DP, Liberti WA, Perkins NL, Liberti DC, Kotton D, Gardner TJ (2020) Hidden neural states underlie canary song syntax. Nature 582:539–544.

Coleman MJ, Roy A, Wild JM, Mooney R (2007) Thalamic gating of auditory responses in telencephalic song control nuclei. J Neurosci 27:10024–10036.

Cooper BG, Goller F (2006) Physiological insights into the social-context-dependent changes in the rhythm of the song motor program. J Neurophysiol 95:3798–3809.

Corneil BD, Munoz DP (2014) Overt responses during covert orienting. Neuron 82:1230–1243.

Cotterill S (2010) Pre-performance routines in sport: Current understanding and future directions. Int Rev Sport Exerc Psychol 3:132–153.

Cunnington R, Iansek R, Johnson KA, Bradshaw JL (1997) Movement-related potentials in Parkinson’s disease. Motor imagery and movement preparation. Brain 120:1339–1353.

Daliparthi VK, Tachibana RO, Cooper BG, Hahnloser RHR, Kojima S, Sober SJ, Roberts TF, Southwestern UT, Sciences L, Worth F, Zurich ETH, Network N (2018) Transitioning between preparatory and precisely sequenced neuronal activity in production of a skilled. :1–26.

Danish HH, Aronov D, Fee MS (2017) Rhythmic syllable-related activity in a songbird motor thalamic nucleus necessary for learned vocalizations. PLoS One 12:e0169568.

Derégnaucourt S, Mitra PP, Fehér O, Pytte C, Tchernichovski O (2005) How sleep affects the developmental learning of bird song. Nature 433:710–716.

Donders FC (1969) On the speed of mental processes. Acta Psychol (Amst) 30:412–431.

Egger R, Tupikov Y, Elmaleh M, Jin DZ, Long MA, Egger R, Tupikov Y, Elmaleh M, Katlowitz KA, Benezra SE, Picardo MA, Jin DZ, Long MA (2020) Local Axonal Conduction Shapes the Spatiotemporal Properties of Neural Sequences. Cell 183:537–548.

Erlich JC, Bialek M, Brody CD (2011) A cortical substrate for memory-guided orienting in the rat. Neuron 72:330–343.

Fee MS, Leonardo a (2001) Miniature motorized microdrive and commutator system for chronic neural recording in small animals. J Neurosci Methods 112:83–94.

Fee MS, Scharff C (2010) The songbird as a model for the generation and learning of complex sequential behaviors. ILAR J 51:362–377.

Fleishman LJ (1992) The influence of the sensory system and the environment on motion patterns in the visual displays of anoline lizards and other vertebrates. Am Nat 139.

Forstmeier W, Birkhead TR (2004) Repeatability of mate choice in the zebra finch: Consistency within and between females. Anim Behav 68:1017–1028.

Franz M, Goller F (2003) Respiratory patterns and oxygen consumption in singing zebra finches. J Exp Biol 206:967–978.

Fried I, Mukamel R, Krieman G (2011) Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. Neuron 69:548–562.

Fritsch G, Hitzig E (1870) “Ueber die elektrische Erregbarkeit des Grosshirns.” Arch fur Anat und Physiol Wissenschaftliche Med 37:300–332.

Fujimoto H, Hasegawa T, Watanabe D (2011) Neural coding of syntactic structure in learned vocalizations in the songbird. J Neurosci 31:10023–10033.

Fuster JM, Alexander GE (1971) Neuron activity related to short-term memory. Science (80-) 173:652–654.

Gao Z, Davis C, Thomas AM, Economo MN, Abrego AM, Svoboda K, De Zeeuw CI, Li N (2018) A cortico-cerebellar loop for motor planning. Nature 563:113–116.

Glaze CM, Troyer TW (2006) Temporal structure in zebra finch song: implications for motor coding. J Neurosci 26:991–1005.

Glaze CM, Troyer TW (2007) Behavioral measurements of a temporally precise motor code for birdsong. J Neurosci 27:7631–7639.

Glaze CM, Troyer TW (2012) A generative model for measuring latent timing structure in motor sequences. PLoS One 7:e37616.

Glaze CM, Troyer TW (2013) Development of temporal structure in zebra finch song. J Neurophysiol 109:1025–1035.

Godschalk M, Lemon RN, Kuypers HGJM, Van Der Steen J (1985) The involvement of monkey premotor cortex neurones in preparation of visually cued arm movements. Behav Brain Res 18:143–157.

Goller F, Cooper BG (2004) Peripheral motor dynamics of song production in the zebra finch. In: Annals of the New York Academy of Sciences, pp 130–152.

Gross CG (2007) The discovery of motor cortex and its background. J Hist Neurosci 16:320–331.

Guo Z V., Li N, Huber D, Ophir E, Gutnisky D, Ting JT, Feng G, Svoboda K (2014) Flow of cortical activity underlying a tactile decision in mice. Neuron 81:179–194.

Hahnloser RHR, Kozhevnikov AA, Fee MS (2002) An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature 797:65–70.

Hahnloser RHR, Kozhevnikov AA, Fee MS (2006) Sleep-related neural activity in a premotor and a basal-ganglia pathway of the songbird. J Neurophysiol 96:794–812.

Hanes DP, Schall JD (1996) Neural Control of Voluntary Movement Initiation 140. 274:427–430.

Hessler N a, Doupe a J (1999) Singing-related neural activity in a dorsal forebrain-basal ganglia circuit of adult zebra finches. J Neurosci 19:10461–10481.

Hocherman S, Wise SP (1991) Effects of hand movement path on motor cortical activity in awake, behaving rhesus monkeys. Exp Brain Res 83:285–302.

Horita H, Wada K, Jarvis ED (2008) Early onset of deafening-induced song deterioration and differential requirements of the pallial-basal ganglia vocal pathway. Eur J Neurosci 28:2519–2532.

Immelmann K (1969) Song development in the zebra finch and other estrildid finches. Bird vocalizations (Hinde RA, ed):61–74.

James LS, Sakata JT (2019) Developmental modulation and predictability of age-dependent vocal plasticity in adult zebra finches. Brain Res 1721:146336.

Kalra S, Yawatkar V, James LS, Sakata JT, Rajan R (2020) Introductory gestures before songbird vocal displays are shaped by learning and biological predispositions. Proc R Soc B Biol Sci.

Kandel ER, Schwartz JH, Jessell TM, Siegelbaum S, Hudspeth AJ, Mack S (2000) Principles of neural science. McGraw-hill New York.

Kao MH, Wright BD, Doupe AJ (2008) Neurons in a forebrain nucleus required for vocal plasticity rapidly switch between precise firing and variable bursting depending on social context. J Neurosci 28:13232–13247.

Katz LC, Gurney ME (1981) Auditory responses in the zebra finch's motor system for song. Brain Res 221:192–197.

Kaufman MT, Churchland MM, Ryu SI, Shenoy K V (2014) Cortical activity in the null space: permitting preparation without movement. Nat Neurosci 17:440–448.

Keele SW (1968) Movement control in skilled motor performance. Psychol Bull 70:387–403.

Kent RD, Kent RD (2000) Research on speech motor control and its disorders: a review and prospective. J Commun Disord 33:391–427; quiz 428.

Kershenbaum A, Bowles AE, Freeberg TM, Jin DZ, Lameira AR, Bohn K (2014) Animal vocal sequences: Not the Markov chains we thought they were. Proc R Soc B Biol Sci 281.

Kojima S, Kao M, Doupe A (2013) Task-related "cortical" bursting depends critically on basal ganglia input and is linked to vocal plasticity. Proc Natl … 110:4756–4761.

Konishi M (1965) The Role of Auditory Feedback in the Control of Vocalization in the White-Crowned Sparrow. Z Tierpsychol 22:770–783.

Kornhuber HH, Deecke L (1965) Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. Pflugers Arch Gesamte Physiol Menschen Tiere 284:1–17.

Kosche G, Vallentin D, Long MA (2015) Interplay of inhibition and excitation shapes a premotor neural sequence. J Neurosci 35:1217–1227.

Kozhevnikov AA, Fee MS (2007) Singing-related activity of identified HVC neurons in the zebra finch. J Neurophysiol 97:4271–4283.

Kozhevnikov AA, Fee MS (2007b) Singing-related activity of identified HVC neurons in the zebra finch. J Neurophysiol 97:4271–4283.

Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D (2017) Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron 93:480–490.

Kutas M, Donchin E (1980) Preparation to respond as manifested by movement-related brain potentials. Brain Res 202:95–115.

Lauay C, Gerlach NM, Adkins-Regan E, Devoogd TJ (2004) Female zebra finches require early song exposure to prefer high-quality song as adults. Anim Behav 68:1249–1255.

Lee IH, Assad JA (2003) Putaminal activity for simple reactions or self-timed movements. J Neurophysiol 89:2528–2537.

Lee IH, Seitz AR, Assad JA (2006) Activity of tonically active neurons in the monkey putamen during initiation and withholding of movement. J Neurophysiol 95:2391–2403.

Leonardo A, Fee MS (2005) Ensemble coding of vocal control in birdsong. J Neurosci 25:652–661.

Leyton AS., Sherrington C. (1917) Observations on the excitable cortex of the chimpanzee, orang-utan and gorilla. Q J Exp Physiol 11:135–222.

Li N, Chen TW, Guo Z V., Gerfen CR, Svoboda K (2015) A motor cortex circuit for motor planning and movement. Nature 519:51–56.

Li N, Daie K, Svoboda K, Druckmann S (2016) Robust neuronal dynamics in premotor cortex during motor planning Preparatory activity in ALM. Nature 532:459.

Lombardino AJ, Nottebohm F (2000) Age at deafening affects the stability of learned song in adult male zebra finches. J Neurosci 20:5054–5064.

Long M a, Jin DZ, Fee MS (2010) Support for a synaptic chain model of neuronal sequence generation. Nature 468:394–399.

Long MA, Fee MS (2008) Using temperature to analyse temporal dynamics in the songbird motor pathway. Nature 456:189–194.

Lynch GF, Okubo TS, Hanuschkin A, Hahnloser RHR, Fee MS (2016) Rhythmic Continuous-Time Coding in the Songbird Analog of Vocal Motor Cortex. Neuron 90:877–892.

Maimon G, Assad JA (2006) Parietal area 5 and the initiation of self-timed movements versus simple reactions. J Neurosci 26:2487–2498.

Mandelblat-Cerf Y, Fee MS (2014) An Automated Procedure for Evaluating Song Imitation. 9:1–13.

McCasland JS (1987) Neuronal control of bird song production. J Neurosci 7:23–39 Available at: http://www.ncbi.nlm.nih.gov/pubmed/3806194.

McCasland JS, Konishi M (1981) Interaction between auditory and motor activities in an avian song control nucleus. Proc Natl Acad Sci U S A 78:7815–7819.

Méndez JM, Dall'asén AG, Goller F (2010) Disrupting vagal feedback affects birdsong motor control. J Exp Biol 213:4193–4204.

Messier J, Kalaska JF (2000) Covariation of primate dorsal premotor cell activity with direction and amplitude during a memorized-delay reaching task. J Neurophysiol 84:152–165.

Mooney R (1992) Synaptic basis for developmental plasticity in a birdsong nucleus. J Neurosci 12:2464–2477.

Mooney R, Prather JF (2005) The HVC microcircuit: The synaptic basis for interactions between song motor and vocal plasticity pathways. J Neurosci 25:1952–1964.

Morris D (1954) The Reproductive Behaviour of the Zebra Finch (Poephila Guttata), With Special Reference To Pseudofemale Behaviour and Displacement Activities. Behaviour 6:271–322.

Murakami M, Vicente MI, Costa GM, Mainen ZF (2014) Neural antecedents of self-initiated actions in secondary motor cortex. Nat Neurosci 17

Nixdorf-Bergweiler BE, Bischof HJ (2007) A stereotaxic atlas of the brain of the zebra finch.

Nordeen KW, Nordeen EJ (1992) Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. Behav Neural Biol 57:58–66.

Nottebohm F, Arnold A (1976) Sexual Dimorphism in Vocal Control Areas of the Songbird Brain. Science (80-):211–213.

Nottebohm F, Paton JA, & Kelley DB (1982) Connections of vocal control nuclei in the canary telencephalon. J Comp Neurol 207:344–357.

Nottebohm F, Stokes TM, Leonard CM (1976) Central control of song in the canary, Serinus canarius. J Comp Neurol 165:457–486.

Nottebohm F, Stokes TM, Leonard CM (1976b) Central control of song in the canary, Serinus canarius. J Comp Neurol 165:457–486.

Nowicki S, Searcy WA, Peters S (2002) Brain development, song learning and mate choice in birds: A review and experimental test of the “nutritional stress hypothesis.” J Comp Physiol A Neuroethol Sensory, Neural, Behav Physiol 188:1003–1014.

Okubo TS, Mackevicius EL, Fee MS (2014) In vivo recording of single-unit activity during singing in zebra finches. Cold Spring Harb Protoc 2014:1273–1283.

Ord TJ, Stamps JA (2008) Alert signals enhance animal communication in “noisy” environments. Proc Natl Acad Sci U S A 105:18830–18835.

Otchy TM, Ölveczky BP (2012) Design and assembly of an ultra-light motorized microdrive for chronic neural recordings in small animals. J Vis Exp.

Penfield W, Boldrey E (1937) Somatic Motor and Sensory Representation in Man. Brain:389–443.

Price PH (1979) Developmental determinants of structure in zebra finch song. J Comp Physiol Psychol 93:260–277.

Pytte CL, Gerson M, Miller J, Kirn JR (2007) Increasing stereotypy in adult zebra finch song correlates with a declining rate of adult neurogenesis. Dev Neurobiol 67:1699–1720.

Rajan R (2018) Pre-bout neural activity changes in premotor nucleus HVC correlate with successful initiation of learned song sequence. J Neurosci 38:5925–5938.

Rajan R, Doupe AJ (2013) Behavioral and Neural Signatures of Readiness to Initiate a Learned Motor Sequence. Curr Biol:1–7.

Rao D, Kojima S, Rajan R (2019) Sensory feedback independent pre-song vocalizations correlate with time to song initiation. J Exp Biol:jeb.199042.
doi: 10.1242/jeb.199042. www.https://journals.biologists.com/jeb/article/222/7/jeb199042/20647

Richards, Douglas G (1981) Alerting and Message Components in Songs of Rufous-Sided Towhees. Behaviour 76:223–249.

Riebel K, Naguib M, Gil D (2009) Experimental manipulation of the rearing environment influences adult female zebra finch song preferences. Anim Behav 78:1397–1404.

Rigato J, Murakami M, Mainen Z (2014) Spontaneous decisions and free will: Empirical results and philosophical considerations. Cold Spring Harb Symp Quant Biol 79:177–184.

Roberts TF, Klein ME, Kubke MF, Wild JM, Mooney R (2008) Telencephalic neurons monosynaptically link brainstem and forebrain premotor networks necessary for song. J Neurosci 28:3479–3489.

Romo R, Schultz W (1987) Neuronal activity preceding self-initiated or externally timed arm movements in area 6 of monkey cortex. Exp Brain Res 67:656–662.

Rosenbaum D a (1980) Human movement initiation: specification of arm, direction, and extent. J Exp Psychol 109:444–474.

Roy A, Mooney R (2007) Auditory plasticity in a basal ganglia-forebrain pathway during decrystallization of adult birdsong. J Neurosci 27:6374–6387.

Scharff C, Nottebohm F (1991) A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. J Neurosci 7.

Shenoy K V., Kaufman MT, Sahani M, Churchland MM (2011) A dynamical systems view of motor preparation: Implications for neural prosthetic system design. Prog Brain Res 192:33–58.

Shenoy K V, Sahani M, Churchland MM (2013) Cortical control of arm movements: a dynamical systems perspective. Annu Rev Neurosci 36:337–359.

Shiovitz KA. (1975) The Process of Species-Specific Song Recognition by the Indigo Bunting , Passerina cyanea , and Its Relationship to the Organization of Avian Acoustical Behavior Published by : Brill Stable URL : Behaviour 55:128–179.

Simpson H, Vicario D (1990) Brain Pathways for Learned and Unlearned Vocalizations Differ in Zebra Finches. J Neurosci 10.

Sober SJ, Wohlgemuth MJ, Brainard MS (2008) Central Contributions to Acoustic Variation in Birdsong. J Neurosci 28:10370–10379.

Soha JA, Marler P (2000) A species-specific acoustic cue for selective song learning in the white-crowned sparrow. Anim Behav 60:297–306.

Sohrabji F, Nordeen EJ, Nordeen KW (1990) Selective Impairment of Song Learning following Lesions of a Forebrain Nucleus in the Juvenile Zebra Finch. 63:51–63.

Sossinka R, Böhner J (1980) Song Types in the Zebra Finch. Z Tierpsychol 53:123–132.

Spencer K., Buchanan K., Goldsmith a. ., Catchpole C. (2003) Song as an honest signal of developmental stress in the zebra finch (Taeniopygia guttata). Horm Behav 44:132–139.

Spiro JE, Dalva MB, Mooney R, Basista MJ, Elliott KC, Wu W, Hyson RL, Bertram R, Larson TA, Wang T, Gale SD, Miller KE, Thatra NM, Caras ML, Perkel DJ, Brenowitz EA (1999) Long-Range Inhibition Within the Zebra Finch Song Nucleus RA Can Coordinate the Firing of Multiple Projection Neurons Long-Range Inhibition Within the Zebra Finch Song Nucleus RA Can Coordinate the Firing of Multiple Projection Neurons. J Neurophysiol 81:3007–3020.

Suri H, Rajan R (2018) Distance-dependent changes in courtship song amplitude reflect song state changes. BioRxiv, 277210.

Suthers RA, Margoliash D (2002) Motor control of birdsong. Curr Opin Neurobiol 12:684–690.

Svoboda K, Li N (2018) Neural mechanisms of movement planning: motor cortex and beyond. Curr Opin Neurobiol 49:33–41.

Tamura M, Marler P (1962) Song " Dialects " in Three Populations of White-Crowned Sparrows. Condor 64:368–377.

Tanji J, Evarts E V. (1976) Anticipatory activity of motor cortex neurons in relation to direction of an intended movement. J Neurophysiol 39:1062–1068.

Tchernichovski O, Nottebohm F, Ho C, Pesaran B, Mitra P, Systems N (2000) A procedure for an automated measurement of song similarity. Anim Behav 59:1167–1176.

Tomaszycki ML, Adkins-Regan E (2005) Experimental alteration of male song quality and output affects female mate choice and pair bond formation in zebra finches. Anim Behav 70:785–794.

Vallentin D, Long M a. (2015) Motor Origin of Precise Synaptic Inputs onto Forebrain Neurons Driving a Skilled Behavior. J Neurosci 35:299–307.

Vicario DS (1991) Contributions of syringeal muscles to respiration and vocalization in the zebra finch. J Neurobiol 22:63–73.

Vicario DS (2004) Using learned calls to study sensory-motor integration in songbirds. Ann N Y Acad Sci 1016:246–262.

Vu ET, Schmidt MF, Mazurek ME (1998) Interhemispheric coordination of premotor neural activity during singing in adult zebra finches. J Neurosci 18:9088–9098.

Vyssotski AL, Stepien AE, Keller GB, Hahnloser RHR (2016) A Neural Code That Is Isometric to Vocal Output and Correlates with Its Sensory Consequences. PLoS Biol 14:1–21.

Wang CA, Brien DC, Munoz DP (2015) Pupil size reveals preparatory processes in the generation of pro-saccades and anti-saccades. Eur J Neurosci 41:1102–1110.

Wang D, Forstmeier W, D’Amelio PB, Martin K, Kempenaers B (2021) Is female mate choice repeatable across males with nearly identical songs? Anim Behav 181:137–149.

Warren TL, Charlesworth JD, Tumer EC, Brainard MS (2012) Variable sequencing is actively maintained in a well learned motor skill. J Neurosci 32:15414–15425.

Wild JM, Goller F, Suthers RA (1998) Inspiratory muscle activity during bird song. J Neurobiol 36:441–453.

Williams H, Kilander K, Sotanski M Lou (1993) Untutored song, reproductive success and song learning. Anim Behav 45:695–705.

Williams H, McKibben JR (1992) Changes in stereotyped central motor patterns controlling vocalization are induced by peripheral nerve injury. Behav Neural Biol 57:67–78.

Williams H, Vicario DS (1993) Temporal patterning of song production: participation of nucleus uvaeformis of the thalamus. J Neurobiol 24:903–912.

Wittenbach JD, Bouchard KE, Brainard MS, Jin DZ (2015) An Adapting Auditory-motor Feedback Loop Can Contribute to Generating Vocal Repetition. PLOS Comput Biol 11:e1004471.

Wong AL, Haith AM, Krakauer JW (2015) Motor planning. Neuroscientist 21:385–398.

Woolley SC, Doupe AJ (2008) Social context-induced song variation affects female behavior and gene expression. PLoS Biol 6:0525–0537.

Woolley SC, Rajan R, Joshua M, Doupe AJ (2014) Emergence of context-dependent variability across a basal ganglia network. Neuron 82:208–223.

Yu a C, Margoliash D (1996) Temporal hierarchical control of singing in birds. Science 273:1871–1875.

Zann R (1993) Sequence and Evolution of Song Elements in Wild Australian Zebra Finches. Auk 110:702–715.

Zann RA (1996) The zebra finch: A synthesis of laboratory and field studies. Oxford University Press, Oxford.

P-value of statistical tests in Chapter 3				
S.No	Data compared	Parameters	Test	P-value
1	IN-song standard deviation of acoustic features	Duration (ms)	signrank	pvalue 0.017378
2	IN-song standard deviation of acoustic features	Mean Frequency (Hz)	signrank	pvalue 0.078731
3	IN-song standard deviation of acoustic features	Entropy	signrank	pvalue 0.00064304
4	IN-song standard deviation of acoustic features	Log Amplitude (dB)	signrank	pvalue 0.00053124
5	IN-song standard deviation of acoustic features	Pitch Goodness	signrank	pvalue 0.00043778
6	IN-song standard deviation of acoustic features	Frequency Modulation (°)	signrank	pvalue 0.00043778
7	IN-song standard deviation of acoustic features	Amplitude Modulation (1/ms)	signrank	pvalue 0.00043778
8	IN-song standard deviation of acoustic features	Entropy Variance	signrank	pvalue 0.53493
9	IN-song mean of acoustic features	Duration (ms)	signrank	pvalue 0.0004
10	IN-song mean of acoustic features	Mean Frequency (Hz)	signrank	pvalue 0.0011
11	IN-song mean of acoustic features	Entropy	signrank	pvalue 0.3520
12	IN-song mean of acoustic features	Log Amplitude (dB)	signrank	pvalue 0.0004
13	IN-song mean of acoustic features	Pitch Goodness	signrank	pvalue 0.0151
14	IN-song mean of acoustic features	Frequency Modulation (°)	signrank	pvalue 0.4380
15	IN-song mean of acoustic features	Amplitude Modulation (1/ms)	signrank	pvalue 0.0016
16	IN-song mean of acoustic features	Entropy Variance	signrank	pvalue 0.0019
17	Change in days (day-day: <1yr to >1yr, >1 yr)	Mean IN number	Kruskal-Wallis Test	Pvalue 0.021175, Dunn’s test Qcrit=2.39
18	Change in days (day-day: <1yr to >1yr, >1 yr)	mean last IN Interval (ms)	Kruskal-Wallis Test	pvalue 0.69735
19	Change in days (day-day: <1yr to >1yr, >1 yr)	mean ratio of intervals	Kruskal-Wallis Test	pvalue 0.98699
20	Change in days (day-day: <1yr to >1yr, >1 yr)	mean first IN interval (ms)	Kruskal-Wallis Test	pvalue 0.18363
21	Change in days (day-day: <1yr to >1yr, >1 yr)	mean last IN acoustic D2L	Kruskal-Wallis Test	pvalue 0.2768
22	Change in days (day-day: <1yr to >1yr, >1 yr)	mean ratio of acoustic D2L	Kruskal-Wallis Test	pvalue 0.78969
23	Change in days (day-day: <1yr to >1yr, >1 yr)	CV IN number	Kruskal-Wallis Test	pvalue 0.38459
24	Change in days (day-day: <1yr to >1yr, >1 yr)	CV last IN Interval	Kruskal-Wallis Test	pvalue 0.43052
25	Change in days (day-day: <1yr to >1yr, >1 yr)	CV first IN interval	Kruskal-Wallis Test	pvalue 0.14958

26	Change in days (day-day: <1yr to >1yr, >1 yr)	IN mean Duration (ms)	Kruskal-Wallis Test	pvalue 0.076222
27	Change in days (day-day: <1yr to >1yr, >1 yr)	IN mean Mean Frequency (Hz)	Kruskal-Wallis Test	pvalue 0.95198
28	Change in days (day-day: <1yr to >1yr, >1 yr)	IN mean Entropy	Kruskal-Wallis Test	pvalue 0.50265
29	Change in days (day-day: <1yr to >1yr, >1 yr)	IN mean Log Amplitude (dB)	Kruskal-Wallis Test	pvalue 0.46338
30	Change in days (day-day: <1yr to >1yr, >1 yr)	IN mean Pitch Goodness	Kruskal-Wallis Test	pvalue 0.28124
31	Change in days (day-day: <1yr to >1yr, >1 yr)	mean FM (°)	Kruskal-Wallis Test	pvalue 0.067983
32	Change in days (day-day: <1yr to >1yr, >1 yr)	IN mean AM (1/ms	Kruskal-Wallis Test	Pvalue 0.039514, Dunn’s test Qcrit=2.39
33	Change in days (day-day: <1yr to >1yr, >1 yr)	IN mean Entropy Variance	Kruskal-Wallis Test	pvalue 0.083511
34	Change in days (day-day: <1yr to >1yr, >1 yr)	IN SD Duration (ms)	Kruskal-Wallis Test	pvalue 0.07248
35	Change in days (day-day: <1yr to >1yr, >1 yr)	IN SD Mean Frequency (Hz)	Kruskal-Wallis Test	pvalue 0.067983
36	Change in days (day-day: <1yr to >1yr, >1 yr)	IN SD Entropy	Kruskal-Wallis Test	pvalue 0.50816
37	Change in days (day-day: <1yr to >1yr, >1 yr)	IN SD Log Amplitude (dB)	Kruskal-Wallis Test	pvalue 0.82742
38	Change in days (day-day: <1yr to >1yr, >1 yr)	IN SD Pitch Goodness	Kruskal-Wallis Test	pvalue 0.92182
39	Change in days (day-day: <1yr to >1yr, >1 yr)	SD FM (°)	Kruskal-Wallis Test	pvalue 0.1375
40	Change in days (day-day: <1yr to >1yr, >1 yr)	IN SD AM (1/ms)	Kruskal-Wallis Test	pvalue 0.058635
41	Change in days (day-day: <1yr to >1yr, >1 yr)	IN SD Entropy Variance	Kruskal-Wallis Test	pvalue 0.18018
42	Change in days (day-day: <1yr to >1yr, >1 yr)	Mean Motif length (ms)	Kruskal-Wallis Test	Pvalue 0.0027448, Dunn’s test Qcrit=2.39
43	Change in days (day-day: <1yr to >1yr, >1 yr)	CV motif length	Kruskal-Wallis Test	pvalue 0.54546
44	Change in days (day-day: <1yr to >1yr, >1 yr)	Within-day temporal similarity	Kruskal-Wallis Test	pvalue 0.2413
45	Change in days (day-day: <1yr to >1yr, >1 yr)	Across-day temporal similarity	Kruskal-Wallis Test	pvalue 0.064197
46	Age (<1yr;1-2yr;>2yr)	Similarity Index within-bout	Kruskal-Wallis Test	pvalue 0.023119
47	Age (<1yr;1-2yr;>2yr)	Similarity Index First motif	Kruskal-Wallis Test	pvalue 0.21925
48	Change in days (day-day: <1yr to >1yr, >1 yr)	mean SimIndex	Kruskal-Wallis Test	Pvalue 0.0099177 Dunn’s test Qcrit=2.39
49	Change in days (day-day: <1yr to >1yr, >1 yr)	mean MotifTempoSim	Kruskal-Wallis Test	Pvalue 0.0061719 Dunn’s test Qcrit=2.39
50	Change in days (day-day: <1yr to >1yr, >1 yr)	mean motiflengths (ms)	Kruskal-Wallis Test	Pvalue 0.014835 Dunn’s test Qcrit=2.39
51	Change in days (day-day: <1yr to >1yr, >1 yr)	mean num motif	Kruskal-Wallis Test	pvalue 0.078624
52	Change in days (day-day: <1yr to >1yr, >1 yr)	CV num motif	Kruskal-Wallis Test	pvalue 0.9075
53	Change in days (day-day: <1yr to >1yr, >1 yr)	meansong interval (ms)	Kruskal-Wallis Test	pvalue 0.38284
54	Change in days (day-day: <1yr to >1yr, >1 yr)	meanDuration (ms)	Kruskal-Wallis Test	pvalue 0.05989
55	Change in days (day-day: <1yr to >1yr, >1 yr)	meanMean Frequency (°)	Kruskal-Wallis Test	pvalue 0.13646

56	Change in days (day-day: <1yr to >1yr, >1 yr)	meanEntropy	Kruskal-Wallis Test	pvalue 0.92387
57	Change in days (day-day: <1yr to >1yr, >1 yr)	meanLog Amplitude (dB)	Kruskal-Wallis Test	pvalue 0.25542
58	Change in days (day-day: <1yr to >1yr, >1 yr)	meanPitch Goodness	Kruskal-Wallis Test	Pvalue 0.00077881 Dunn’s test Qcrit=2.39
59	Change in days (day-day: <1yr to >1yr, >1 yr)	MeanFM (°)	Kruskal-Wallis Test	pvalue 0.24087
60	Change in days (day-day: <1yr to >1yr, >1 yr)	meanAM (1/ms)	Kruskal-Wallis Test	pvalue 0.050535
61	Change in days (day-day: <1yr to >1yr, >1 yr)	meanEntropy Variance	Kruskal-Wallis Test	pvalue 0.9008
62	Change in days (day-day: <1yr to >1yr, >1 yr)	SDDuration (ms)	Kruskal-Wallis Test	pvalue 0.36448
63	Change in days (day-day: <1yr to >1yr, >1 yr)	SDMean Frequency (Hz)	Kruskal-Wallis Test	pvalue 0.94465
64	Change in days (day-day: <1yr to >1yr, >1 yr)	SDEntropy	Kruskal-Wallis Test	pvalue 0.94396
65	Change in days (day-day: <1yr to >1yr, >1 yr)	SDLog Amplitude (dB)	Kruskal-Wallis Test	pvalue 0.42457
66	Change in days (day-day: <1yr to >1yr, >1 yr)	SDPitch Goodness	Kruskal-Wallis Test	pvalue 0.49956
67	Change in days (day-day: <1yr to >1yr, >1 yr)	SDFM (°)	Kruskal-Wallis Test	Pvalue 0.0059411 Dunn’s test Qcrit=2.39
68	Change in days (day-day: <1yr to >1yr, >1 yr)	SDAM (1/ms)	Kruskal-Wallis Test	Pvalue 0.01197 Dunn’s test Qcrit=2.39
69	Change in days (day-day: <1yr to >1yr, >1 yr)	SDEntropy Variance	Kruskal-Wallis Test	pvalue 0.65763
70	Change in days (day-day: <1yr to >1yr, >1 yr)	mean num motif	Kruskal-Wallis Test	pvalue 0.071358
71	Change in days (day-day: <1yr to >1yr, >1 yr)	CV num motif	Kruskal-Wallis Test	pvalue 0.9075
72	Change in days (day-day: <1yr to >1yr, >1 yr)	mean num motif	Kruskal-Wallis Test	pvalue 0.071358
73	Change in days (day-day: <1yr to >1yr, >1 yr)	CV num motif	Kruskal-Wallis Test	pvalue 0.9075
74	Change in days (day-day: <1yr to >1yr, >1 yr)	mean num motif	Kruskal-Wallis Test	pvalue 0.071358
75	Change in days (day-day: <1yr to >1yr, >1 yr)	CV num motif	Kruskal-Wallis Test	pvalue 0.9075
76	Change in days (day-day: <1yr to >1yr, >1 yr)	SC	Kruskal-Wallis Test	pvalue 0.91088
77	Change in days (day-day: <1yr to >1yr, >1 yr)	SE	Kruskal-Wallis Test	Pvalue 0.024716 Dunn’s test Qcrit=2.39

© 2019. Published by The Company of Biologists Ltd | Journal of Experimental Biology (2019) 222, jeb199042. doi:10.1242/jeb.199042

RESEARCH ARTICLE

Sensory feedback independent pre-song vocalizations correlate with time to song initiation

Divya Rao[1], Satoshi Kojima[2] and Raghav Rajan[1,*]

ABSTRACT

The song of the adult male zebra finch is a well-studied example of a learned motor sequence. Song bouts begin with a variable number of introductory notes (INs) before actual song production. Previous studies have shown that INs progress from a variable initial state to a stereotyped final state before each song. This progression is thought to represent motor preparation, but the underlying mechanisms remain poorly understood. Here, we assessed the role of sensory feedback in the progression of INs to song. We found that the mean number of INs before song and the progression of INs to song were not affected by removal of two sensory feedback pathways (auditory or proprioceptive). In both feedback-intact and feedback-deprived birds, the presence of calls (other non-song vocalizations), just before the first IN, was correlated with fewer INs before song and an initial state closer to song. Finally, the initial IN state correlated with the time to song initiation. Overall, these results show that INs do not require real-time sensory feedback for progression to song. Rather, our results suggest that changes in IN features and their transition to song are controlled by internal neural processes, possibly involved in getting the brain ready to initiate a learned movement sequence.

KEY WORDS: Bird song, Preparatory vocalizations, Zebra finch, Motor preparation, Sequence initiation

INTRODUCTION

The song motif (referred to as song) of the adult male zebra finch, consisting of a stereotyped sequence of sounds (syllables) interleaved with silent gaps (Fig. 1), is a well-established model for understanding learned movement sequences (Fee and Scharff, 2010). How such learned movement sequences are initiated in the brain remains poorly understood. Song is learned by young birds from a conspecific tutor during a critical period (Fee and Scharff, 2010). While song is typically part of a courtship ritual for mate attraction, birds also sing when they are alone (undirected song) (Sossinka and Böhner, 1980; Zann, 1996), making this an excellent model system to study motor preparation before self-initiated, learned movement sequences.

Song is preceded by the bird repeating a short vocalization called an introductory note (IN; Fig. 1) (Price, 1979; Sossinka and Böhner, 1980). Each song bout consists of a variable number of such INs followed by multiple repeats of the song. We have previously shown that intervals between successive INs and the acoustic properties of successive INs progress from a variable initial state (first IN in each song bout) to a more consistent ‘ready’ state (last IN in each song bout) just before the start of each song (Rajan and Doupe, 2013). Given the similarity to the reduction in variability associated with neural preparatory activity before the onset of simple movements (Churchland et al., 2006c), INs may represent vocalizations that help prepare the zebra finch brain to produce song. However, the mechanisms underlying IN progression to song remain unclear.

One possibility is that real-time sensory feedback could drive the progression of INs to song. Sensory feedback is important for song learning and maintenance in zebra finches (Konishi, 1965; Nordeen and Nordeen, 1992). In a related species, the Bengalese finch, recent work has shown that removal of auditory feedback changes the repeat number of individual syllables within song (Wittenbach et al., 2015). As INs are also repeating syllables, real-time sensory feedback could drive IN progression to song. Consistent with this hypothesis, previous studies disrupting proprioceptive feedback or auditory feedback have reported changes in the number of INs before song in some birds (Bottjer and Arnold, 1984). However, these changes have not been quantified rigorously and the specificity of these changes to removal of feedback has not been determined.

In order to assess the role of sensory feedback, we analyzed the number and properties of INs soon after removal of two important forms of sensory feedback, namely proprioceptive feedback from the syringeal muscles (Bottjer and Arnold, 1984; Vicario, 1991; Williams and McKibben, 1992) and auditory feedback (Konishi, 1965). We found that mean IN number before song and progression of INs to song were not affected by removal of either form of feedback. Further, the progression of INs to song was not affected by removal of neural input to the syringeal muscles. Finally, we found fewer INs and a quicker transition to song when the first IN was produced soon after calls (non-song vocalizations that are different from INs and song syllables). These data demonstrate that INs do not provide sensory feedback. Rather, INs may reflect internal neural processes, potentially involved in getting the zebra finch brain ‘ready’ to produce the learned song sequence.

MATERIALS AND METHODS

Experimental procedures performed at IISER Pune were approved by the Institute Animal Ethical Committee in accordance with the guidelines of the Committee for the Purpose of Control and Supervision of Experiments on Animals (CPCSEA, New Delhi). Experiments performed at UCSF (CA, USA) were approved by the UCSF Institutional Animal Care and Use Committee in accordance with NIH guidelines.

Birds and song recording

All birds (n=42) used in this study were >100 days post-hatch at the time of the experiment and were either purchased from an outside

---

[1]Division of Biology, Indian Institute of Science Education and Research Pune, Pune, Maharashtra 411008, India. [2]Department of Structure and Function of Neural Network, Korea Brain Research Institute, Dong-gu, Daegu 701-300, Republic of Korea.

*Author for correspondence (raghav@iiserpune.ac.in)

R.R., 0000-0002-7770-3711

Received 1 January 2019; Accepted 11 March 2019

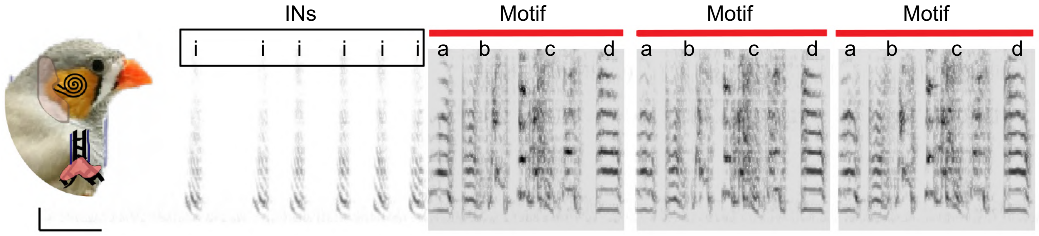

Fig. 1. Spectrogram of a zebra finch song bout with introductory notes (INs) and motifs. A male zebra finch showing two important sensory feedback sources, namely auditory feedback from the cochlea and proprioceptive feedback from the syringeal muscles. Spectrogram of the song of an adult male zebra finch. ‘i’ denotes INs and ‘a’ to ‘d’ represent the syllables of the song motif (gray shading). Scale bar: 200 ms (horizontal) and 1000 Hz (vertical).

vendor (n=13) or bred at IISER Pune (n=16) or UCSF (n=13). Birds were kept in separate sound isolation boxes (Newtech Engineering Systems, Bangalore, India, or Acoustic Systems, Austin, TX, USA) for the duration of the experiment. All songs were recorded by placing a microphone (AKG Acoustics C417PP omnidirectional condenser microphone or B3 lavalier microphone, Countryman Associates, CA, USA) at the top of the cage. For birds in the tracheosyringeal nerve surgery (ts-cut) and sham-surgery groups (see below), we kept the position of the microphone the same for recording songs before and after surgical manipulations. Signals from the microphone were amplified using a mixer (Behringer XENYX 802) and then digitized on a computer at a sampling rate of 44,100 Hz using custom-written software. Songs were recorded in ‘triggered’ mode before and after surgery, such that data were saved when the microphone signal crossed a pre-set threshold. Along with the data that crossed the threshold, 1–3 s of data before and after threshold crossing were also saved. For a subset of birds, data were saved in ‘continuous’ mode, i.e. all of the data for the entire recording period. All songs were recorded in the ‘undirected’ condition. Songs of three of the birds used for the analysis of calls and their influence on song initiation have been used in a previous study for analysis of INs before song (Rajan and Doupe, 2013). The influence of calls on song initiation was not considered in the previous study. For the analysis of day-to-day changes in IN number and properties, we used data from 14 birds that were recorded on two different days (range: 1–3 days apart). Of these 14 birds, one bird was also used at a later time point for ts-cut surgery with a new set of pre- and post-surgery recordings and nine birds were also used for analysis of the influence of calls on INs. Pre-surgery recordings for 18/21 birds (n=5 ts-cut, n=6 sham surgery and n=7 deaf) were performed 0–2 days before surgery. For the remaining three birds (n=3 ts-cut), pre-surgery recordings were made 18, 14 and 5 days before surgery, respectively.

Tracheosyringeal nerve cut and sham surgery

Tracheosyringeal nerves were surgically cut using previously described protocols (Bottjer and Arnold, 1984; Vicario, 1991; Williams and McKibben, 1992). Briefly, birds (n=9) were deeply anesthetized by intramuscular injection of ketamine (30 mg kg^{-1}), xylazine (3 mg kg^{-1}) and diazepam (7 mg kg^{-1}). Absence of a response to toe pinch was used to assess the depth of anesthesia. Birds were then placed on a platform with the ventral side facing up. A rolled tissue under the neck served to stretch and give easy access to the throat. Feathers were plucked and an incision of ∼10 mm was made. The trachea was exposed by removal of fat tissue. Using fine forceps, the tracheosyringeal (ts) nerve bundle on either side of the trachea was pulled away from the trachea and part of the nerve (n=9 birds, median length cut 4 mm, range 2–7 mm) was cut out on both sides using spring scissors (Fine Science Tools, Foster City, CA, USA). The skin was then glued using tissue adhesive (Vetbond, 3M). For sham surgeries (n=6), the same procedure was followed but the ts nerves were not cut. In two of the sham-surgery birds, some cuts were made on the thick membrane enclosing the esophagus. Birds typically resumed singing within 10 days of surgery. We considered songs produced on the second day of singing after surgery (sham surgery: 2–5 days and ts-cut: 3–10 days after surgery) for analysis because of the higher number of songs. For one bird, we did not have pre-surgery songs in the undirected condition, so we excluded this bird from analyses involving comparison with the pre-surgery condition (Figs 2–6). Data from this bird were included only for analysis of the influence of calls on the number and properties of INs (Figs 7 and 8). For the ts-cut group, birds with first motif syllables that were significantly different in duration from the IN were chosen; this made it easier to recognize the onset of the motif after surgery. The number of INs was not a consideration while choosing birds for the sham-surgery or ts-cut groups. For the first two birds in which we attempted ts nerve surgery, we did not observe any changes to song after surgery. While doing the other ts nerve surgeries, we realized that we had not cut the ts nerve in these two birds and so both birds were assigned to the sham-surgery group. The rest of the sham-surgery birds were chosen at random based on availability in our colony. IN number and song were not taken into consideration.

Deafening

Deafening was done by bilateral removal of the cochlea under equithesin anesthesia using previously described protocols (Kojima et al., 2013; Konishi, 1965). All of the deaf birds (n=7) were also used in a previous study that examined the effects of deafening on song (Kojima et al., 2013). Here, we only analyzed the effects of deafening on IN number and properties. As we were interested in the role of real-time auditory feedback in progression from INs to song, we only analyzed IN number and properties for songs recorded 1 day post-deafening. Birds for deafening were chosen on the basis of their motif structure; the number of INs was not a consideration.

Data analysis

All analysis was performed using custom-written scripts in MATLAB. All data and scripts for analysis are available on request from the corresponding author (raghav@iiserpune.ac.in).

Song analysis

Audio files were segmented into syllables based on a user-defined amplitude threshold. Syllables with less than 5 ms between them were merged and syllables with a duration shorter than 10 ms were discarded. Individual syllables were given labels in a semi-automatic manner. They were first assigned labels based on a

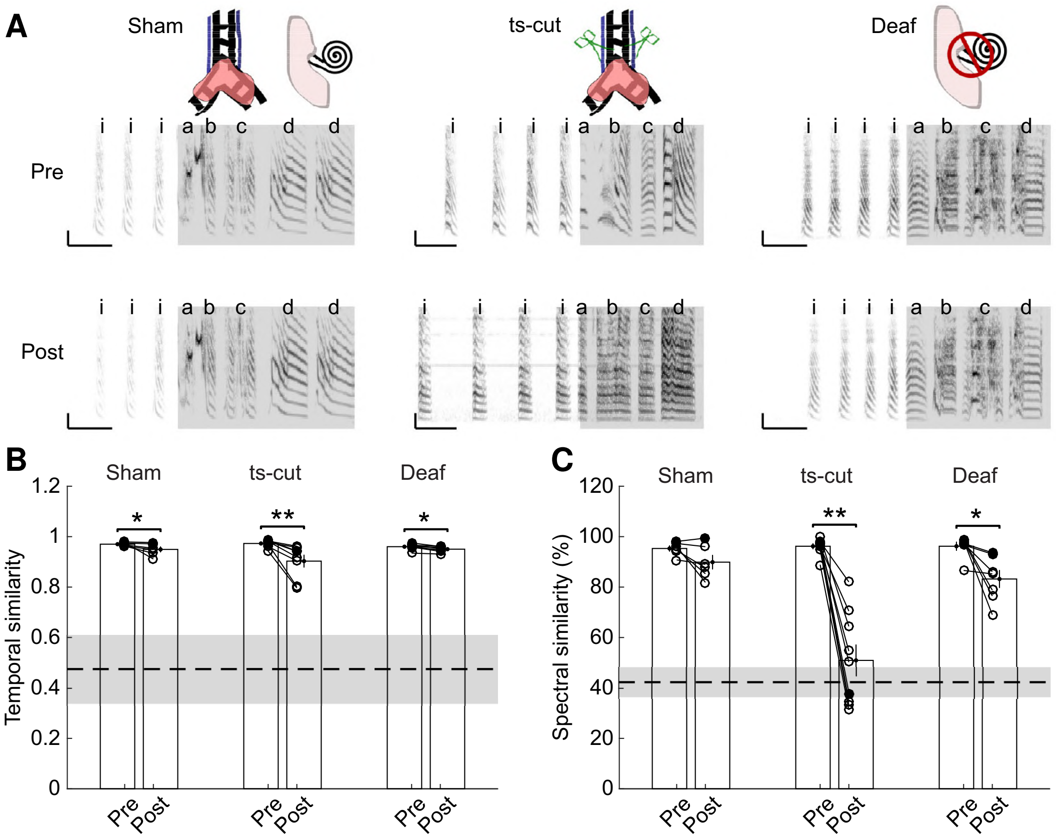

Fig. 2. Song structure after sham surgery, tracheosyringeal nerve cut or deafening. (A) Representative spectrograms of song for individual birds before and after sham surgery (left), tracheosyringeal (ts) nerve surgery (middle) or deafening (right). Scale bar: 200 ms and 1000 Hz. 'i' denotes INs and 'a' to 'd' represent song motif syllables. (B,C) Temporal similarity (B) and spectral similarity (C) to pre-surgery song for songs produced before and after sham surgery (*n*=6 birds), ts nerve surgery (*n*=8 birds) or deafening (*n*=7 birds). Each circle represents one bird and lines connect data from the same bird before and after surgery. Bars and whiskers represent means±s.e.m. across birds. Dashed lines and shading represent means and 95% confidence intervals for similarity between random birds. Filled circles represent values for the birds shown in A. *$P<0.05$, **$P<0.01$, Wilcoxon signed-rank test.

modified template-matching procedure (Glaze and Troyer, 2006) or clustering based on acoustic features calculated using Sound Analysis Pro. Clustering was done using KlustaKwik (http://klustakwik.sourceforge.net/). Labels were then manually checked for all files.

The repetitive sequence or song motif for each bird was identified. Song bouts were defined as groups of vocalizations with at least one motif syllable that were separated from other such groups by more than 2 s of silence (Sossinka and Böhner, 1980). For a subset of birds (*n*=7 deaf birds; *n*=6 birds for analysis of call-song bouts and *n*=7 birds for analysis of day-to-day changes in IN number and properties) with triggered recordings, a number of files did not have 2 s of silence before the first vocalization in the file. However, as these were triggered recordings, we assumed that there was silence before the start of the file too and so we considered files with >0.5 s silence at the beginning of the file as valid bouts. For a given bird, we used the same criterion before and after surgery to ensure that the criterion did not affect our results. Syllables that were produced in isolation outside of song bouts were identified as calls. All kinds of calls (distance calls, short calls and intermediate calls) (Zann, 1996) were combined together.

As described earlier (Price, 1979; Rajan and Doupe, 2013; Sossinka and Böhner, 1980), syllables that were repeated at the beginning of a bout were considered as INs. Calls were not considered as INs. As described previously (Zann, 1993), 76.2% of our birds (*n*=32/42) produced only one IN type. The rest of the birds produced two IN types (*n*=10/42). For all the analyses described, we combined the multiple types of INs together.

For ts-cut birds, syllables and INs lost their characteristic acoustic structure and were reduced to harmonic stacks without any modulation (Fig. 2A, middle). However, durations of individual syllables and INs remained the same as pre-surgery (Fig. 2A,B, middle, Fig. 3; Fig. S3). In these birds, syllables were labeled using cluster analysis as described above and INs and motif syllables were matched to pre-surgery INs and motif syllables by examining plots for duration versus mean frequency for all syllables.

On average, we analyzed 124 song bouts per bird (median 98 song bouts per bird; range 11–428 song bouts per bird).

Temporal and spectral similarity

We quantified changes in song after removal of sensory feedback using temporal and spectral similarity. Temporal similarity was calculated as the maximum of the cross-correlation function between the normalized amplitude envelopes of a pre-surgery template motif and other pre-/post-surgery motifs (*n*=9 randomly chosen motifs from pre-surgery and *n*=10 randomly chosen motifs from post-surgery) (Roy and Mooney, 2007). The template motif was proportionally stretched ±20% to account for differences in duration of the entire motif. As a measure of random temporal similarity between any two zebra finches, we

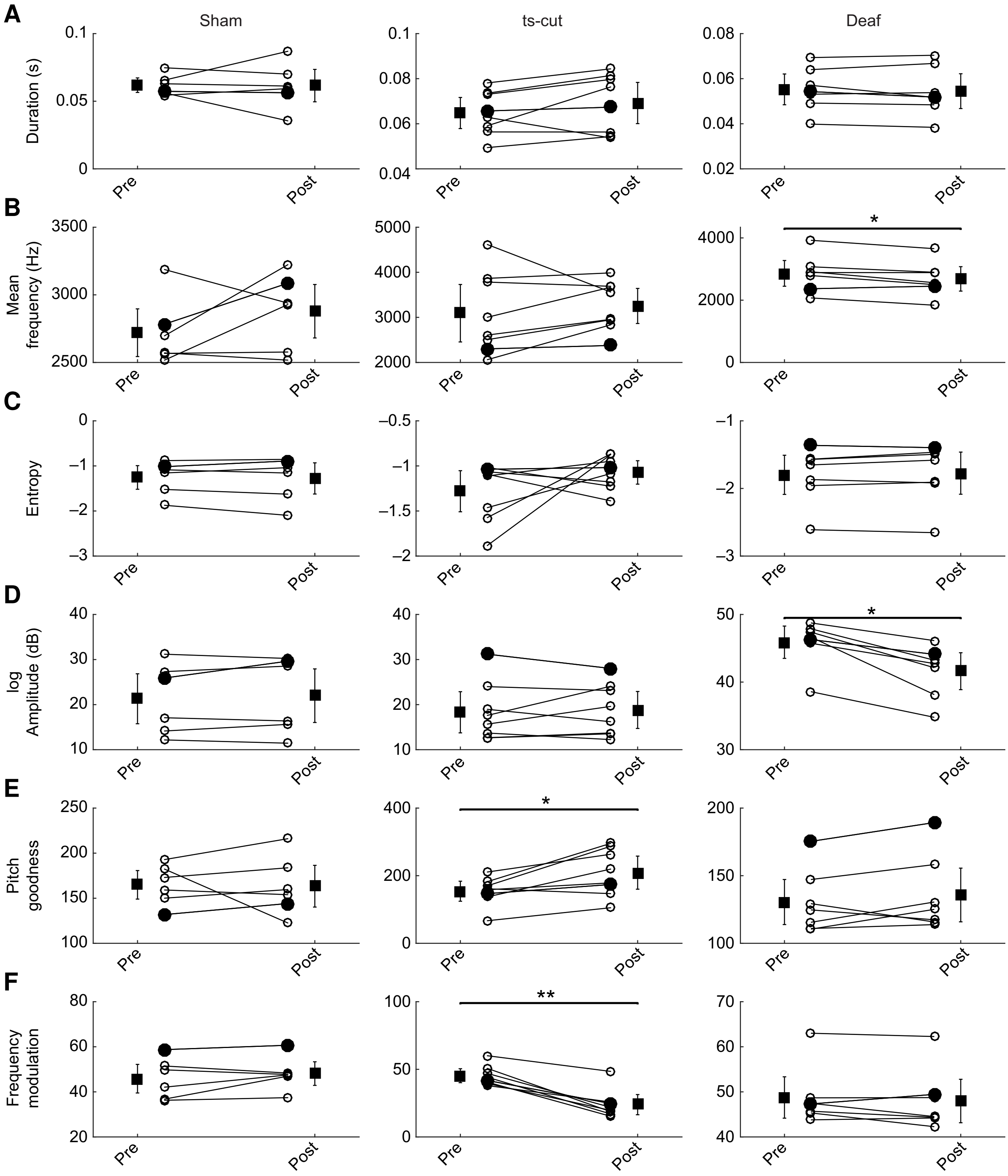

Fig. 3. Changes in IN acoustic features after removal of sensory feedback. (A–F) Acoustic properties of INs before and after sham surgery (left column, *n*=6 birds), ts nerve surgery (middle column, *n*=8 birds) or deafening (right column, *n*=7 birds). Circles represent individual birds and lines connect data from the same bird. Squares and whiskers represent means±s.e.m. across birds. Acoustic features plotted are duration (A), mean frequency (B), entropy (C), log amplitude (D), goodness of pitch (E) and frequency modulation (F). The filled circles in the columns from left to right represent the values for the birds shown in Fig. 2A–C, respectively. **P*<0.05, ***P*<0.01, Wilcoxon signed-rank test.

calculated temporal similarity for motifs from 10 random pairs of birds (*n*=10 motifs each). Spectral similarity (% similarity) was calculated using Sound Analysis Pro (five motifs pre-surgery were compared with five other motifs pre-/post-surgery) (Tchernichovski et al., 2000). Random spectral similarity was measured for 10 random pairs of birds (*n*=5 motifs each).

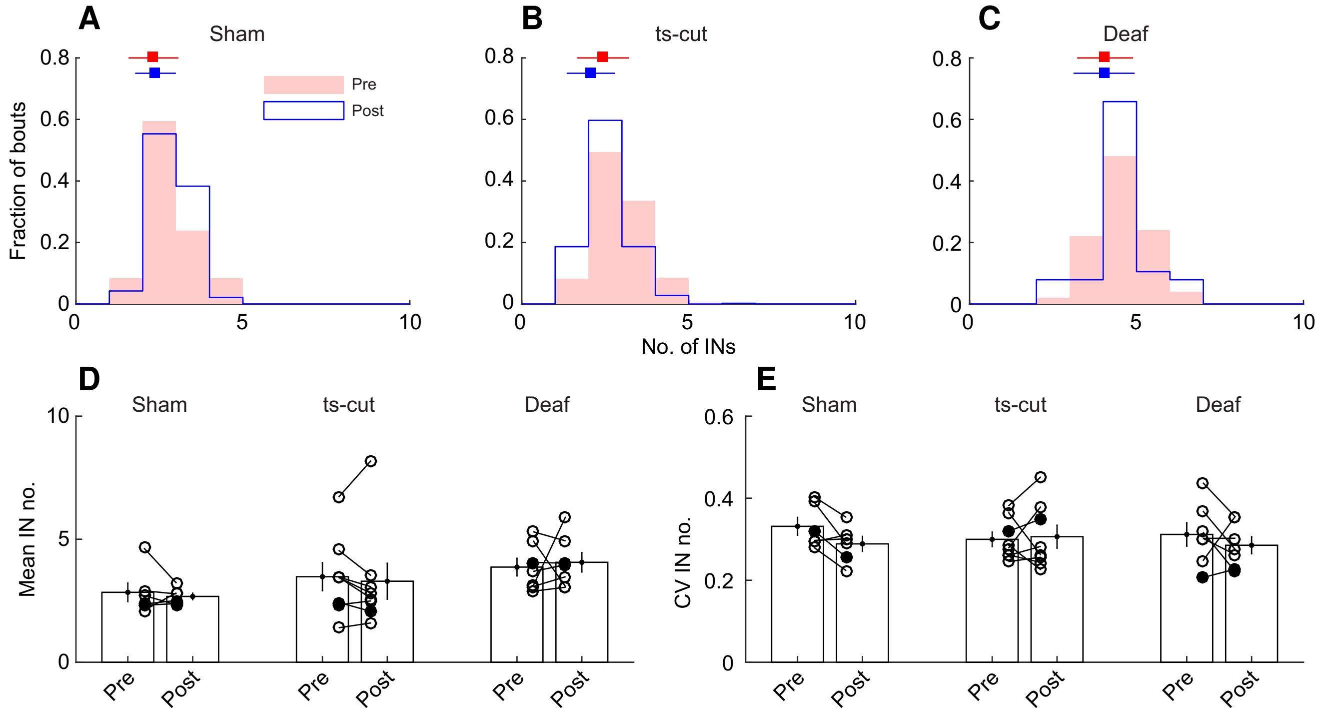

Fig. 4. Mean IN number and variability are not affected by removal of sensory feedback. (A–C) Distribution of number of INs for a representative bird before (red shaded) and after (blue outline) sham surgery (A), ts nerve surgery (B) and deafening (C). Squares and whiskers represent means and s.d. of the distributions before (red) and after surgery (blue). (D,E) Mean IN number (D) and coefficient of variation (CV) of IN number (E) before and after surgery for all sham surgery (*n*=6 birds), ts-cut (*n*=8 birds) and deaf birds (*n*=7 birds). Circles represent individual birds and lines connect data from the same bird. Bars and whiskers represent means±s.e.m. across birds. Filled circles represent the birds shown in A–C. *P*>0.05, Wilcoxon signed-rank test between pre- and post-surgery.

Characterization of IN progression

In each song bout, the last set of consecutive INs with inter-IN intervals <500 ms before the first motif syllable were considered for counting IN number (Rajan and Doupe, 2013; Sossinka and Böhner, 1980). All of our analysis was restricted to such sequences of INs present at the beginning of each bout.

Intervals between INs were measured as the duration between the end of an IN to the start of the next IN. The first interval was the interval between the first two INs satisfying the above criteria. The last interval was measured as the interval between the last IN and the first motif syllable. As a measure of the progression of IN timing, we quantified the ratio between successive IN intervals across all IN sequences. Ratios were averaged across bouts to obtain a mean ratio for each bird. A ratio <1 indicated a speeding up of successive intervals as shown earlier (Rajan and Doupe, 2013). The coefficient of variation (CV) was measured as the standard deviation divided by the mean.

To characterize acoustic properties of INs and their progression to song, we used the acoustic distance to the last IN and the ratio of the distance of successive INs, respectively. The acoustic distance is an inverse measure of similarity in acoustic properties between an IN and all last INs (Rajan and Doupe, 2013). We calculated four acoustic features, namely duration, log amplitude, entropy and mean frequency for each IN using the MATLAB code for Sound Analysis Pro (http://soundanalysispro.com/matlab-library). For each day, we randomly chose 50% of the last INs as the reference. The distance of the remaining last INs and the corresponding first INs in the same bouts was measured as the Mahalanobis distance of the IN from the reference last INs in the 4-dimensional space formed by the four acoustic features. As a measure of acoustic progression of INs, we calculated the ratio of distances of successive INs from the last IN for each IN sequence at the beginning of a bout (50% of the bouts were excluded as the last INs from these bouts were chosen as the reference). A ratio <1 indicated that successive INs became closer in distance (or more similar) to the last IN, as seen in intact birds (Rajan and Doupe, 2013).

Analysis of the influence of calls on the number and properties of INs

The influence of calls on IN number and properties was analyzed in 16 normal, unmanipulated birds. Bouts where the first IN began <2000 ms after the end of a call were considered as call-song bouts. Bouts with only INs at the beginning were considered as IN song bouts. Birds with a minimum of five IN song bouts and five call-song bouts were considered for this analysis. For each bird, the mean number of INs in IN song bouts was subtracted from the number of INs in each call-song bout. For each bird, the change in IN number in call-song bouts was then binned at 100 ms resolution starting at 40 ms after the end of the call to 1940 ms after the end of the call. Across all birds, we fitted an exponential function (MATLAB fit function) to characterize the dependence of this change in IN number on time between the end of the call and the start of the first IN. Similarly, we also fitted exponential functions to the change in the interval between the first two INs and change in acoustic properties of the first IN (Fig. S4).

For many of the feedback-deprived birds, we did not have enough call-song bouts to carry out a similar analysis. Instead, we divided call-song bouts into two categories: (1) bouts where the first IN started <200 ms after the end of the call and (2) bouts where the first IN started >200 ms after the end of the call. A 200 ms period was chosen based on the exponential fit (Fig. 7B) and data availability in the feedback-deprived birds. We calculated mean IN number, mean and variability of the interval between the first two INs and acoustic distance of the first IN for both these bout categories and compared

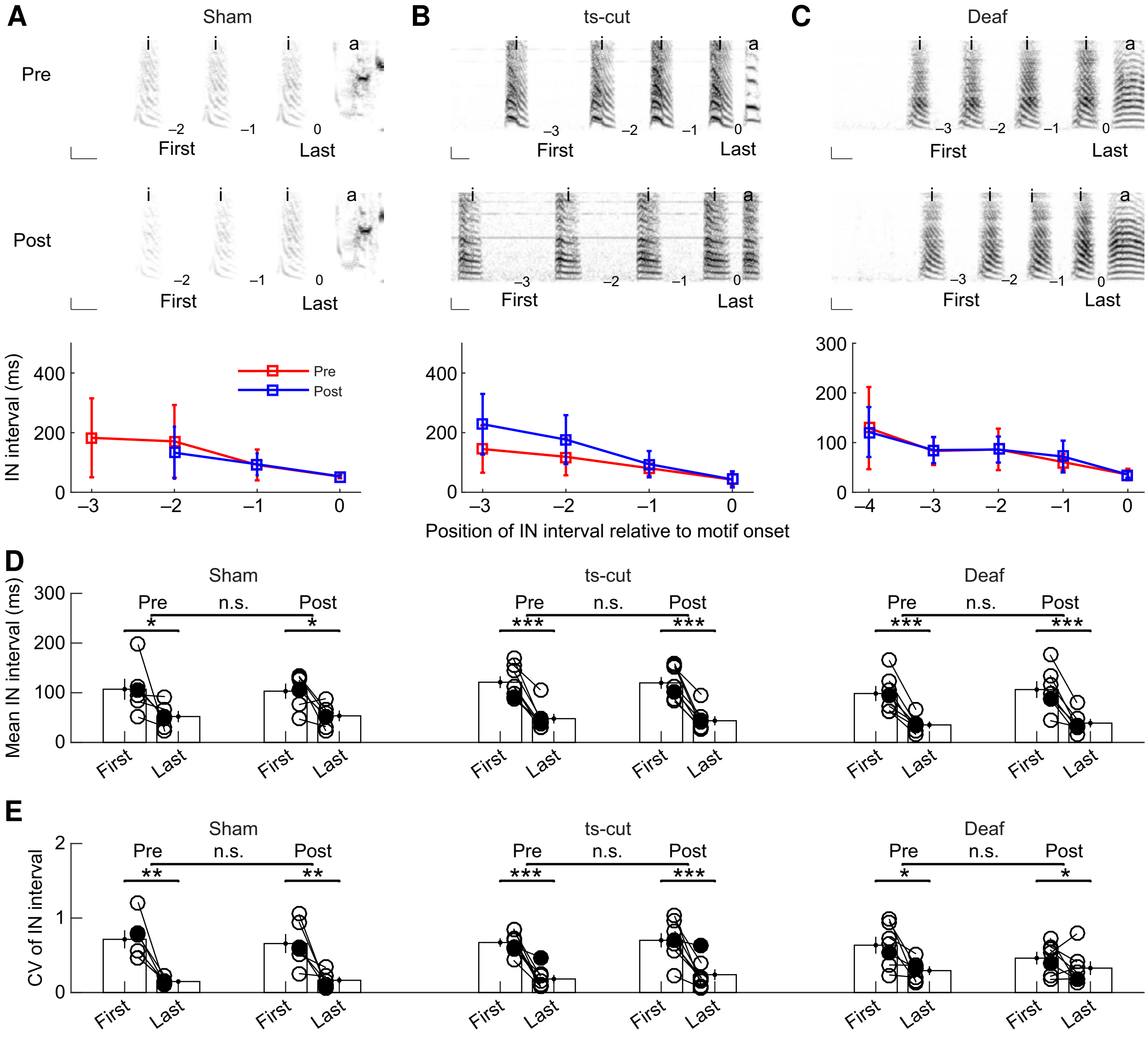

Fig. 5. Progression of intervals between successive INs is independent of sensory feedback. (A–C) Top: spectrograms of a sequence of INs before the first motif syllable before and after surgery. The position of each IN interval relative to the motif onset and the first and last interval is marked. Bottom: interval between successive INs versus position of the interval relative to motif onset for three example birds before (red) and after (blue) sham surgery (A), ts nerve surgery (B) or deafening (C). Squares and whiskers represent mean and s.d., *n*>5 bouts. Scale bars for A–C: 50 ms (horizontal) and 1000 Hz (vertical). (D) Mean interval between the first two INs in a bout and mean interval between the last IN and song across all birds before and after sham surgery (left, *n*=6 birds), ts nerve surgery (middle, *n*=8 birds) or deafening (right, *n*=7 birds). Circles represent individual birds and lines connect data from the same bird before and after surgery. Bars and whiskers represent means±s.e.m. across all birds. Filled circles represent the birds shown in A–C. **P*<0.05, ***P*<0.01, ****P*<0.001, repeated-measures two-way ANOVA. (E) CV of the interval between the first two INs in a bout and the interval between the last IN and song across all birds before and after sham surgery (left, *n*=6 birds), ts nerve surgery (middle, *n*=8 birds) or deafening (right, *n*=7 birds). Circles represent individual birds and lines connect data from the same bird before and after surgery. Bars and whiskers represent means±s.e.m. across all birds. Filled circles represent the birds shown in A–C. **P*<0.05, ***P*<0.01, ****P*<0.001, n.s. represents not significant (*P*>0.05), repeated-measures two-way ANOVA.

them with the corresponding properties for IN song bouts (Figs 7 and 8). Only birds with >3 call-song bouts in both of these categories were considered for analysis. Further, we combined data for ts-cut and deaf birds as our previous results showed that both manipulations had no effect on IN number and properties.

Statistics

We did not perform any power calculations to determine sample sizes for each group. However, sample sizes are comparable with other studies. As detailed below and in Table S1, we used non-parametric tests for most of our statistical comparisons because of small sample sizes (<10). We used a parametric test – repeated-measures one-way ANOVA and repeated-measures two-way ANOVA – only for analysis of the effect of calls on the number and properties of INs and the analysis of progression of IN features after feedback removal, respectively. Birds were excluded from the analysis only if there were no pre-surgery undirected song recordings or if there were too few song bouts (number specified in earlier sections). Both of these conditions were established before the start of the analysis. The investigators were not blinded for both the choice of animals in each group and the analysis of data. However, IN number or properties were not considered while choosing birds as described in an earlier section.

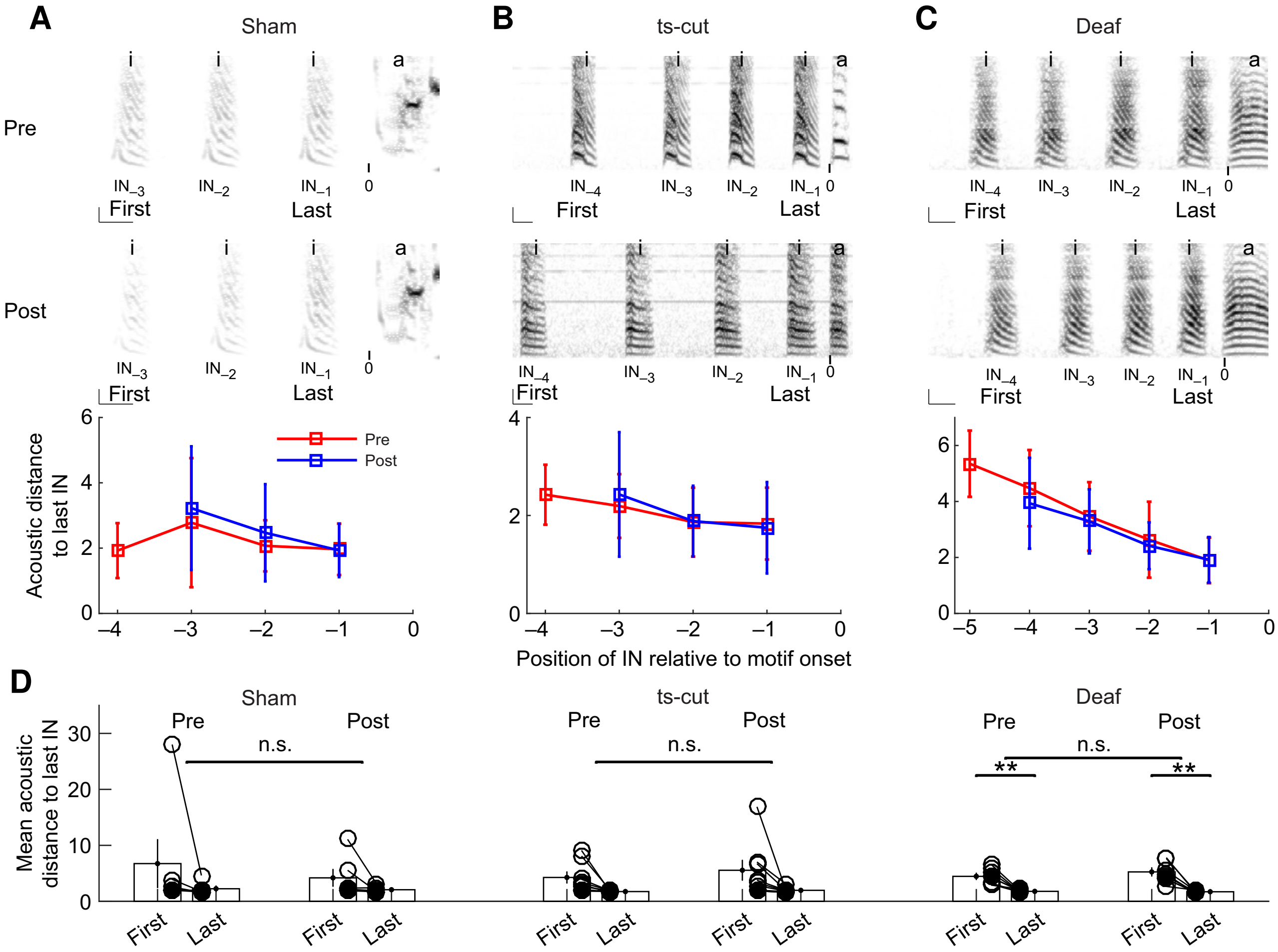

Fig. 6. Progression of acoustic properties of successive INs is independent of sensory feedback. (A–C) Top: spectrograms of a sequence of INs before the first motif syllable before and after surgery. The position of each IN relative to motif onset, and the first and last INs are marked. Bottom: acoustic distance of successive INs from the last IN for three example birds before (red) and after (blue) sham surgery (A), ts nerve surgery (B) or deafening (C). Squares and whiskers represent means and s.d., $n>5$ bouts. Scale bars for A–C: 50 ms (horizontal) and 1000 Hz (vertical). (D) Mean acoustic distance for the first IN and last IN in a bout across all birds before and after sham surgery (left, $n=6$ birds), ts nerve surgery (middle, $n=8$ birds) or deafening (right, $n=7$ birds). Circles represent individual birds and lines connect data from the same bird before and after surgery. Bars and whiskers represent means and s.e.m. across all birds. Filled circles represent the birds shown in A–C. $^{**}P<0.01$, n.s. represents not significant ($P>0.05$), repeated-measures two-way ANOVA.

Wilcoxon signed-rank test was used for paired comparisons of temporal similarity (Fig. 2B), spectral similarity (Fig. 2C), changes in IN/motif syllable acoustic features (Fig. 3; Fig. S1), mean IN number (Fig. 4D), IN number CV (Fig. 4E) and progression in IN features (Fig. S3). For comparing progression in IN timing and IN acoustic structure after removal of feedback (Figs 5D,E and 6D), we used repeated-measures two-way ANOVA using IN position (first versus last) as one factor and time (pre-surgery versus post-surgery) as the second factor (MATLAB code from https://in.mathworks.com/matlabcentral/fileexchange/6874-two-way-repeated-measures-anova). For comparing IN number and properties in bouts where calls preceded the first IN, we used repeated-measures one-way ANOVA (Figs 7C,D and 8). If the ANOVA *P*-value was <0.05, we used a *post hoc* Tukey–Kramer test to identify groups that were significantly different (Figs 7C,D and 8). For comparing changes in IN number and properties after surgery with day-to-day changes, we used Kruskal–Wallis ANOVA (Fig. S2). Pearson's correlation coefficient was used to assess the correlation between first IN properties and time to song initiation (Fig. 9).

All of the tests used and the associated *P*-values are provided in Table S1. A significance level of $P=0.05$ was used throughout.

RESULTS

To see whether sensory feedback plays a role in song initiation, we analyzed the number and progression of INs after removal of either proprioceptive ($n=8$ birds) or auditory feedback ($n=7$ birds). As a control, sham surgeries were performed in a separate group of birds ($n=6$). As we were interested in self-initiated movement sequences, we focused on undirected songs produced when the bird was alone.

Song spectral structure alone is affected after ts nerve surgery

Proprioceptive feedback was removed by bilaterally cutting the ts nerve (Bottjer and Arnold, 1984; Vicario, 1991; Williams and McKibben, 1992) (ts-cut; $n=8$ birds, see Materials and Methods) and auditory feedback was eliminated by bilateral removal of the cochlea (Konishi, 1965) (deaf; $n=7$ birds, see Materials and Methods).

Despite small changes, songs of birds subjected to sham surgery remained more similar to pre-surgery songs than expected by chance in both spectral and temporal structure (Fig. 2). As described earlier (Konishi, 1965; Price, 1979), only minor changes to song characteristics (both temporal and spectral) were seen after deafening. Songs post-surgery remained more similar than expected

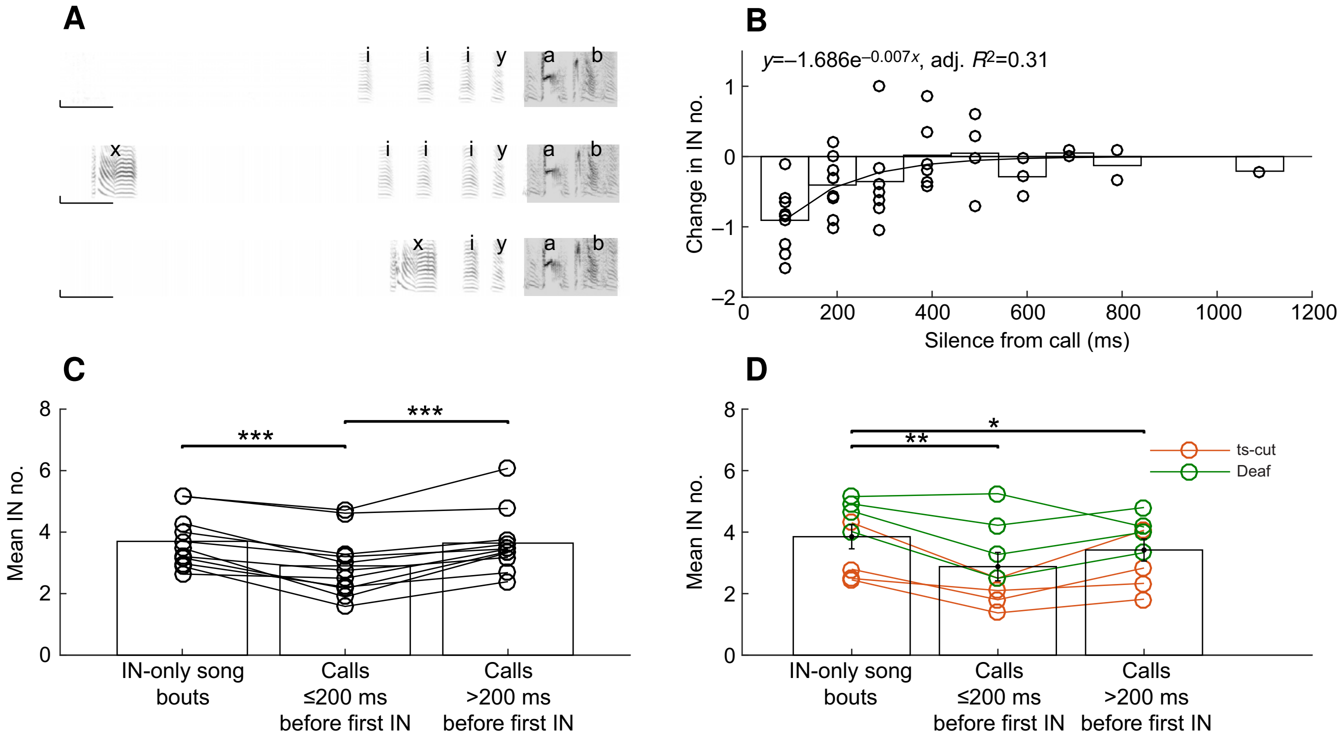

Fig. 7. Calls just before the first IN correlate with fewer INs at the start of the song bout. (A) Spectrograms of the start of an IN song bout (top) and two call song bouts with the call occurring well before the first IN (middle) or just before the first IN (bottom). ‘i’ and ‘y’ represent INs, ‘a’ and ‘b’ represent motif syllables and ‘x’ represents a call. Scale bars: 200 ms and 1000 Hz. (B) Silence between the end of the call and the beginning of the first IN versus change in IN number relative to the mean IN number in IN song bouts (n=16 birds). Each circle represents one bird. Bars represent means across birds and the line represents an exponential fit to the data ($y=-1.686e^{-0.007x}$, adjusted R^2=0.31). (C,D) Mean IN number in IN song bouts, call bouts with calls occurring ≤200 ms before the first IN and call bouts with calls occurring >200 ms before the first IN for unmanipulated, feedback-intact birds (n=11 birds) (C), and for feedback-deprived birds (D; orange – post-ts-cut, n=4 birds; green – post-deafening, n=4 birds). Circles represent individual birds and lines connect data from the same bird. Bars and whiskers represent means±s.e.m. across birds. *P<0.05, **P<0.01, ***P<0.001, repeated-measures one-way ANOVA.

by chance to pre-surgery songs (Fig. 2). The ts nerve contains both efferent and afferent nerves carrying motor input to the syringeal muscles and proprioceptive feedback from the syringeal muscles, respectively (Bottjer and Arnold, 1984). Nerve cuts disrupted both efferent and afferent nerves, resulting in the loss of song spectral structure immediately after nerve cut surgery (Fig. 2A,C) (Roy and Mooney, 2007; Vicario, 1991; Williams and McKibben, 1992). However, as described earlier, song temporal structure remained more similar than expected by chance to that before ts nerve surgery (Fig. 2A middle and Fig. 2B) as motor input to the respiratory muscles was not affected (Bottjer and Arnold, 1984; Roy and Mooney, 2007; Vallentin and Long, 2015; Vicario, 1991; Williams and McKibben, 1992). Thus, consistent with earlier studies (Bottjer and Arnold, 1984; Roy and Mooney, 2007; Vallentin and Long, 2015; Vicario, 1991; Williams and McKibben, 1992), we also found that cutting the ts nerve altered both proprioceptive and auditory feedback, while deafening disrupted only auditory feedback.

IN acoustic structure, not duration, is affected by ts-cut surgery

We next quantified changes to the acoustic structure of INs after removal of either proprioceptive or auditory feedback. Similar to changes in song syllable structure (Fig. 2A, middle column), INs also became harmonic stacks after surgery in ts-cut birds as seen by increased goodness of pitch and decreased frequency modulation (Fig. 2A middle, Fig. 3E,F, P<0.05, Wilcoxon signed-rank test). Despite this change, we could identify INs because IN duration, mean frequency, entropy and amplitude did not change significantly after surgery (Fig. 3A–D ts-cut; see Materials and Methods for details of IN identification procedure in ts-cut birds). The position of INs at the beginning of the bout was also maintained (Fig. 2A middle). Post-deafening, INs were softer and had reduced mean frequency (Fig. 3B, D, P<0.05, Wilcoxon signed-rank test). However, song syllables were also softer after deafening (Fig. S1, right column, P<0.05, Wilcoxon signed-rank test), suggesting that these changes could have been the result of a change in microphone position after surgery. No significant changes in IN acoustic structure were seen after sham surgery (Fig. 3, P>0.05, Wilcoxon signed-rank test).

Mean IN number before song is not affected by removal of proprioceptive or auditory feedback

We next analyzed the mean and variability of IN number before each song (Fig. 4A–C; see Materials and Methods). The mean number of INs before song (Fig. 4D) and the variability in IN number (measured by the CV – Fig. 4E) did not change significantly soon after surgery in sham-surgery, ts-cut and deaf birds (Fig. 4D and E, P>0.05, Wilcoxon signed-rank test). In fact, changes in mean IN number post-surgery for feedback-deprived birds were not different from day-to-day changes in IN number seen in normal, unmanipulated birds (Fig. S2A). This further strengthened our conclusion that IN number was unaffected by removal of proprioceptive or auditory feedback.

Progression of IN timing to song is not affected by removal of proprioceptive or auditory feedback

We have previously shown that progression of INs to song is accompanied by changes in both the timing and acoustic structure of

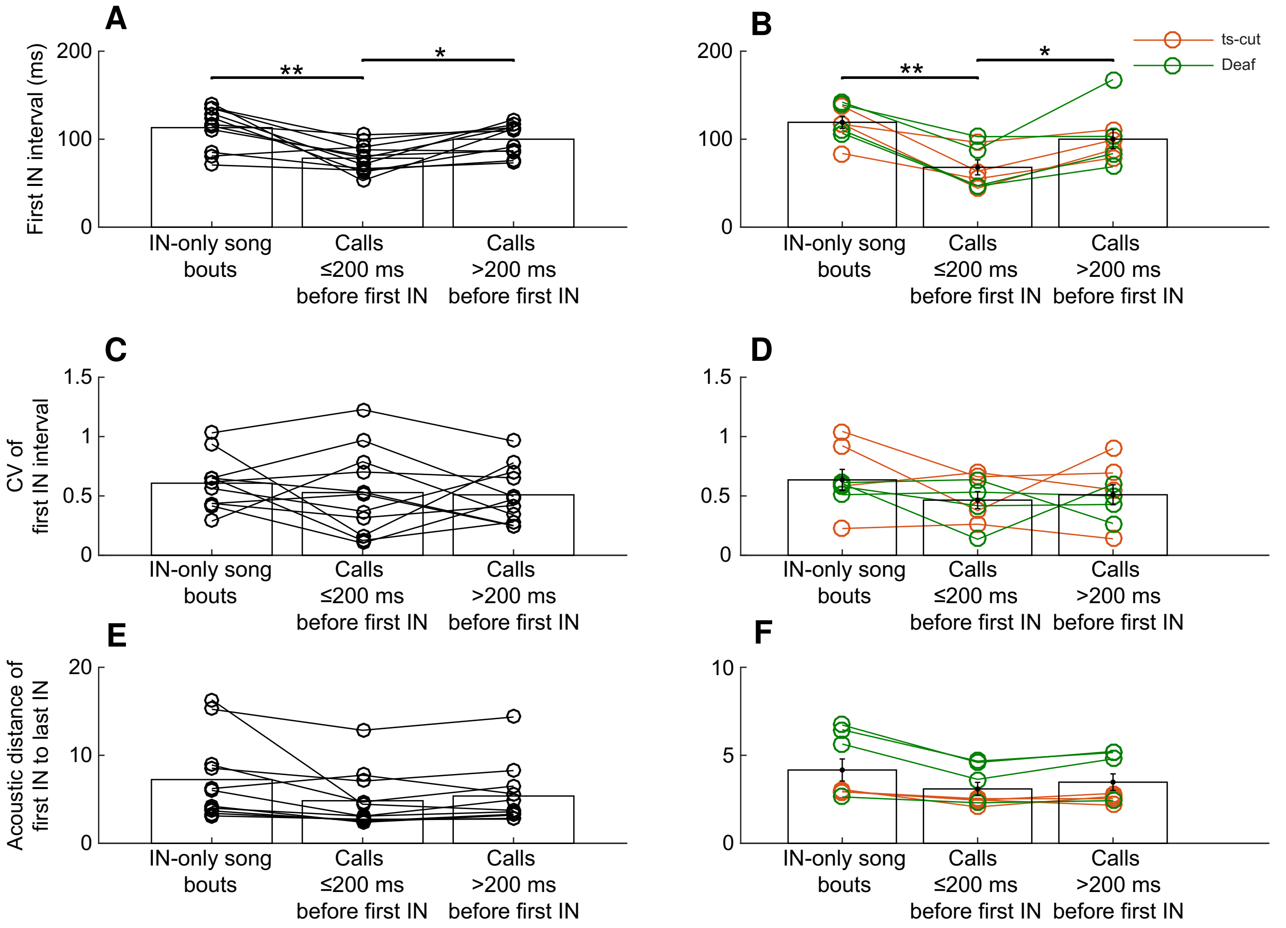

Fig. 8. Calls just before the first IN correlate with a shorter interval between the first two INs. (A,B) Mean duration of the interval between the first two INs in IN song bouts, call bouts with calls occurring ≤200 ms before the first IN and call bouts with calls occurring >200 ms before the first IN for unmanipulated, feedback-intact birds (A; *n*=11 birds) and for feedback-deprived birds (B; orange: post-ts-cut, *n*=4 birds; green: post-deafening, *n*=4 birds). Circles represent individual birds and lines connect data from the same bird. Bars and whiskers represent means±s.e.m. across birds. (C,D) Variability of the interval between the first two INs in IN song bouts, call bouts with calls occurring ≤200 ms before the first IN and call bouts with calls occurring >200 ms before the first IN for unmanipulated, feedback-intact birds (*n*=11 birds) (C) and for feedback-deprived birds (D; orange: post ts-cut, *n*=4 birds; green: post-deafening, *n*=4 birds). Circles represent individual birds and lines connect data from the same bird. Bars and whiskers represent means±s.e.m. across birds. (E,F) Mean acoustic distance of the first IN from the last IN in IN song bouts, call bouts with calls occurring ≤200 ms before the first IN and call bouts with calls occurring >200 ms before the first IN for unmanipulated, feedback-intact birds (*n*=11 birds) (E) and for feedback-deprived birds (F; orange: post-ts-cut, *n*=4 birds; green: post-deafening, *n*=4 birds). Circles represent individual birds and lines connect data from the same bird. Bars and whiskers represent means±s.e.m. across birds. **P*<0.05, ***P*<0.01, repeated-measures one-way ANOVA.

INs within a bout (Rajan and Doupe, 2013). We first considered IN timing. Specifically, intervals between successive INs progress from a longer, more variable first interval to a shorter, more stereotyped interval between the last IN and song (Rajan and Doupe, 2013). This progression in IN timing was unchanged after surgery in sham-surgery, ts-cut and deaf birds (Fig. 5A–C). After surgery, the interval between the first two INs remained longer and more variable than the interval between the last IN and song in ts-cut, deaf and sham-surgery birds (Fig. 5D,E; *P*<0.05 for first versus last, repeated-measures two-way ANOVA). Importantly, removal of

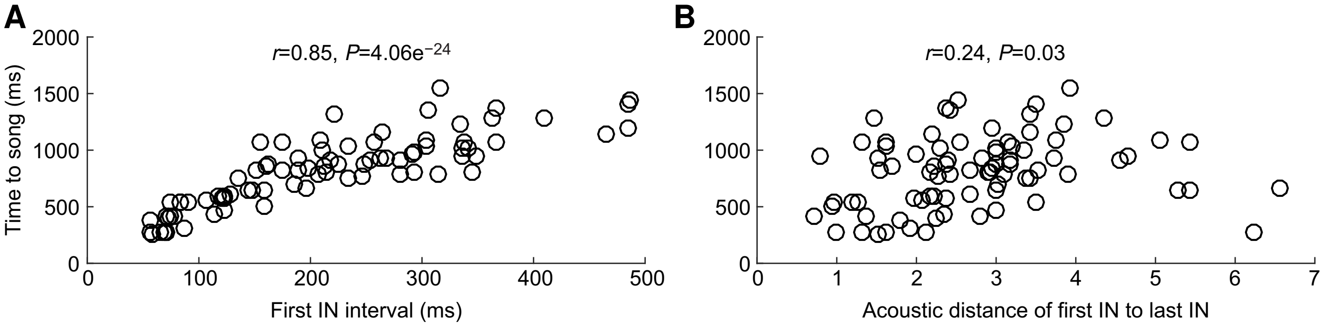

Fig. 9. 'Initial' state of IN progression correlates with song initiation delay. (A,B) Correlation between the interval between the first two INs (A) or the acoustic distance of the first IN to the last IN (B) and the time to the start of song (onset of the first motif syllable) for one bird. Circles represent data from individual bouts.

feedback did not alter either the mean or variability of both the first interval and the last interval (Fig. 5D,E; $P>0.05$ for pre- versus post-surgery, repeated-measures two-way ANOVA). A number of other aspects of IN timing were also not affected by removal of auditory or proprioceptive feedback and changes in IN timing post-surgery were similar to day-to-day changes seen in unmanipulated birds (Figs S2B–D and S3A – see Materials and Methods). These results showed that the timing of INs and their progression did not depend on intact sensory feedback.

Progression of IN acoustic features to song is not affected by removal of proprioceptive or auditory feedback

Similar to IN timing, the acoustic structure of INs has also been shown to progress to a consistent last-IN state just before song (Rajan and Doupe, 2013). Although individual INs in each bout looked very similar (Fig. 6A–C top), we have previously shown that the first IN is less similar to the last IN across bouts. We quantified this by calculating the similarity between the first IN and the last IN before and after surgery (acoustic distance to the last IN: the smaller the distance, the greater the similarity and vice versa; see Materials and Methods; see Fig. 6A–C for representative examples for sham-surgery, ts-cut and deaf birds). As we were interested in the progression, we calculated similarity to the last IN on the same day (pre-surgery last IN for pre-surgery and post-surgery last IN for post-surgery; see Materials and Methods). For each day, half of the last INs across all bouts were randomly chosen as a reference. The rest of the last INs and all of the first INs were then compared with this reference using the acoustic distance as an inverse measure of similarity (see Materials and Methods). The first IN was significantly different from the last IN (larger distance – Fig. 6D) before and after surgery in deaf birds ($P<0.05$ for first versus last IN, repeated-measures two-way ANOVA) and almost reached significance in ts-cut birds ($P=0.0544$ for first versus last IN, repeated-measures two-way ANOVA). However, this difference was smaller in sham-surgery birds both before and after surgery and did not reach significance (Fig. 6D, $P=0.2469$ for first versus last IN in sham-surgery birds, repeated-measures two-way ANOVA). Importantly, in all groups of birds, removal of feedback did not affect any of the measures of progression ($P>0.05$, pre- versus post-surgery, repeated-measures two-way ANOVA). These results showed that INs still progressed from a first IN that was significantly different from the last IN to a more consistent last IN even in the absence of auditory or proprioceptive feedback. A number of other aspects of IN acoustic structure progression to song were also not affected by removal of auditory or proprioceptive feedback and remained similar to day-to-day changes seen in unmanipulated birds (Figs S2E–G and S3B – see Materials and Methods). As mentioned earlier, ts-cut birds lacked neural input to the syringeal muscles in addition to the loss of proprioceptive feedback from the syringeal muscles. The continued progression of IN acoustic features suggested that this progression is a result of changing respiratory drive, as neural input to the respiratory muscles remained intact in these birds.

Overall, these results show that IN number and progression are not dependent on intact sensory feedback (auditory and proprioceptive). This suggested that IN progression to song is controlled by internal neural processes.

IN number is reduced when calls precede the first IN of a song bout

If IN progression is controlled by internal neural processes, we next asked whether the presence of calls (other non-song vocalizations) just before the first IN influenced the number of INs before song. Calls are partially learned or unlearned vocalizations that are acoustically distinct from song and are initiated by separate neural pathways (Simpson and Vicario, 1990; Vicario, 2004; Zann, 1996). Many aspects of calls are controlled by song motor nuclei and increased neural activity is seen in many of the song motor nuclei before and during calls (Benichov et al., 2016; Danish et al., 2017; Hahnloser et al., 2002; Kozhevnikov and Fee, 2007; Long and Fee, 2008; Rajan, 2018; Simpson and Vicario, 1990; Vyssotski et al., 2016; Yu and Margoliash, 1996). Further, we have previously shown the presence of higher levels of preparatory activity in the premotor nucleus HVC before the first IN when calls precede the first IN of a song bout (Rajan, 2018). Given these changes in neural activity when calls are present before the first IN, we expected changes in IN number. We tested this in a separate set of unmanipulated birds ($n=16$) by examining the number of INs in song bouts where calls (other non-song vocalizations) preceded the first IN (call-song bouts – see Materials and Methods).

Calls occurred at variable times before the first IN in a small fraction of bouts (Fig. 7A; mean±s.e.m. interval between end of call and start of first IN: 468.6±48 ms, mean±s.e.m. CV of interval between end of call and start of first IN: 0.85±0.07; $n=16$ birds). We observed fewer INs when calls occurred before the first IN ($n=16$ birds; mean±s.e.m. for IN song bouts: 3.7±0.24, for call song bouts: 3.4±0.25, $P=0.03$, Wilcoxon signed-rank test). This reduction was dependent on the time between the end of the call and the start of the IN: the shorter the time, the greater the reduction (Fig. 7B, adjusted $R^2=0.31$ for an exponential fit, see Materials and Methods). In both feedback-intact and feedback-deprived birds, song bouts where the first IN began <200 ms after the end of a call had fewer INs when compared with song bouts with only INs at the beginning or song bouts where the first IN began >200 ms after the end of the call (Fig. 7C,D, $P<0.05$, repeated-measures one-way ANOVA and *post hoc* Tukey–Kramer test). These results showed that the presence of calls just before the first IN of a song bout correlated with fewer INs in both feedback-intact and feedback-deprived birds and further strengthened our conclusion that IN progression to song may be controlled by internal neural processes.

Calls just before the first IN of a song bout correlate with altered ‘initial’ state

Given that both IN timing and acoustic features progress towards a consistent ‘ready’ state just before song, we hypothesized that calls might reduce IN number by speeding up this progression. Consistent with this idea, song bouts where the first IN began <200 ms after the end of call had a significantly shorter interval between the first two INs when compared with song bouts with only INs or song bouts where the first IN began >200 ms after the end of the call (Fig. 8A,B; $P<0.05$, repeated-measures one-way ANOVA followed by *post hoc* Tukey–Kramer test). This was true both in feedback-intact (Fig. 8A) and feedback-deprived birds (Fig. 8B). In feedback-intact birds, the decrease in interval between the first two INs in bouts with calls was correlated with the time between the end of the call and the start of the first IN, though the strength of the correlation was moderate (Fig. S4A, adjusted R^2 for exponential fit=0.16). In contrast to the changes in IN timing, neither the variability of the interval between the first two INs nor the acoustic structure of the first IN showed any differences based on whether calls were present before the first IN or not (feedback-intact birds: Fig. 8C,E; feedback-deprived birds: Fig. 8D,F; $P>0.05$, repeated-measures one-way ANOVA). However, in feedback-intact birds, relative to bouts with only INs, the acoustic structure of the first IN after a call was more similar to that of the last IN (Fig. S4B). The change in acoustic structure was correlated with the time between the end of the call and the start of the first IN, but the strength of the correlation was weak (Fig. S4B, adjusted R^2 for exponential fit=0.09). Overall, these results showed that the presence of calls correlated with a change in IN timing (shorter interval between the first two INs), potentially causing the reduction in IN number before song.

‘Initial’ state of IN progression correlates with time to song initiation

Our results suggested that the progression of IN timing and acoustic features are controlled by internal neural processes possibly related to motor preparation. In other systems, neural preparatory activity is strongly correlated with the time to movement initiation: the greater the progress of preparation, the shorter the time to movement initiation (Churchland et al., 2006a; Shenoy et al., 2011, 2013). Similar to this, we found a significant correlation between the length of the interval between the first two INs and the time to song initiation in all birds (see example data from one bird in Fig. 9A; across all 16 birds, mean r=0.77, range=0.57–0.90). How similar the first IN was to the last IN was also correlated with the time to song initiation, albeit to a weaker extent in 14/16 birds (see example from one bird in Fig. 9B; significant in 14/16 birds, mean r=0.32, range=−0.39–0.62). These data suggested that IN timing and acoustic features reflect internal neural processes, possibly involved in preparing the zebra finch brain for song initiation.

DISCUSSION

In this study, we show that real-time auditory and/or proprioceptive feedback is not required for initiation of adult zebra finch song. We also show that the progression of INs, the repeated pre-song vocalizations, from a variable initial state to a more stereotyped final state is also independent of real-time sensory (auditory and/or proprioceptive) feedback. Further, we show, in both feedback-intact and feedback-deprived birds, that fewer INs are present when the first IN of a song bout occurs within 200 ms of the end of a call (other non-song vocalization). In such cases, IN timing was closer to the final state. Finally, the ‘initial’ state of IN progression was correlated with the time to song initiation. Overall, these results demonstrate that the progression of INs to song does not require real-time sensory feedback. Rather, progression of INs to song is controlled by internal neural processes possibly involved in preparing the motor system for song initiation.

Contributions of respiratory feedback to song initiation

One feedback that we did not alter is respiratory feedback from the air sacs (Méndez et al., 2010). However, previous work strongly suggests that respiratory feedback does not contribute to IN initiation. First, one earlier study showed that disrupting respiratory pressure during short syllables (of the order of 60 ms) did not disrupt song progression (Amador et al., 2013). Given that INs are short syllables of the order of 60 ms, INs may not require real-time respiratory feedback for progression to the next syllable (or song). Second, unilateral disruption of vagal feedback mostly affected syllables at the end of a song (Méndez et al., 2010). Finally, sparse, patterned neural activity of one class of neurons in the premotor nucleus HVC during singing was also not affected by removal of sensory feedback including respiratory feedback (Vallentin and Long, 2015). All of these data suggest that respiratory feedback does not play a role in IN progression.

Long-term requirement for sensory feedback

Song production in adult birds does not depend on real-time sensory feedback (Bottjer and Arnold, 1984; Konishi, 1965) and our results show that song initiation also does not depend on real-time sensory feedback. However, long-term song maintenance does require intact sensory feedback, as shown by song degradation starting many weeks after deafening (Horita et al., 2008; Nordeen and Nordeen, 1992; Williams and McKibben, 1992). Similarly, it is possible that sensory feedback could be necessary in the longer term for maintenance of IN progression to song (our study focused on songs produced within 10 days of removal of feedback). It would also be interesting to see whether song degradation seen at later time points after deafening is linked to (or caused by) a change in IN progression to song. If INs represent preparatory vocalizations, such a link would be expected as small changes in the neural preparatory state in primates are correlated with changes in features of the upcoming movement (Afshar et al., 2011; Churchland et al., 2006b).

Comparison of INs to motor preparation in other systems

Preparatory neural activity has been described as a slow change in neural activity, starting as early as 1 s before the start of a movement (Chen et al., 2017; Churchland et al., 2006b; Gao et al., 2018; Lee and Assad, 2003; Li et al., 2015; Maimon and Assad, 2006; Murakami et al., 2014; Romo and Schultz, 1987; Tanji and Evarts, 1976). One important characteristic of this preparatory activity appears to be a decrease in variability across trials (Churchland et al., 2006a,c). The decrease in variability as INs progress to song (Rajan and Doupe, 2013) is very similar to the decrease in variability in neural activity seen before the start of a movement. Together with our current data showing that sensory feedback is not important for progression of INs to song, these results suggest that INs may represent preparatory activity. Additionally, earlier studies have shown the presence of preparatory neural activity in song control areas well before the first IN of undirected song bouts (Hessler and Doupe, 1999; Kao et al., 2008; Rajan, 2018; Woolley et al., 2014) and directed song bouts (Daliparthi et al., 2018preprint). Thus, INs may reflect a continuation of this preparatory activity that begins hundreds of milliseconds before the first IN.

Overt movements in other systems as motor preparation

Our results suggest that overt vocalizations (INs) serve as preparatory activity. Previous studies describing neural preparatory activity in primates and rodents before the onset of a movement have not described similar overt movements as motor preparation (Chen et al., 2017; Churchland et al., 2006a; Gao et al., 2018; Murakami et al., 2014; Romo and Schultz, 1987; Tanji and Evarts, 1976). However, all of these studies have involved training animals to perform a task and animals are rewarded for maintaining stable posture without movements until a ‘go’ signal is provided for movement initiation. Therefore, overt preparatory movements, if present during the early stages of learning, would not be reinforced. This raises two interesting questions for further experiments. (1) Are overt movements present at early stages of learning in primates and rodents too? (2) Given that songbirds learn their song with internal reinforcement cues that only reinforce similarity to the tutor song (or tutor song memory) (Fee and Scharff, 2010), are INs learned similar to song learning? Additionally, there are human studies showing the presence of small eye movements (microsaccades) and small limb movements while waiting for a ‘go’ cue to perform an eye or limb movement (Betta and Turatto, 2006; Cohen and Rosenbaum, 2007; Corneil and Munoz, 2014). Changes in pupil size have also been shown to correlate with preparatory activity (Wang et al., 2015). This suggests that overt movements like INs may be more common before the start of naturally learned movements and may reflect motor preparation.

Mechanisms for IN progression to song

Our results show that sensory feedback is not essential for IN progression to song. Rather, the properties of INs correlate with the time to song initiation. How do the properties of INs change to progress to song? In our current study, we showed that the presence of calls prior to the first IN was correlated with shorter intervals between the first two INs and fewer INs before song. Similarly, shorter intervals between the first two INs have also been observed when neural preparatory activity in the premotor nucleus HVC precedes the first IN (Rajan, 2018). As calls are also associated with increased neural activity in many song control areas (Benichov et al., 2016; Danish et al., 2017; Hahnloser et al., 2002; Kozhevnikov and Fee, 2007; Vyssotski et al., 2016; Yu and Margoliash, 1996), the intervals between successive INs may reflect a history of increased activity within these inter-connected motor regions. The shorter interval might also lead to short-term plasticity that could facilitate song initiation by speeding up IN progression. Such short-term plasticity has been observed in the inputs to the premotor nucleus HVC (Coleman et al., 2007). Further experiments disrupting short-term plasticity or disrupting activity in motor control regions during IN production could help us to understand the mechanisms of IN progression to song.

Overall, our results show that real-time sensory feedback is not essential for INs to progress to song. Rather, changes in IN properties just before song initiation may reflect internal neural processes, potentially involved in preparing the zebra finch brain for initiation of the learned song sequence.

Acknowledgements
We would like to thank Prakash Raut for help with bird colony maintenance. We would also like to thank Michael Long, Hamish Mehaffey, Anand Krishnan, Deepa Subramanyam, Girish Deshpande, Sanjay Sane, Upi Bhalla and members of the Rajan and Krishnan Labs for useful discussions and comments on the manuscript. We also thank Allison Doupe, in whose lab the deafening experiments were carried out.

Competing interests
The authors declare no competing or financial interests.

Author contributions
Conceptualization: D.R., R.R.; Methodology: D.R., R.R.; Software: D.R., R.R.; Formal analysis: D.R., R.R.; Investigation: D.R., S.K.; Writing - original draft: D.R., R.R.; Writing - review & editing: D.R., S.K., R.R.; Visualization: D.R.; Supervision: R.R.; Project administration: R.R.; Funding acquisition: R.R.

Funding
This work was supported by a Department of Biotechnology, Ministry of Science and Technology (DBT) Ramalingaswami Fellowship (BT/HRD/35/02/2006) and a grant from the Department of Science and Technology, Ministry of Science and Technology (DST SERB EMR/2015/000829) and intramural support from the Indian Institute of Science Education and Research (IISER) Pune to R.R. We would also like to acknowledge a graduate student fellowship from IISER Pune and travel support from Department of Biotechnology, Ministry of Science and Technology-Conference, Travel, Exhibition and Popular Lectures (DBT/CTEP/02/2018 0847433) and the Infosys Foundation Travel Award (IISER-P/InfyFnd/Trv/116) to D.R. The deafening experiments, carried out at UCSF were supported by a National Institutes of Health award R01 (MH55987) to Allison Doupe. Deposited in PMC for release after 12 months.

Supplementary information
Supplementary information available online at
http://jeb.biologists.org/lookup/doi/10.1242/jeb.199042.supplemental

References

Afshar, A., Santhanam, G., Yu, B. M., Ryu, S. I., Sahani, M. and Shenoy, K. V. (2011). Single-trial neural correlates of arm movement preparation. *Neuron* **71**, 555-564. doi:10.1016/j.neuron.2011.05.047

Amador, A., Perl, Y. S., Mindlin, G. B. and Margoliash, D. (2013). Elemental gesture dynamics are encoded by song premotor cortical neurons. *Nature* **495**, 59-64. doi:10.1038/nature11967

Benichov, J. I., Benezra, S. E., Vallentin, D., Globerson, E., Long, M. A. and Tchernichovski, O. (2016). The forebrain song system mediates predictive call timing in female and male zebra finches. *Curr. Biol.* **26**, 309-318. doi:10.1016/j.cub.2015.12.037

Betta, E. and Turatto, M. (2006). Are you ready? I can tell by looking at your microsaccades. *Neuroreport* **17**, 1001-1004. doi:10.1097/01.wnr.0000223392.82198.6d

Bottjer, S. W. and Arnold, A. P. (1984). The role of feedback from the vocal organ. I. Maintenance of stereotypical vocalizations by adult zebra finches. *J. Neurosci.* **4**, 2387-2396. doi:10.1523/JNEUROSCI.04-09-02387.1984

Chen, T.-W., Li, N., Daie, K. and Svoboda, K. (2017). A map of anticipatory activity in mouse motor cortex. *Neuron* **94**, 866-879.e4. doi:10.1016/j.neuron.2017.05.005

Churchland, M. M., Yu, B. M., Ryu, S. I., Santhanam, G. and Shenoy, K. V. (2006a). Neural variability in premotor cortex provides a signature of motor preparation. *J. Neurosci.* **26**, 3697-3712. doi:10.1523/JNEUROSCI.3762-05.2006

Churchland, M. M., Afshar, A. and Shenoy, K. V. (2006b). A central source of movement variability. *Neuron* **52**, 1085-1096. doi:10.1016/j.neuron.2006.10.034

Churchland, M. M., Santhanam, G. and Shenoy, K. V. (2006c). Preparatory activity in premotor and motor cortex reflects the speed of the upcoming reach. *J. Neurophysiol.* **96**, 3130-3146. doi:10.1152/jn.00307.2006

Cohen, R. G. and Rosenbaum, D. A. (2007). Directional bias of limb tremor prior to voluntary movement. *Psychol. Sci.* **18**, 8-12. doi:10.1111/j.1467-9280.2007.01840.x

Coleman, M. J., Roy, A., Wild, J. M. and Mooney, R. (2007). Thalamic gating of auditory responses in telencephalic song control nuclei. *J. Neurosci.* **27**, 10024-10036. doi:10.1523/JNEUROSCI.2215-07.2007

Corneil, B. D. and Munoz, D. P. (2014). Overt responses during covert orienting. *Neuron* **82**, 1230-1243. doi:10.1016/j.neuron.2014.05.040

Daliparthi, V. K., Tachibana, R. O., Cooper, B. G., Hahnloser, R., Kojima, S., Sober, S. J. and Roberts, T. F. (2018). Transitioning between preparatory and precisely sequenced neuronal activity in production of a skilled behavior. *bioRxiv*, 481960. doi:10.1101/481960

Danish, H. H., Aronov, D. and Fee, M. S. (2017). Rhythmic syllable-related activity in a songbird motor thalamic nucleus necessary for learned vocalizations. *PLoS ONE* **12**, e0169568. doi:10.1371/journal.pone.0169568

Fee, M. S. and Scharff, C. (2010). The songbird as a model for the generation and learning of complex sequential behaviors. *ILAR J. Natl. Res. Counc. Inst. Lab. Anim. Resour.* **51**, 362-377. doi:10.1093/ilar.51.4.362

Gao, Z., Davis, C., Thomas, A. M., Economo, M. N., Abrego, A. M., Svoboda, K., De Zeeuw, C. I. and Li, N. (2018). A cortico-cerebellar loop for motor planning. *Nature* **563**, 113-116. doi:10.1038/s41586-018-0633-x

Glaze, C. M. and Troyer, T. W. (2006). Temporal structure in zebra finch song: implications for motor coding. *J. Neurosci.* **26**, 991-1005. doi:10.1523/JNEUROSCI.3387-05.2006

Hahnloser, R. H. R., Kozhevnikov, A. A. and Fee, M. S. (2002). An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* **419**, 65-70. doi:10.1038/nature00974

Hessler, N. A. and Doupe, A. J. (1999). Singing-related neural activity in a dorsal forebrain–basal ganglia circuit of adult zebra finches. *J. Neurosci.* **19**, 10461-10481. doi:10.1523/JNEUROSCI.19-23-10461.1999

Horita, H., Wada, K. and Jarvis, E. D. (2008). Early onset of deafening-induced song deterioration and differential requirements of the pallial-basal ganglia vocal pathway. *Eur. J. Neurosci.* **28**, 2519-2532. doi:10.1111/j.1460-9568.2008.06535.x

Kao, M. H., Wright, B. D. and Doupe, A. J. (2008). Neurons in a forebrain nucleus required for vocal plasticity rapidly switch between precise firing and variable bursting depending on social context. *J. Neurosci.* **28**, 13232-13247. doi:10.1523/JNEUROSCI.2250-08.2008

Kojima, S., Kao, M. H. and Doupe, A. J. (2013). Task-related "cortical" bursting depends critically on basal ganglia input and is linked to vocal plasticity. *Proc. Natl. Acad. Sci. USA* **110**, 4756-4761. doi:10.1073/pnas.1216308110

Konishi, M. (1965). The role of auditory feedback in the control of vocalization in the white-crowned sparrow 1. *Z. Für Tierpsychol.* **22**, 770-783.

Kozhevnikov, A. A. and Fee, M. S. (2007). Singing-related activity of identified HVC neurons in the zebra finch. *J. Neurophysiol.* **97**, 4271-4283. doi:10.1152/jn.00952.2006

Lee, I. H. and Assad, J. A. (2003). Putaminal activity for simple reactions or self-timed movements. *J. Neurophysiol.* **89**, 2528-2537. doi:10.1152/jn.01055.2002

Li, N., Chen, T.-W., Guo, Z. V., Gerfen, C. R. and Svoboda, K. (2015). A motor cortex circuit for motor planning and movement. *Nature* **519**, 51-56. doi:10.1038/nature14178

Long, M. A. and Fee, M. S. (2008). Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* **456**, 189-194. doi:10.1038/nature07448

Maimon, G. and Assad, J. A. (2006). Parietal area 5 and the initiation of self-timed movements versus simple reactions. *J. Neurosci.* **26**, 2487-2498. doi:10.1523/JNEUROSCI.3590-05.2006

Méndez, J. M., Dall'asén, A. G. and Goller, F. (2010). Disrupting vagal feedback affects birdsong motor control. *J. Exp. Biol.* **213**, 4193-4204. doi:10.1242/jeb.045369

Murakami, M., Vicente, M. I., Costa, G. M. and Mainen, Z. F. (2014). Neural antecedents of self-initiated actions in secondary motor cortex. *Nat. Neurosci.* **17**, 1574-1582. doi:10.1038/nn.3826

Nordeen, K. W. and Nordeen, E. J. (1992). Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behav. Neural Biol.* **57**, 58-66. doi:10.1016/0163-1047(92)90757-U

Price, P. H. (1979). Developmental determinants of structure in zebra finch song. *J. Comp. Physiol. Psychol.* **93**, 260. doi:10.1037/h0077553

Rajan, R. (2018). Pre-bout neural activity changes in premotor nucleus HVC correlate with successful initiation of learned song sequence. *J. Neurosci.* **38**, 5925-5938. doi:10.1523/JNEUROSCI.3003-17.2018

Rajan, R. and Doupe, A. J. (2013). Behavioral and neural signatures of readiness to initiate a learned motor sequence. *Curr. Biol.* **23**, 87-93. doi:10.1016/j.cub.2012.11.040

Romo, R. and Schultz, W. (1987). Neuronal activity preceding self-initiated or externally timed arm movements in area 6 of monkey cortex. *Exp. Brain Res.* **67**, 656-662. doi:10.1007/BF00247297

Roy, A. and Mooney, R. (2007). Auditory plasticity in a Basal Ganglia–forebrain pathway during decrystallization of adult birdsong. *J. Neurosci.* **27**, 6374-6387. doi:10.1523/JNEUROSCI.0894-07.2007

Shenoy, K. V., Kaufman, M. T., Sahani, M. and Churchland, M. M. (2011). A dynamical systems view of motor preparation: implications for neural prosthetic system design. *Prog. Brain Res.* **192**, 33-58. doi:10.1016/B978-0-444-53355-5.00003-8

Shenoy, K. V., Sahani, M. and Churchland, M. M. (2013). Cortical control of arm movements: a dynamical systems perspective. *Annu. Rev. Neurosci.* **36**, 337-359. doi:10.1146/annurev-neuro-062111-150509

Simpson, H. B. and Vicario, D. S. (1990). Brain pathways for learned and unlearned vocalizations differ in zebra finches. *J. Neurosci. Off. J. Soc. Neurosci.* **10**, 1541-1556. doi:10.1523/JNEUROSCI.10-05-01541.1990

Sossinka, R. and Böhner, J. (1980). Song Types in the Zebra Finch Poephila guttata castanotis1. *Z. Für Tierpsychol.* **53**, 123-132. doi:10.1111/j.1439-0310.1980.tb01044.x

Tanji, J. and Evarts, E. V. (1976). Anticipatory activity of motor cortex neurons in relation to direction of an intended movement. *J. Neurophysiol.* **39**, 1062-1068. doi:10.1152/jn.1976.39.5.1062

Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B. and Mitra, P. P. (2000). A procedure for an automated measurement of song similarity. *Anim. Behav.* **59**, 1167-1176. doi:10.1006/anbe.1999.1416

Vallentin, D. and Long, M. A. (2015). Motor origin of precise synaptic inputs onto forebrain neurons driving a skilled behavior. *J. Neurosci.* **35**, 299-307. doi:10.1523/JNEUROSCI.3698-14.2015

Vicario, D. S. (1991). Contributions of syringeal muscles to respiration and vocalization in the zebra finch. *J. Neurobiol.* **22**, 63-73. doi:10.1002/neu.480220107

Vicario, D. S. (2004). Using learned calls to study sensory-motor integration in songbirds. *Ann. N. Y. Acad. Sci.* **1016**, 246-262. doi:10.1196/annals.1298.040

Vyssotski, A. L., Stepien, A. E., Keller, G. B. and Hahnloser, R. H. R. (2016). A neural code that is isometric to vocal output and correlates with its sensory consequences. *PLoS Biol.* **14**, e2000317. doi:10.1371/journal.pbio.2000317

Wang, C.-A., Brien, D. C. and Munoz, D. P. (2015). Pupil size reveals preparatory processes in the generation of pro-saccades and anti-saccades. *Eur. J. Neurosci.* **41**, 1102-1110. doi:10.1111/ejn.12883

Williams, H. and McKibben, J. R. (1992). Changes in stereotyped central motor patterns controlling vocalization are induced by peripheral nerve injury. *Behav. Neural Biol.* **57**, 67-78. doi:10.1016/0163-1047(92)90768-Y

Wittenbach, J. D., Bouchard, K. E., Brainard, M. S. and Jin, D. Z. (2015). An adapting auditory-motor feedback loop can contribute to generating vocal repetition. *PLoS Comput. Biol.* **11**, e1004471. doi:10.1371/journal.pcbi.1004471

Woolley, S. C., Rajan, R., Joshua, M. and Doupe, A. J. (2014). Emergence of context-dependent variability across a basal ganglia network. *Neuron* **82**, 208-223. doi:10.1016/j.neuron.2014.01.039

Yu, A. C. and Margoliash, D. (1996). Temporal hierarchical control of singing in birds. *Science* **273**, 1871-1875. doi:10.1126/science.273.5283.1871

Zann, R. (1993). Structure, sequence and evolution of song elements in wild Australian zebra finches. *The Auk* **110**, 702-715. doi:10.2307/4088626

Zann, R. A. (1996). *The Zebra Finch: A Synthesis of Field and Laboratory Studies*. Oxford University Press.

SUPPLEMENTARY INFORMATION

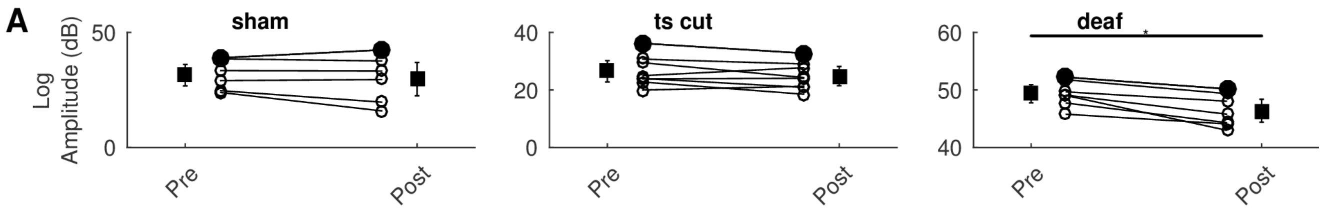

Figure S1 Changes in motif syllable log amplitude after removal of sensory feedback

(A) Log amplitude of motif syllables pre and post sham surgery (n=6 birds, left column), ts-cut surgery (n=8 birds, middle column) or deafening (n=7 birds, right column). Circles represent average across all motif syllables for individual birds and lines connect data from the same bird. Squares and whiskers represent mean ± s.e.m. across birds. The filled circles in the first, second and third columns (left to right) represent the values for the birds shown in Figure 2 (A), (B) and (C) respectively. * represents $p < 0.05$, ** represents $p < 0.01$, Wilcoxon sign-rank test.

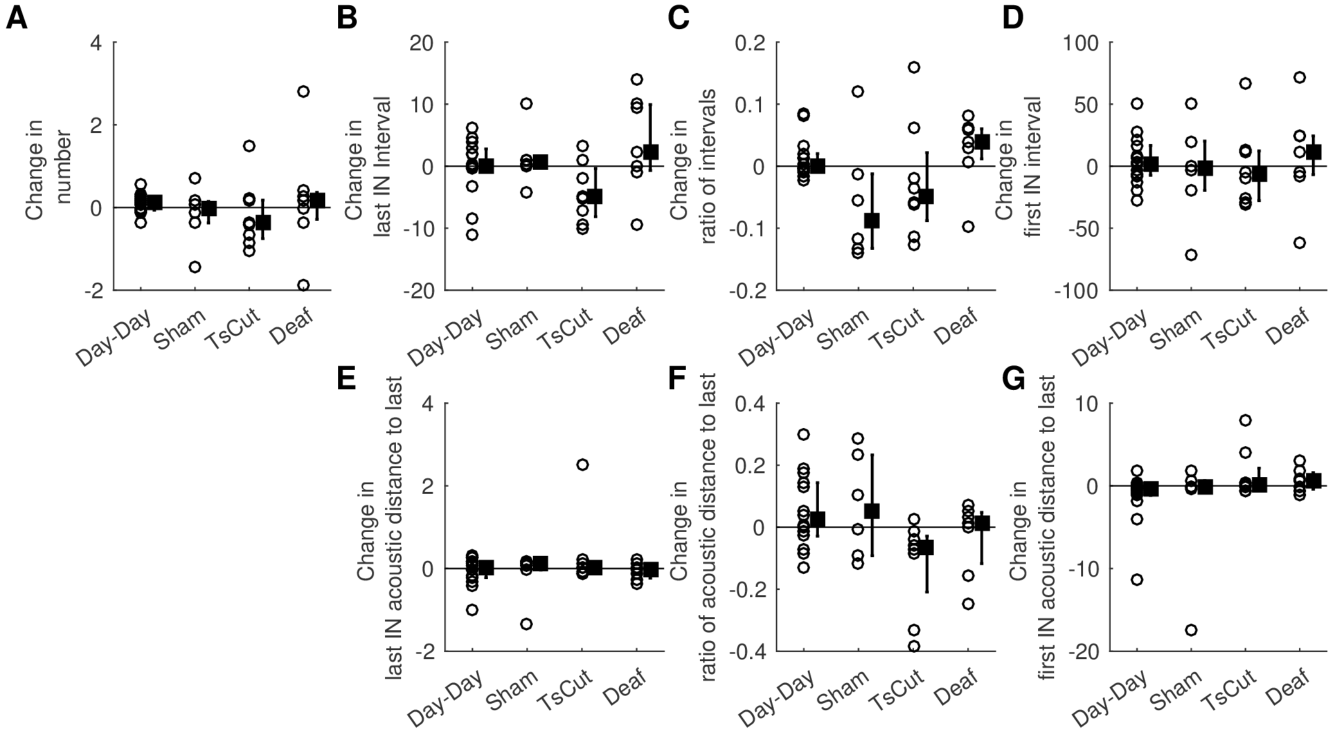

Figure S2 Comparison of day-to-day changes in IN properties with changes observed after surgery for sham-surgery, ts-cut and deaf birds

(A), (B), (C), (D), (E), (F) and (G) Comparison of change in IN properties before and after sham-surgery (n=6 birds), ts-cut surgery (n=8 birds) or deafening (n=7 birds) with day-to-day changes in IN properties in n=14 birds. Changes in IN number (A), Last IN interval (B), Ratio of successive intervals (C), first IN interval (D), last IN acoustic distance to last IN (E), ratio of acoustic distances of successive INs to the last IN (F) and first IN acoustic distance to the last IN (G) are plotted. Circles represent individual birds. Squares and whiskers represent median and inter-quartile range across all birds. (A)-(G) $p > 0.05$, KruskalWallis ANOVA.

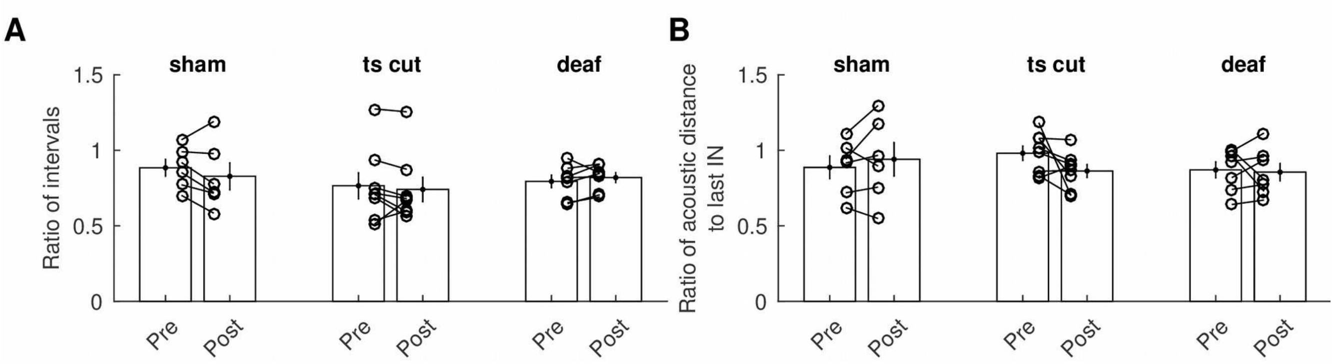

Figure S3 Progression of IN timing and acoustic features is not altered by removal of auditory or proprioceptive feedback

(A), (B) Ratio of successive inter-IN intervals (A) and ratio of acoustic distances of successive INs from the last IN (B) for all birds before and after sham-surgery (n=6 birds), ts-cut surgery (n=8 birds) or deafening (n=7 birds). Each circle represents data from an individual bird and lines connect data from the same bird. Bars and whiskers represent mean ± s.e.m. across birds. $p > 0.05$, Wilcoxon sign-rank test between pre and post surgery.

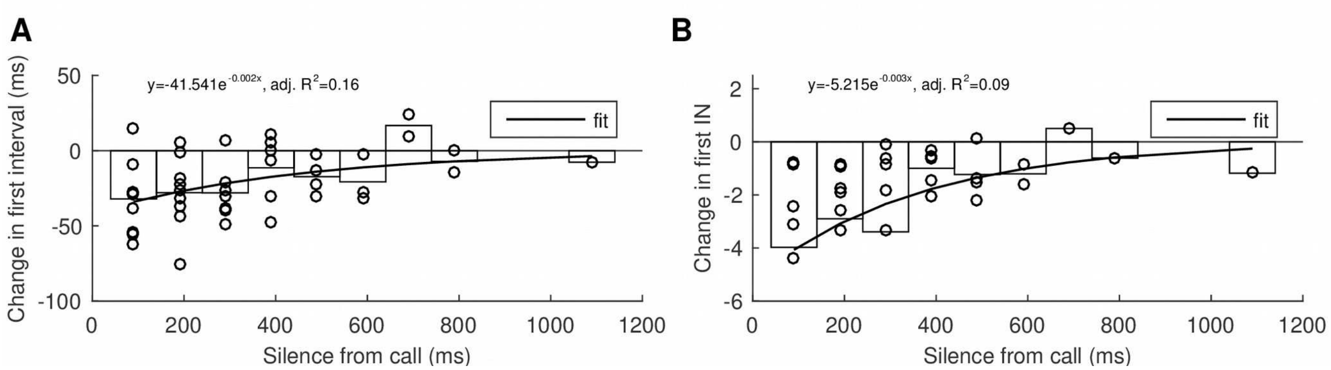

Figure S4 Changes in IN "initial" state in bouts with calls before the first IN also depend on the time between the end of the call and the start of the first IN

(A) Silence between the end of the call and the beginning of the first IN vs. change in interval between the first two INs in call song bouts relative to the interval between the first two INs in IN song bouts. Each circle represents one bird. Bars represent mean across birds (n=16 birds) and the line represent an exponential fit to the data ($y=-41.541e^{-0.002x}$, adjusted $R^2 = 0.16$).

(B) Silence between the end of the call and the beginning of the first IN vs. change in acoustic distance of the first IN to the last IN in call song bouts relative to acoustic distance of the first IN from the last IN in IN song bouts. Each circle represents one bird. Bars represent mean across birds (n=16 birds) and the line represent an exponential fit to the data ($y=-5.215e^{-0.003x}$, adjusted $R^2 = 0.09$).

Table S1 Details of statistical tests and associated p-values for all analyses

S.No	Figure No.	Paremeters	Groups Compared	Test	RESULTS					
1	2B	Temporal Similarity	Sham: Pre-Post	signrank	p-value	0.0313				
			Ts cut: Pre-Post	signrank	p-value	0.0078				
			Deaf: Pre-Post	signrank	p-value	0.0156				
			Random birds-Sham Post	kruskal-wallis ANOVA	p-value	0.0011				
			Random birds-Ts cut Post	kruskal-wallis ANOVA	p-value	0.0004				
			Random birds-Deaf Post	kruskal-wallis ANOVA	p-value	0.0006				
2	2C	Spectral Similarity	Sham: Pre-Post	signrank	p-value	0.0937				
			Ts cut: Pre-Post	signrank	p-value	0.0039				
			Deaf: Pre-Post	signrank	p-value	0.0156				
			Random birds-Ts cut Post	kruskal-wallis ANOVA	p-value	0.3272				
			Random birds-Deaf Post	kruskal-wallis ANOVA	p-value	0.0006				
3	3A	Duration of INs	Sham: Pre-Post	signrank	p-value	1				
			Ts cut: Pre-Post	signrank	p-value	0.1953				
			Deaf: Pre-Post	signrank	p-value	0.5781				
4	3B	Mean frequency of INs	Sham: Pre-Post	signrank	p-value	0.3125				
			Ts cut: Pre-Post	signrank	p-value	0.3125				
			Deaf: Pre-Post	signrank	p-value	0.0469				
5	3C	Entropy of INs	Sham: Pre-Post	signrank	p-value	1				
			Ts cut: Pre-Post	signrank	p-value	0.3828				
			Deaf: Pre-Post	signrank	p-value	0.2188				
6	3D	Log Amplitude of INs	Sham: Pre-Post	signrank	p-value	0.4375				
			Ts cut: Pre-Post	signrank	p-value	0.8438				
			Deaf: Pre-Post	signrank	p-value	0.0156				
7	3E	Pitch Goodness of INs	Sham: Pre-Post	signrank	p-value	0.5625				
			Ts cut: Pre-Post	signrank	p-value	0.0156				
			Deaf: Pre-Post	signrank	p-value	0.2188				
8	3F	Frequency modulation of INs	Sham: Pre-Post	signrank	p-value	0.5625				
			Ts cut: Pre-Post	signrank	p-value	0.0078				
			Deaf: Pre-Post	signrank	p-value	0.375				
9	4D	Mean Number of INs	Sham: Pre-Post	signrank	p-value	0.8438				
			Ts cut: Pre-Post	signrank	p-value	0.3828				
			Deaf: Pre-Post	signrank	p-value	0.6875				
10	4E	CV Number of INs	Sham: Pre-Post	signrank	p-value	0.1563				
			Ts cut: Pre-Post	signrank	p-value	0.7422				
			Deaf: Pre-Post	signrank	p-value	0.5781				
11	5D	Mean IN interval	Sham: first and last interval on pre and post day	Repeated Measures two-wayANOVA	Source'	SS'	df'	MS'	F'	p-value
					Pre-Post'	9.7247	1	9.7247	0.0215	0.8891
					First-Last'	16433.383	1	16433.383	6.6752	0.0492
					Pre-Post x First-Last'	43.1520	1	43.1520	0.1068	0.7571
		Mean IN interval	Ts cut: first and last interval on pre and post day	Repeated Measures two-wayANOVA	Source'	SS'	df'	MS'	F'	p-value
					Pre-Post'	61.7874	1	61.7874	0.2585	0.6268
					First-Last'	44717.540	1	44717.540	91.7874	0.0000
					Pre-Post x First-Last'	16.4337	1	16.4337	0.0558	0.8200
		Mean IN interval	Deaf: first and last interval on pre and post day	Repeated Measures two-wayANOVA	Source'	SS'	df'	MS'	F'	p-value
					Pre-Post'	236.1533	1	236.1533	0.6847	0.4397
					First-Last'	29916.057	1	29916.057	94.9916	0.0001
					Pre-Post x First-Last'	32.9580	1	32.9580	0.0646	0.8078
12	5E	CV of Interval	Sham: first and last interval on pre and post day	Repeated Measures two-wayANOVA	Source'	SS'	df'	MS'	F'	p-value
					Pre-Post'	0.0024	1	0.0024	0.1551	0.7099
					First-Last'	1.6773	1	1.6773	19.2118	0.0071
					Pre-Post x First-Last'	0.0080	1	0.0080	0.3355	0.5875
			Ts cut: first and last interval on pre and post day	Repeated Measures two-wayANOVA	Source'	SS'	df'	MS'	F'	p-value
					Pre-Post'	0.0155	1	0.0155	1.3947	0.2762
					First-Last'	1.8026	1	1.8026	35.8553	0.0005
					Pre-Post x First-Last'	0.0014	1	0.0014	0.1947	0.6723
			Deaf: first and last interval on pre and post day	Repeated Measures two-wayANOVA	Source'	SS'	df'	MS'	F'	p-value
					Pre-Post'	0.0343	1	0.0343	0.6613	0.4472
					First-Last'	0.3953	1	0.3953	6.9461	0.0388
					Pre-Post x First-Last'	0.0752	1	0.0752	2.2303	0.1859

13	6D	Acoustic distance to last IN	Sham: first and last IN on pre and post day	Repeated Measures two-wayANOVA	Source'	SS'	df'	MS'	F'	p-value
					Pre-Post'	11.1136	1	11.1136	0.7358	0.4302
					First-Last'	65.5787	1	65.5787	1.7185	0.2469
					Pre-Post x First-Last'	8.4402	1	8.4402	0.8231	0.4059
			Ts cut: first and last IN on pre and post day	Repeated Measures two-wayANOVA	Source'	SS'	df'	MS'	F'	p-value
					Pre-Post'	4.5421	1	4.5421	1.5773	0.2494
					First-Last'	74.5198	1	74.5198	5.3249	0.0544
					Pre-Post x First-Last'	2.2832	1	2.2832	1.1463	0.3198
			Deaf: first and last IN on pre and post day	Repeated Measures two-wayANOVA	Source'	SS'	df'	MS'	F'	p-value
					Pre-Post'	0.9207	1	0.9207	2.7493	0.1484
					First-Last'	67.7705	1	67.7705	24.9848	0.0025
					Pre-Post x First-Last'	1.3146	1	1.3146	3.5509	0.1085
14	7C	Number of INs	Feedback-intact: IN song bouts, calls <=200ms before the first IN, calls >200ms before the first IN	Repeated Measures one-wayANOVA		SumSq	df	MeanSq	F	p-value
					(Intercept):BoutType	1.8560	2	0.9280	8.4085	0.0026
					Bird:BoutType	0.2150	2	0.1075	0.9741	0.3966
					Error (BoutType)	1.9866	18	0.1104	1.0000	0.5000
				Multcompare						
				Bout Type 1	Bout Type 2	Difference	Std Err	pValue	Lower 95 % CI	Upper 95 % CI
				IN song bouts	Calls < 200 ms before first IN	0.7968	0.1355	0.0006	0.4184	1.1752
				IN song bouts	Calls>= 200 ms before first IN	0.0561	0.1613	0.9359	-0.3941	0.5064
				Calls < 200 ms before first IN	IN song bouts	-0.7968	0.1355	0.0006	-1.1752	-0.4184
				Calls < 200 ms before first IN	Calls>= 200 ms before first IN	-0.7407	0.1258	0.0006	-1.0920	-0.3894
				Calls>= 200 ms before first IN	IN song bouts	-0.0561	0.1613	0.9359	-0.5064	0.3941
				Calls>= 200 ms before first IN	Calls < 200 ms before first IN	0.7407	0.1258	0.0006	0.3894	1.0920
15	7D	Number of INs	Feedback-deprived: IN song bouts, calls <=200ms before the first IN, calls >200ms before the first IN	Repeated Measures one-wayANOVA		SumSq	df	MeanSq	F	p-value
					(Intercept):BoutType	2.3666	2	1.1833	7.8697	0.0066
					Bird:BoutType	0.7949	2	0.3974	2.6431	0.1119
					Error (BoutType)	1.8043	12	0.1504	1.0000	0.5000
				Multcompare						
				Bout Type 1	Bout Type 2	Difference	Std Err	pValue	Lower 95 % CI	Upper 95 % CI
				IN song bouts	Calls < 200 ms before first IN	0.9735	0.2100	0.0085	0.3290	1.6179
				IN song bouts	Calls>= 200 ms before first IN	0.4323	0.1208	0.0270	0.0618	0.8029
				Calls < 200 ms before first IN	IN song bouts	-0.9735	0.2100	0.0085	-1.6179	-0.3290
				Calls < 200 ms before first IN	Calls>= 200 ms before first IN	-0.5411	0.2325	0.1273	-1.2546	0.1723
				Calls>= 200 ms before first IN	IN song bouts	-0.4323	0.1208	0.0270	-0.8029	-0.0618
				Calls>= 200 ms before first IN	Calls < 200 ms before first IN	0.5411	0.2325	0.1273	-0.1723	1.2546
16	8A	Fisrt IN interval mean	Feedback-intact: IN song bouts, calls <=200ms before the first IN, calls >200ms before the first IN	Repeated Measures one-wayANOVA		SumSq	df	MeanSq	F	p-value
					(Intercept):BoutType	3930.1301	2	1965.0651	7.4006	0.0045
					Bird:BoutType	875.8583	2	437.9291	1.6493	0.2199
					Error (BoutType)	4779.5195	18	265.5289	1.0000	0.5000
				Multcompare						
				Bout Type 1	Bout Type 2	Difference	Std Err	pValue	Lower 95 % CI	Upper 95 % CI
				IN song bouts	Calls < 200 ms before first IN	34.7853	8.0407	0.0049	12.3356	57.2350
				IN song bouts	Calls>= 200 ms before first IN	13.1458	6.3450	0.1509	-4.5694	30.8610
				Calls < 200 ms before first IN	IN song bouts	-34.7853	8.0407	0.0049	-57.2350	-12.3356
				Calls < 200 ms before first IN	Calls>= 200 ms before first IN	-21.6395	6.3184	0.0187	-39.2805	-3.9985
				Calls>= 200 ms before first IN	IN song bouts	-13.1458	6.3450	0.1509	-30.8610	4.5694
				Calls>= 200 ms before first IN	Calls < 200 ms before first IN	21.6395	6.3184	0.0187	3.9985	39.2805
17	8C	Fisrt IN interval CV	Feedback-intact: IN song bouts, calls <=200ms before the first IN, calls >200ms before the first IN	Repeated Measures one-wayANOVA		SumSq	df	MeanSq	F	p-value
					(Intercept):BoutType	0.1654	2	0.0827	1.7250	0.2064
					Bird:BoutType	0.1338	2	0.0669	1.3957	0.2732
					Error (BoutType)	0.8629	18	0.0479	1.0000	0.5000
18	8E	First IN distance to last IN	Feedback-intact: IN song bouts, calls <=200ms before the first IN, calls >200ms before the first IN	Repeated Measures one-wayANOVA		SumSq	df	MeanSq	F	p-value
					(Intercept):BoutType	9.0163	2	4.5082	0.9289	0.4131
					Bird:BoutType	0.4322	2	0.2161	0.0445	0.9566
					Error (BoutType)	87.3577	18	4.8532	1.0000	0.5000
19	8B	First IN interval mean	Feedback-deprived: IN song bouts, calls <=200ms before the first IN, calls >200ms before the first IN	Repeated Measures one-wayANOVA		SumSq	df	MeanSq	F	p-value
					(Intercept):BoutType	2217.5005	2	1108.7503	4.1573	0.0425
					Bird:BoutType	253.0544	2	126.5272	0.4744	0.6334
					Error (BoutType)	3200.3707	12	266.6976	1.0000	0.5000
				Multcompare						
				Bout Type 1	Bout Type 2	Difference	Std Err	pValue	Lower 95 % CI	Upper 95 % CI
				IN song bouts	Calls < 200 ms before first IN	51.1831	7.7612	0.0014	27.3697	74.9966
				IN song bouts	Calls>= 200 ms before first IN	19.1572	8.1192	0.1221	-5.7548	44.0693
				Calls < 200 ms before first IN	IN song bouts	-51.1831	7.7612	0.0014	-74.9966	-27.3697
				Calls < 200 ms before first IN	Calls>= 200 ms before first IN	-32.0259	8.5945	0.0228	-58.3961	-5.6557
				Calls>= 200 ms before first IN	IN song bouts	-19.1572	8.1192	0.1221	-44.0693	5.7548
				Calls>= 200 ms before first IN	Calls < 200 ms before first IN	32.0259	8.5945	0.0228	5.6557	58.3961

20	8D	Fisrt IN interval CV	Feedback-deprived: IN song bouts, calls <=200ms before the first IN, calls >200ms before the first IN	Repeated Measures one-wayANOVA		SumSq	df	MeanSq	F	p-value
					(Intercept):BoutType	0.0287	2	0.0144	0.4083	0.6737
					Bird:BoutType	0.0016	2	0.0008	0.0224	0.9779
					Error (BoutType)	0.4225	12	0.0352	1.0000	0.5000
21	8F	First IN distance to last IN	Feedback-deprived: IN song bouts, calls <=200ms before the first IN, calls >200ms before the first IN	Repeated Measures one-wayANOVA		SumSq	df	MeanSq	F	p-value
					(Intercept):BoutType	0.2693	2	0.1347	0.7387	0.4983
					Bird:BoutType	0.2794	2	0.1397	0.7663	0.4862
					Error (BoutType)	2.1878	12	0.1823	1.0000	0.5000
36	S1A	Log amplitude of motif syllables	Sham: Pre-Post	signrank	p-value	0.4375				
			Ts cut: Pre-Post	signrank	p-value	0.1484				
			Deaf: Pre-Post	signrank	p-value	0.0156				
22	S2A	Change in IN Number	Intact (da-day), Sham (pre-post), Ts cut (pre-post), Deaf (pre-post)	kruskal-wallis ANOVA	p-value	0.4539				
23	S2B	Change in last IN interval	Intact (da-day), Sham (pre-post), Ts cut (pre-post), Deaf (pre-post)	kruskal-wallis ANOVA	p-value	0.1543				
24	S2C	Change in ratio of IN intervals	Intact (da-day), Sham (pre-post), Ts cut (pre-post), Deaf (pre-post)	kruskal-wallis ANOVA	p-value	0.0519				
25	S2D	Change in first IN interval	Intact (da-day), Sham (pre-post), Ts cut (pre-post), Deaf (pre-post)	kruskal-wallis ANOVA	p-value	0.7167				
26	S2E	Change in last IN distance to last IN	Intact (da-day), Sham (pre-post), Ts cut (pre-post), Deaf (pre-post)	kruskal-wallis ANOVA	p-value	0.8503				
27	S2F	Change in ratio of distance to last IN	Intact (da-day), Sham (pre-post), Ts cut (pre-post), Deaf (pre-post)	kruskal-wallis ANOVA	p-value	0.0905				
28	S2G	Change in first IN distance to last IN	Intact (da-day), Sham (pre-post), Ts cut (pre-post), Deaf (pre-post)	kruskal-wallis ANOVA	p-value	0.2122				
29	S3A	Ratio of intervals	Sham: Pre-Post	signrank	p-value	0.2188				
			Ts cut: Pre-Post	signrank	p-value	0.6749				
			Deaf: Pre-Post	signrank	p-value	0.8471				
30	S3B	Ratio of distance to last IN	Sham: Pre-Post	signrank	p-value	0.5625				
			Ts cut: Pre-Post	signrank	p-value	0.0781				
			Deaf: Pre-Post	signrank	p-value	0.9375				