

Lesions in a songbird vocal circuit increase variability in song syntax

Avani Koparkar, Timothy L. Warren, Jonathan D. Charlesworth, Sooyoon Shin, Michael S. Brainard, Lena Veit 

Neurobiology of Vocal Communication, Institute for Neurobiology, University of Tübingen, Germany • Center for Integrative Neuroscience, University of California, San Francisco, USA and Howard Hughes Medical Institute • Departments of Horticulture and Integrative Biology, Oregon State University, Corvallis, OR • Department of Clinical Research, Noctrix Health, Inc., Pleasanton, CA

Reviewed Preprint

Revised by authors after peer review.

About eLife's process

Reviewed preprint version 2

March 21, 2024 (this version)

Reviewed preprint version 1

January 10, 2024

Posted to preprint server

November 6, 2023

Sent for peer review

October 18, 2023

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

Abstract

Complex motor skills like speech and dance are composed of ordered sequences of simpler elements, but the neuronal basis for syntactic ordering of individual actions into sequences is poorly understood. Birdsong is a learned vocal behavior composed of syntactically ordered sequences of individual syllables. Activity in song premotor nucleus HVC (proper name) has been linked to the control of syllable sequencing, but sequencing may also be affected by its recurrent inputs. We here test the contribution of one of HVC's inputs, mMAN (*medial magnocellular nucleus of the anterior nidopallium*), to the variable songs of adult male Bengalese finches (*Lonchura striata domestica*). The syntax of Bengalese song includes several patterns: 1) *chunks*, where syllables follow stereotypical order 2) *branch points*, where a given syllable can be followed by two or more different syllables in a probabilistic manner and 3) *repeat phrases*, where an individual syllable is repeated a variable number of times. We found that after bilateral lesions of mMAN, the acoustic structure of syllables remained largely intact, but sequencing became more variable for each of these patterns, seen by 'breaks' in previously stereotyped chunks, increased uncertainty at branch points and increased variability of repeat numbers. This increase in sequencing variability highlights the potential importance of regions projecting to HVC in the ordering of vocal elements. Previous studies on zebra finches found little effect of mMAN lesions on their relatively stereotyped adult song. In contrast, our results suggest that mMAN contributes to sequencing the variable songs of Bengalese finches and highlight the utility of species with more complex song syntax in investigating neuronal control of ordered motor sequences.

eLife assessment

Songbirds provide a tractable model system to study mechanisms of vocal production and sequencing, and past work showed that the lesions to LMAN, the output of a basal ganglia thalamocortical loop, reduced vocal variability, consistent with a role in motor exploration. In this **fundamental** work, the authors rigorously examined how lesions to an understudied neighboring region, MMAN, part of a parallel basal ganglia loop, affect singing in Bengalese finches, whose songs exhibit complex sequential transitions. The authors provide **compelling** evidence that MMAN lesions resulted in increased sequential variability but do not affect syllable acoustic structure, showing that distinct frontal systems can have distinct functions for producing and sequencing song syllables.

Introduction

Complex behaviors are composed of sequences of simpler motor elements. For example, in human speech, sentences are formed by syllables that are flexibly sequenced according to syntactic rules (Berwick *et al.*, 2011 [↗](#)), which define the order of individual elements. Production of motor sequences is important not only for speech, but also for most other natural movements. Despite the ubiquity of syntactically organized sequential behaviors, little is known about how the brain produces flexible motor sequences (Lashley, 1951 [↗](#); Tanji, 2001 [↗](#); Aldridge and Berridge, 2003 [↗](#); Wiltchko *et al.*, 2015 [↗](#); Veit *et al.*, 2021 [↗](#)).

Birdsong is a learned behavior that is well suited for the study of the neural control of motor sequences. Song is composed of discrete elements, called syllables, that are organized according to syntactic rules. The complexity of song syntax varies widely across different songbird species - from the simple and repetitive songs of owl finches (Wang *et al.*, 2019 [↗](#)) and zebra finches (Zann, 1996 [↗](#)) to the variable songs of Bengalese finches (Honda and Okanoya, 1999 [↗](#)) and canaries (Nottebohm, Stokes and Leonard, 1976 [↗](#); Cohen *et al.*, 2020 [↗](#)) up to the immense repertoire of nightingales (Hultsch *et al.*, 2004 [↗](#); Costalunga *et al.*, 2023 [↗](#)). The neuronal mechanisms of birdsong production have been primarily studied in the zebra finch, partly by leveraging the structure of zebra finch song, in which syllables are produced in a relatively stereotyped and linear sequence. Consequently, we have a detailed understanding of the neuronal mechanisms for producing individual song syllables in this species, but we know little about how syllables are organized into the more complex and variable sequences that are characteristic of many other species (Ivanitskii and Marova, 2022 [↗](#)).

Song production is controlled by neuronal activity in a set of brain nuclei called the song motor pathway, including premotor nucleus HVC (proper name). HVC neurons projecting to primary motor nucleus RA (*robust nucleus of the arcopallium*) encode syllable identity and have been shown to play a role in controlling syllable timing (Simpson and Vicario, 1990 [↗](#); Vu, Mazurek and Kuo, 1994 [↗](#); Hahnloser, Kozhevnikov and Fee, 2002 [↗](#); Long and Fee, 2008 [↗](#); Aronov *et al.*, 2011 [↗](#); Ölveczky *et al.*, 2011 [↗](#); Lynch *et al.*, 2016 [↗](#); Picardo *et al.*, 2016 [↗](#); Zhang *et al.*, 2017 [↗](#)). HVC activity in Bengalese finches (Sakata and Brainard, 2008 [↗](#); Fujimoto, Hasegawa and Watanabe, 2011 [↗](#)) and canaries (Cohen *et al.*, 2020 [↗](#)) has also been shown to encode syllable identity, and can additionally influence sequencing and reflect the sequential context in which syllables are produced (Fujimoto, Hasegawa and Watanabe, 2011 [↗](#); Zhang *et al.*, 2017 [↗](#); Cohen *et al.*, 2020 [↗](#)). The nuclei of the motor pathway are essential for song production (Nottebohm,

Stokes and Leonard, 1976 [↗](#); Scharff and Nottebohm, 1991 [↗](#)), but the extent to which these sequence-selective activity patterns in HVC reflect internal neural dynamics or are shaped by recurrent inputs is poorly understood.

Although HVC is often considered to be at the top of a hierarchy in the song control circuit, it is part of a network of brain nuclei involved in song production and receives multiple inputs that might also contribute to premotor activity (McCasland, 1987 [↗](#); Hosino and Okanoya, 2000 [↗](#); Jin, 2009 [↗](#); Hamaguchi, Tanaka and Mooney, 2016 [↗](#); Vyssotski *et al.*, 2016 [↗](#)). One prominent source of input to HVC is mMAN (*medial magnocellular nucleus of the anterior nidopallium*) which is part of a recurrent loop that bilaterally connects the two hemispheres through RA and bilateral connections from thalamic nucleus DMP (*dorsomedial nucleus of the posterior thalamus*, **Figure 1E** [↗](#)) and could thus be ideally positioned to convey sequence-related activity across hemispheres (Vates, Vicario and Nottebohm, 1997 [↗](#); Schmidt, Ashmore and Vu, 2004 [↗](#); Williams, Nast and Coleman, 2012 [↗](#)). A parallel recurrent pathway, which projects to RA through nucleus IMAN (*lateral magnocellular nucleus of the anterior nidopallium*, **Figure 1E** [↗](#)) has been shown to contribute variability to syllable pitch, a parameter that is controlled by RA premotor activity (Kao, Doupe and Brainard, 2005 [↗](#); Sober, Wohlgemuth and Brainard, 2008 [↗](#); Miller, Cheung and Brainard, 2017 [↗](#)). IMAN-guided variability can serve as a form of motor exploration which is essential for learning (Sober and Brainard, 2012 [↗](#); Dhawale, Smith and Ölveczky, 2017 [↗](#)), regulating social context-dependent changes in pitch variability (Kao and Brainard, 2006 [↗](#); Hampton, Sakata and Brainard, 2009 [↗](#)), and biasing pitch in the direction of improved motor performance during learning (Andalman and Fee, 2009 [↗](#); Warren *et al.*, 2012 [↗](#); Tian and Brainard, 2017 [↗](#)). This raises the question of whether mMAN, as the output of a parallel recurrent pathway that projects onto HVC, might have a similar role in contributing to premotor activity in its projection target, with the potential to influence syllable identity or sequencing variability (Kubikova, Turner and Jarvis, 2007 [↗](#); Seki and Okanoya, 2008 [↗](#); Ali *et al.*, 2013 [↗](#)). Previous studies in zebra finches showed that mMAN lesions in juveniles disrupt normal song learning, while lesions in adults affected song initiation, but not subsequent production of song in this species (Foster and Bottjer, 2001 [↗](#); Horita, Wada and Jarvis, 2008 [↗](#); Ali *et al.*, 2013 [↗](#)).

We tested the hypothesis that mMAN contributes to syllable sequencing by studying the effect of bilateral mMAN lesions on song production in adult Bengalese finches. We found that lesions had little effect on the acoustic structure of individual syllables but led to an increase in the variability of multiple aspects of syllable sequencing. These results highlight the potential importance of recurrent inputs such as mMAN in shaping the syntactical structure of adult birdsongs.

Methods

Subjects and sound recordings

Experiments were carried out on seven adult male Bengalese finches (*Lonchura striata domestica*) obtained from the Brainard lab's breeding colony at University of California, San Francisco (median age 640 days post-hatch, range 149-1049 at start of experiment). Birds were raised with their parents and then housed in same-sex group cages. For the experiments, birds were placed in individual sound-attenuating boxes (Acoustic Systems, Austin, TX) and maintained on a 14:10 hour light: dark period. Song was recorded using an omnidirectional microphone above the cage using custom LabView software for continuous monitoring of song output (Turner and Brainard, 2007 [↗](#)). All procedures were performed in accordance with animal care protocols approved by the University of California, San Francisco Institutional Animal Care and Use Committee (IACUC).

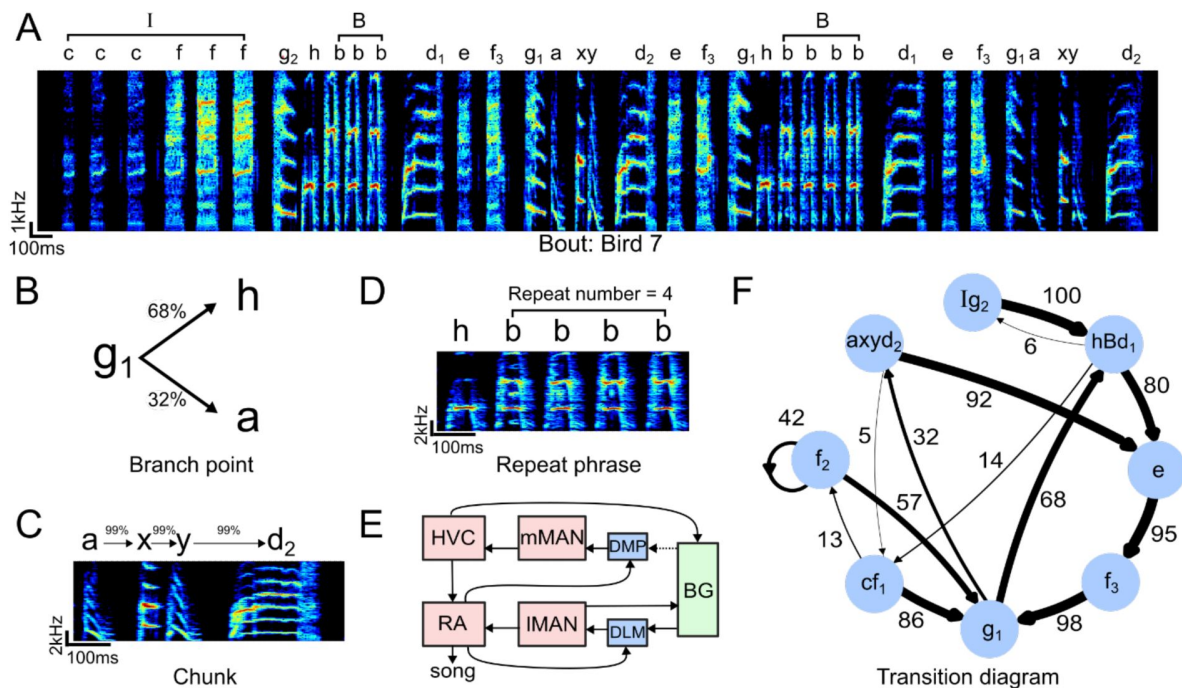


Figure 1

Structure of Bengalese finch song (A) Example spectrogram (bird 7) depicting an entire song bout with introductory state 'I', repeat phrase 'B' and individual syllables. (B) Example transition diagram depicting a branchpoint with variable sequencing. Numbers above the arrows denote transition probabilities in percent. (C) Example spectrogram of a chunk. Chunks are defined as highly stereotyped sequences of syllables which only have a single in- and output branch and are condensed into one state in the transition diagram (see Methods). (D) Example spectrogram of a repeat phrase, summarized by capital letter 'B' in the transition diagram. The repeating syllable (here, syllable 'b') repeats a variable number of times across different instances of the repeat phrase. (E) Schematic showing recurrent pathways projecting onto motor pathway nuclei through IMAN and mMAN. Red: pallial nuclei, blue: thalamic nuclei, green: basal ganglia. Dotted line indicates suspected connection by Kubikova et al. 2007. Abbrev: mMAN: medial magnocellular nucleus of the anterior nidopallium, DMP: dorsomedial nucleus of the posterior thalamus, BG: basal ganglia, DLM: medial portion of the dorsolateral thalamus, IMAN: lateral magnocellular nucleus of the anterior nidopallium, RA: robust nucleus of the arcopallium. (F) Example of a transition diagram. Nodes denote chunk or syllable labels, numbers denote transition probabilities (in percent, % symbol omitted for clarity), d1/d2, g1/g2 denote different states of syllables d and g respectively based on different sequential contexts, capital letters denote repeat phrases. Edges at each node may not sum to 100% because branches smaller than 5% are omitted for clarity.

mMAN lesions

After several days of sound recording, birds were anesthetized using ketamine and midazolam with isoflurane. mMAN was localized using stereotaxic coordinates and was destroyed bilaterally using current injections of 100uA current for 80s at four sites per hemisphere. After conclusion of the study, birds were deeply anesthetized and perfused with paraformaldehyde. 40um sections were cut on a microtome and processed with either Nissl stain (cresyl violet) or calcitonin gene-related peptide (CGRP) (Hampton, Sakata and Brainard, 2009 [↗](#)). mMAN lesion size was estimated by experienced observers and estimated to be either complete (bird 1, birds 3-6) or greater than 75% (bird 2 and 7) for all birds used in this analysis (Figure 1-figure supplement 1 [↗](#)).

Sequence analysis

Bengalese finch songs were recorded for several days before and after lesions. Singing activity resumed on average 3.8 days post lesions (range 2-6 days, Figure 1-figure supplement 1 [↗](#)), and we analyzed songs starting on average 5.3 days post lesions (range 3-11 days) up to 7.4 days post lesions (range 5-12 days). We analyzed at least two days pre and post lesions, and subsampled the larger of the two datasets pooled over all days such that the pre and post datasets were equal in size for each bird (average of 299 song bouts, range 102-601, for all analyses). All our main analyses were carried out on this dataset. To analyze specifically the persistence of effects after the lesion, Figure 1-figure supplement 2 [↗](#) includes four additional data points (14, 19, 33 and 33 days post lesion for birds 2, 3, 6 and 7 respectively) after additional behavioral manipulation, from which birds typically recover (Warren *et al.*, 2012 [↗](#)).

Annotation

We annotated the entire syllable sequence in an automated way. We first used a subset of the data to generate a training set for classification. Separate classifiers were trained for each bird. In the training data, syllable onsets and offsets were determined by amplitude thresholding. We then performed dimensionality-reduction and unsupervised clustering of the spectrograms of detected syllables using UMAP (McInnes *et al.*, 2018 [↗](#); Sainburg, Thielk and Gentner, 2020 [↗](#)), in order to determine the number of different syllables in an objective way (mean 13, range 11-14). On average, we used 181 files (range 70-505) for creating this training dataset, which was then used to train a deep neural network (TweetyNet) for the annotation of birdsong (Cohen *et al.*, 2022 [↗](#)). TweetyNet was used to segment and annotate all songs, followed by semi-automated hand-checking using custom-written software in MATLAB R2021b to ensure quality annotation. In rare cases, where there was some ambiguity in the assignment of syllable identity during hand-checking, we additionally took into account the sequential context, for example during stereotyped chunks, to assign the most likely label. Thus, any errors in assignment in such cases would have tended to reduce rather than accentuate the magnitude of the lesions' effects on reported sequencing changes.

Simplification of the syllable sequence

We defined song bouts as continuous sequences of syllables separated by at least 2 s of silence. We then simplified the syllable sequence by merging introductory notes into a single introductory state, i.e., a single node 'T' in the transition diagram (Figure 1A,F [↗](#)). Introductory notes were defined as up to three syllable types occurring at the start of bouts which were quieter, and more variable in timing and structure than other syllable types (Rajan and Doupe, 2013 [↗](#); Veit *et al.*, 2021 [↗](#)). Similarly, syllables which were repeated a variable number of at least two times were reduced to a single state (e.g. 'bbbbbbbbb' to 'B', Figure 1A,F [↗](#)) in the syllable sequence, and the number of repetitions within these repeat phrases was analyzed separately (Hampton, Sakata and Brainard, 2009 [↗](#); Zhang *et al.*, 2017 [↗](#); Jaffe and Brainard, 2020 [↗](#)). We only analyzed repeat phrases with variable numbers of syllable repetitions in this way, therefore, we did not consider cases in which a syllable was repeated an exact number of times on each occurrence (occurring with repeat numbers of 2 or 3 in our dataset, see Figure 3A [↗](#)) a variable repeat phrase.

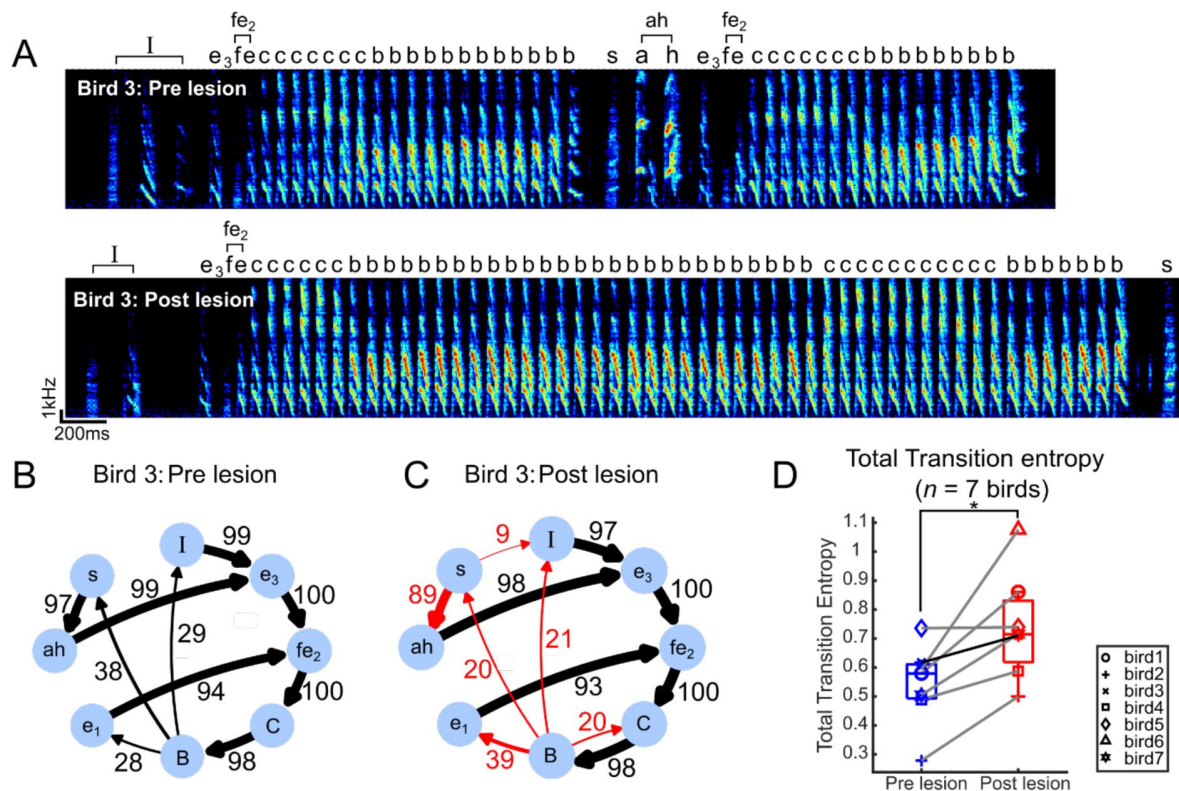


Figure 2

Transition entropy increased after bilateral mMAN lesions (A) Example spectrogram (bird 3) pre lesion (above) and post lesion (below). *I* denotes introductory state, and '*fe2*' and '*ah*' denote chunks, shown as single nodes in the transition diagram. (B) Pre lesion transition diagram, as in 1F. Note that '*fe2CB*' is also a chunk before the lesions but is shown as three separate nodes in order to align with the post lesion diagram in C. (C) Post lesion transition diagram. Arrows in red mark example nodes with relatively high increase in transition entropy, including the introduction of new branches after mMAN lesions. (D) Total transition entropy for 7 birds before and after mMAN lesions (* $p < 0.05$, $n = 7$, Wilcoxon signed-rank test). Example bird from A-C is shown as darker line.

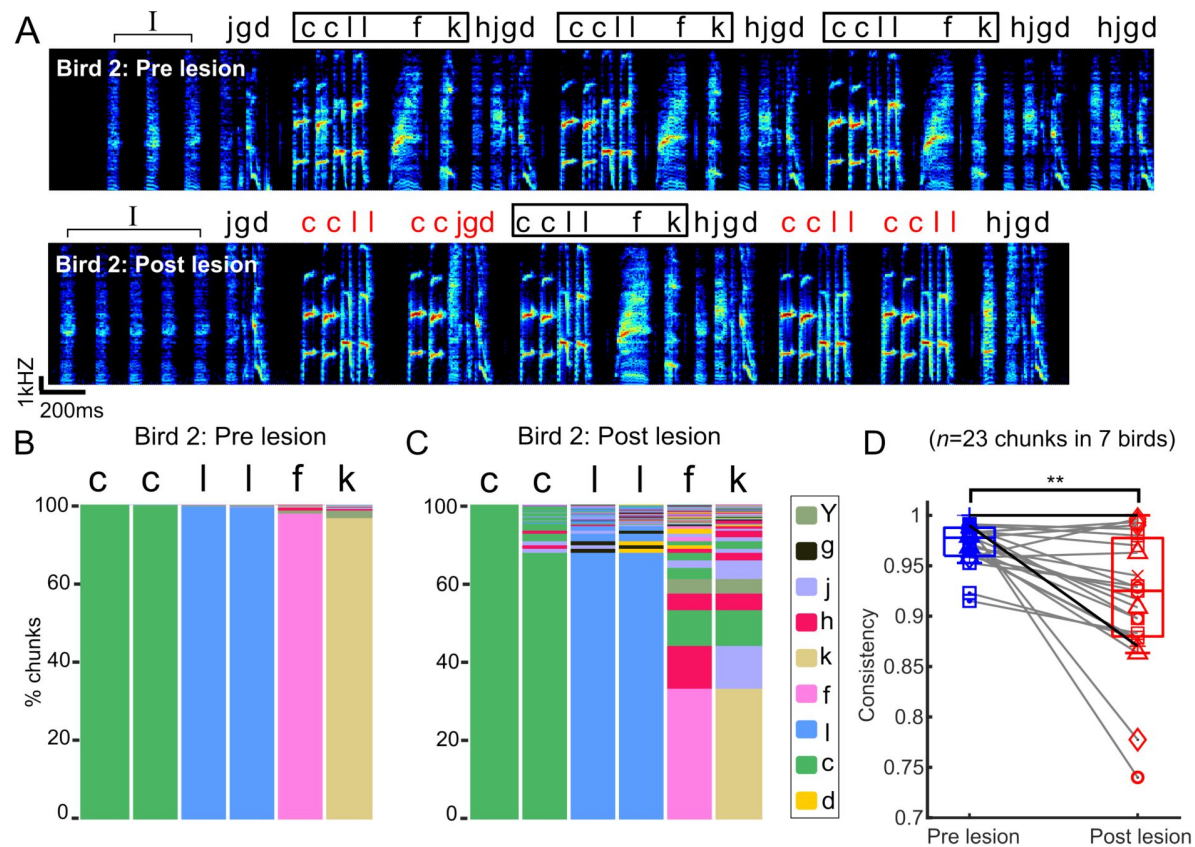


Figure 3

Chunks became more variable after bilateral mMAN lesions. (A) Example spectrogram (bird 2) before and after bilateral mMAN lesions. Atypical chunk sequences are highlighted in red. (B,C) Transitions following the first 'c' of the 'ccll/fk' chunk from (A) before and after mMAN lesions. Different column colors represent different syllables. (D) Chunk consistency before and after bilateral mMAN lesions (** p<0.01, n=23, Wilcoxon signed rank test). Example bird is shown as darker line.

Separating different syllable states based on sequential context

Transition probabilities between syllables in Bengalese finch song can depend not only on the identity of the syllable, but also on the identity of at least one preceding syllable (Jin and Kozhevnikov, 2011; Katahira *et al.*, 2011; Morita *et al.*, 2021). In order to consider this sequential context-dependency of transition probabilities, we followed the procedure of Katahira *et al.* 2011 to classify spectrally similar syllables (i.e., with the same labels) into different ‘states’ based on the identity of the preceding syllable. For example, in **Figure 1A**, the syllable ‘g’ is preceded by ‘T’ or ‘f’ and is followed by ‘h’ or ‘a’. If ‘g’ is preceded by ‘T’ it is always followed by ‘h’ and if it is preceded by ‘f’, it can be followed by ‘h’ or ‘a’. We would therefore like to introduce two separate states g_1 and g_2 in the transition diagram representing this sequential context. The input for this analysis was the syllable sequence with introductory and repeat phrases already condensed into single states. We compared the distribution of transition probabilities from these syllables by themselves (e.g., syllable ‘g’) to the distribution of transition probabilities from these syllables considering the context of one preceding syllable (e.g., syllable ‘f-g’) using the chi-square goodness of fit test. We split the syllable into a new state (e.g., ‘ g_1 ’) if $p < 0.01/n$ where n is the number of comparisons made (Bonferroni correction for multiple comparisons) (Jin, 2009; Katahira *et al.*, 2013). After completing this for all syllables of a given type (e.g., ‘ g_1 ’, ‘ g_2 ’, ‘ g_3 ’), we compared the different output states to each other using chi-square goodness of fit test and merged them back together if $p > 0.01/n$ (mean 15, range 6-21 additional states/bird).

Determining chunks and branchpoints

To determine chunks (Suge and Okanoya, 2010), we used the modified syllable sequence after chi-square analysis. We calculated a syllable-to-syllable transition diagram, while eliminating any syllables occurring with less than a threshold frequency of 0.9% and omitting branches with a probability of less than 5%. From the resultant graph, we merged all non-overlapping linear sequences (which consisted of nodes with only a single input and output branch) into chunks. The transition probabilities within resultant chunks were typically higher than 90% (range 87-100%). Values smaller than 100% for non-branching paths result from the omission of one or several small branches of less than 5%. We then re-calculate the transition probabilities between chunked nodes. We defined branchpoints as the set of single syllables that are retained after this processing and the end syllables of the newly defined chunks. To simplify diagrams for visualization, we eliminated any syllables which occurred less than once per bout and branches with transition probabilities of less than 5%.

Transition entropy

Transition entropy is a measure of uncertainty of sequence at a given syllable (Sakata and Brainard, 2006). With c different outputs from the given syllable ‘a’ and $P(i)$ the probability of the i^{th} outcome, we calculate the entropy H_a as:

$$H_a = - \sum_{i=1}^c P(i) \log P(i)$$

We call this value ‘transition entropy per branchpoint’ in **Figure 3A**.

To determine the overall variability of song before and after mMAN lesions, we calculated total transition entropy TE , over all syllables ‘b’ as:

$$TE = - \sum_{b=1}^n H_b * P(b)$$

where H_b is the transition entropy at 'b' and $P(b)$ is the frequency of syllable 'b' (Chatfield and Lemon, 1970 [↗](#); Katahira *et al.*, 2013 [↗](#)).

History dependence

History dependence is a previously established metric that measures the extent to which the most common transition at a given syllable is influenced by the transition at the last occurrence of this syllable (Warren *et al.*, 2012 [↗](#)). It has been used to characterize instances of apparent sequence variability, where seemingly variable transitions are always strictly alternating. For example, if the possible transitions from syllable 'a' are 'ab' or 'ac' but these strictly alternated ('ab...ac...ab...ac' and so on) then the seemingly variable branchpoint 'a' is perfectly predictable based on its history (Warren *et al.*, 2012 [↗](#)). Such apparent variability should be largely eliminated in our sequence analysis by the introduction of context dependent states (i.e., in this example, the 'a' would be re-labelled as 'a₁' or 'a₂' depending on the context in which it occurs) and identification of chunks. However, if higher-order dependencies in the song determine the order of chunks, we might still expect some variable transitions to be governed by history dependence. If 'ab' is the most frequently occurring transition from 'a', and 'ac' is the collection of all other transitions from 'a', we define history dependence D of 'a' as:

$$D = |P(ab_n|ab_{n-1}) - P(ab_n|ac_{n-1})|$$

where $P(ab_n|ab_{n-1})$ is the conditional probability of 'ab' transition given that 'ab' transition occurred at the previous instance of 'a' and $P(ab_n|ac_{n-1})$ is the conditional probability of 'ab' transition given that 'ac' transition occurred at the previous instance of 'a'.

Chunk consistency

As defined above, a chunk is defined by a single dominant sequence, but may have a small amount of variability across different instances. To quantify the stereotypy of chunks, we used a measure based on sequence consistency previously defined for relatively stereotyped zebra finch songs (Scharff and Nottebohm, 1991 [↗](#)). Across all instances of a given chunk, we identified the syllable sequence that occurred most often as the 'dominant sequence'. We then defined '*n_{dominant}*' as the number of instances of the dominant sequence, and '*n_{other}*' as the number of instances of other sequence variants for the chunk.

We quantified chunk consistency *C* as the proportion of total instances that were the dominant sequence:

$$C = \frac{n_{dominant}}{n_{dominant} + n_{other}}$$

To compare a chunk before and after mMAN lesions, the dominant sequence for the pre lesion chunk was used as a reference, regardless of whether the same sequence qualified as a chunk post lesion. To quantify chunk consistency post lesion, the most dominant sequence post lesion was used (even if that was not the same as the most dominant sequence pre lesion).

Repeat number variability

To study the influence of mMAN on repeat phrases, we examined the distribution of repeat numbers before and after lesions. We quantified the variability of these distributions as their coefficient of variation ($CV = \text{standard deviation}/\text{mean}$).

Analysis of syllable acoustic structure

Average spectrogram

For visual inspection of preservation of syllable types before and after mMAN lesions, we calculated average spectrograms for 200 instances of each syllable type for 7 birds (**Figure 2-figure supplement 1** [↗](#)). Individual spectrograms of each syllable type were rescaled to have matching time axes and then overlaid.

Syllable similarity analysis

We used Sound Analysis Pro (SAP) software (Tchernichovski *et al.*, 2000 [↗](#)) to quantify the similarity in syllable structure for all syllable types per bird (**Figure 2-figure supplement 2** [↗](#)). We used the settings ‘Symmetric’ and ‘Time Courses’ and compared twenty instances of each syllable type.

We first calculated SAP similarity of all syllables to syllables of the same type from two separate control recordings before mMAN lesions. We termed this as ‘Self Similarity’. We next calculated SAP similarity of the same syllable types before and after mMAN lesions. We termed this as ‘Pre vs Post’ similarity. Lastly, we calculated SAP similarity of all syllable types compared to all other syllable types per bird. This was used as a measure of dissimilarity of structure across syllable types. We termed this as ‘Cross Similarity’.

Analysis of fundamental frequency

We used the method previously described by Hampton, Sakata and Brainard, 2009 [↗](#) to calculate fundamental frequency and coefficient of variation (CV) of fundamental frequency for select syllables for each bird (**Figure 2-figure supplement 3** [↗](#)) (Hampton, Sakata and Brainard, 2009 [↗](#)). Briefly, each syllable was visually inspected to have a segment with clearly defined fundamental frequency. We then found the peak of the autocorrelation function (using parabolic interpolation) of the sound waveform in this segment.

Analysis of acoustic features

To further compare syllable phonology before and after the mMAN lesions, we used the SoundSig python package (Elie and Theunissen, 2016 [↗](#)) to calculate a set of acoustic features for all syllables for 7 birds (**Figure 2-figure supplement 4** [↗](#)). The features used were Entropy of spectral envelope (entS), Temporal centroid for the temporal envelope (meanT), first, second and third formants (F1, F2, F3).

Results

Structure of Bengalese finch song

Bengalese finch songs are composed of a fixed number of syllable types, which are arranged into variable sequences following syntactical rules. In this study, we sought to first characterize the different features of Bengalese finch song, before investigating their change following mMAN lesions. For each bird, we first annotated all songs with labels for each of the syllable types, and then defined a transition diagram for analysis according to the following procedure (see Methods):

- 1) We identified introductory notes (Rajan and Doupe, 2013; Veit *et al.*, 2021) and grouped them together into a single introductory state ‘I’ in the transition diagram. This reduced the overall transition entropy of song for 7 birds from 0.95 to 0.85 (Figure 1A). [↗](#)

2013 [Veit et al., 2021](#)) and grouped them together into a single introductory state 'I' in the transition diagram. This reduced the overall transition entropy of song for 7 birds from 0.95 to 0.85 (**Figure 1A**).

2) Repeat phrases, in which the same syllable is repeated a variable number of times, were also reduced to a single state in the transition diagram (indicated by capital letters) and the distribution of repeat numbers was analyzed separately (Hampton, Sakata and Brainard, 2009; Zhang et al., 2017; Jaffe and Brainard, 2020) (**Figure 1D**). Hampton, Sakata and Brainard, 2009 ; Zhang et al., 2017 ; Jaffe and Brainard, 2020) (**Figure 1D**).

3) In Bengalese finch song a single spectrally defined and labelled syllable can occur in different fixed sequences (for example, syllable 'b' in the sequences a-b-c and d-b-e) and therefore reflects distinct states in the song sequence (Wohlgemuth, Sober and Brainard, 2010; Katahira et al., 2011). We followed the procedure of Katahira et al. to determine if any initially identified syllable types corresponded to more than one state based on whether the preceding syllable carried information on transition probabilities (see Methods, Katahira et al., 2013). After identifying and splitting these 'hidden states' (for example b₁, b₂) for each bird, overall transition entropy was further reduced from 0.85 to 0.54 (**Figure 1F**). 'b' in the sequences a-b-c and d-b-e) and therefore reflects distinct states in the song sequence (Wohlgemuth, Sober and Brainard, 2010 ; Katahira et al., 2011). We followed the procedure of Katahira et al. to determine if any initially identified syllable types corresponded to more than one state based on whether the preceding syllable carried information on transition probabilities (see Methods, Katahira et al., 2013). After identifying and splitting these 'hidden states' (for example b₁, b₂) for each bird, overall transition entropy was further reduced from 0.85 to 0.54 (**Figure 1F**).

4) In the new syllable sequence, we identified 'chunks' (Suge and Okanoya, 2010), sub-sequences of high-probability transitions. We defined these as paths of continuous, non-branching sequences in the transition diagram, after omitting branches that occurred with probability < 5% (**Figure 1C**). Lastly, we identified branch points in the transition diagram as the set of single syllables and endpoints of chunks, i.e., all states where the syllable sequence proceeds in a probabilistic manner (**Figure 1B,F**). Suge and Okanoya, 2010), sub-sequences of high-probability transitions. We defined these as paths of continuous, non-branching sequences in the transition diagram, after omitting branches that occurred with probability < 5% (**Figure 1C**). Lastly, we identified branch points in the transition diagram as the set of single syllables and endpoints of chunks, i.e., all states where the syllable sequence proceeds in a probabilistic manner (**Figure 1B,F**).

Transition entropy increased after bilateral mMAN lesions

After bilateral mMAN lesions, all syllable types were preserved (**Figure 2-figure supplement 1** -**4**), but the sequencing of syllables became more variable. In the example in **Figure 2B,C** , we show the transition diagrams of one example bird before and after mMAN lesions. We observed a change in transition probabilities at existing branch points, as well as the appearance of novel branches. We quantified changes in sequencing variability for all syllables using transition entropy (Sakata and Brainard, 2006 ; Katahira et al., 2013). Total transition entropy increased significantly after mMAN lesions, indicating that song syntax overall became more variable (**Figure 2D**). These effects appeared as soon as the bird resumed singing after the lesions and did not recover over the time for which the birds were observed (**Figure 1-figure supplement 2**).

Figure 2-figure supplement 5A shows the corresponding changes at all individual branch points. To investigate the source of differences in the magnitude of changes across branch points, we considered the history-dependence of each of these branchpoints. History-dependence captures apparent variability governed by long-range dependencies in the sequence and history-

dependent transitions previously have been found to be difficult to modify in a sequence modification training protocol (Warren *et al.*, 2012). They might therefore be less affected by lesions as well. Alternatively, we might expect lesion effects to be stronger for these transitions, if mMAN contributes specifically to long-range dependencies. Consistent with the first possibility, we observed that there was a nonsignificant trend toward larger changes after mMAN lesion for transitions with low history dependence. (Figure 2-figure supplement 5B).

Chunks became more variable after bilateral mMAN lesions

Bengalese finch songs contain short, relatively stereotyped sequences of syllables, here called ‘chunks’ (Okanoya and Yamaguchi, 1997; Suge and Okanoya, 2010; Isola, Vochin and Sakata, 2020; Veit *et al.*, 2021). We speculated that transitions within chunks might be differently affected by mMAN lesions, if for example the relatively fixed sequences within chunks are determined within premotor song nucleus HVC and inputs to HVC are only relevant at variable branch points. We found that the relatively fixed sequences within chunks were altered after mMAN lesions. An example bird (Figure 3A-C) exhibited one prominent chunk which was occasionally observed to ‘break’ after lesions, i.e., branching within the chunk increased to a degree where it would no longer be considered a chunk after the lesions. The branching increased most noticeably in the ‘l-f’ transition (Figure 3C), which is characterized by a longer gap than other transitions within the chunk (Figure 3A). We therefore wondered whether sequencing changes were related to gap durations and found that gap durations within chunks were not significantly correlated with the increase in transition entropy at the corresponding transitions (Figure 3-figure supplement 1A). Overall, we found that changes within chunks were approximately the same magnitude as changes at branch points (Figure 3-figure supplement 1B).

We quantified changes in within-chunk transitions using sequence consistency (Scharff and Nottebohm, 1991), a measure previously used to describe the relatively consistent sequence of zebra finch song. Sequence consistency measures how consistently the most probable sequence is followed (see Methods). Sequence consistency within chunks significantly decreased across 23 chunks from all birds, indicating that this aspect of song structure was affected by mMAN lesions (Figure 3D).

Repeat numbers became more variable after mMAN lesions

We next tested how mMAN lesions affected repeat phrases. Repeat phrases might be governed by separate neural mechanisms than branch points (Fujimoto, Hasegawa and Watanabe, 2011; Jin and Kozhevnikov, 2011; Wittenbach *et al.*, 2015; Zhang *et al.*, 2017). In our dataset of 7 birds, only 5 birds had songs which contained repeat phrases.

In the example bird in Figure 4A,B, the average repeat number for the indicated syllable increased after mMAN lesions. Across birds, the mean repeat number did not change consistently, although individual birds could exhibit quite dramatic effects on repeat number, (Figure 4B,C; $p > 0.05$, $n = 6$, Wilcoxon signed rank test). The distributions of repeat numbers became wider and the coefficient of variation for repeat numbers increased significantly after mMAN lesions (Figure 4D; * $p < 0.05$, $n = 6$, Wilcoxon signed rank test).

Discussion

Bird songs are vocal sequences composed of individual syllables which appear in sequential order dictated by syntactic rules, but the neuronal control of song syntax is largely unknown. Previous studies in zebra finches showed that mMAN lesions in adult birds had minimal effects on song production (Foster and Bottjer, 2001; Horita, Wada and Jarvis, 2008; Ali *et al.*, 2013). We

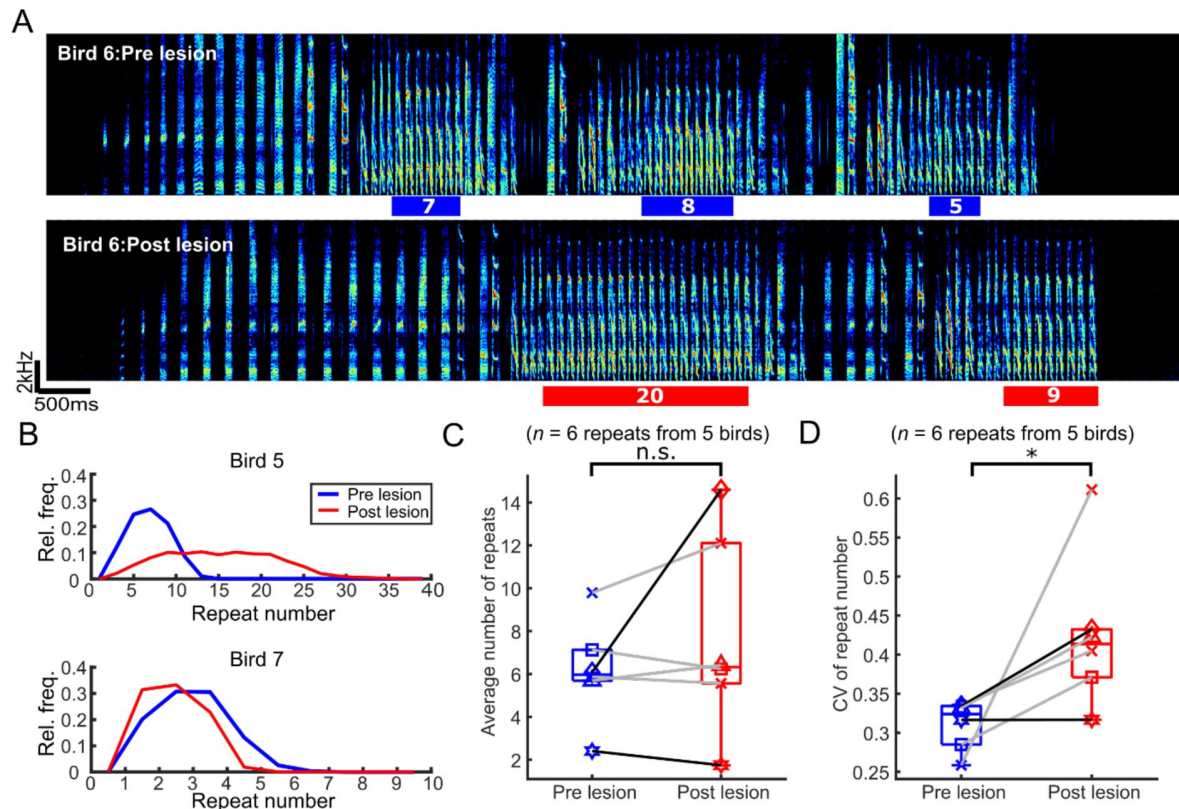


Figure 4

The number of syllables per repeat phrase (repeat number) became more variable after bilateral mMAN lesions. (A) Example spectrogram (bird 6) highlighting one repeat syllable before (blue) and after (red) mMAN lesions. (B) Repeat numbers for two additional example birds before (blue) and after (red) mMAN lesions. (C) Average repeat numbers before and after mMAN lesions for all repeat phrases. (D) Coefficient of variation for distribution of repeat numbers before and after mMAN lesions for all repeat phrases. Example birds from B are shown as darker lines.

speculated that the modest effect of lesions in adult zebra finches might be due to the relatively stereotyped nature of zebra finch song. Consistent with this possibility, our results demonstrate that lesions of pallial nucleus mMAN affect the more complex and variable song syntax in Bengalese finches. We showed that while mMAN lesions did not abolish the production of song itself or grossly disrupt the acoustic structure of individual syllables, they increased the variability of syllable sequencing; the transition probabilities between syllables became less predictable (increased transition entropy), new branches were introduced into previously stereotyped chunks, and the number of repetitions of repeated syllables (repeat number) became more variable. These results indicate that mMAN is involved in controlling aspects of song syntax in a songbird species that exhibits complex syllable sequencing and suggest that mMAN might play a similar or prominent role in additional species with even greater song complexity, such as canaries or nightingales.

Song in passerine birds relies on a network of recurrently connected nuclei, with different pathways for song production and learning. The song motor pathway consists of primary motor nucleus RA, which projects directly to vocal motor nuclei in the brainstem, and is involved in the moment-by-moment control of the structure of individual syllables (Vu, Mazurek and Kuo, 1994; Sober, Wohlgemuth and Brainard, 2008; Miller, Cheung and Brainard, 2017). Neurons in premotor song nucleus HVC that project to RA have been shown to encode syllable identity and contribute to syllable timing (Simpson and Vicario, 1990; Vu, Mazurek and Kuo, 1994; Hahnloser, Kozhevnikov and Fee, 2002; Long and Fee, 2008; Aronov *et al.*, 2011; Ölveczky *et al.*, 2011; Zhang *et al.*, 2017). Consistent with a possible role of HVC in also controlling variable syllable sequencing, cooling HVC influences syntax in Bengalese finches (Zhang *et al.*, 2017). However, lesions of one input to HVC – the nucleus Nif (*nucleus interfascialis of the anterior nidopallium*) – can also influence syllable sequencing in the Bengalese finch by making the sequence more stereotyped (Hosino and Okanoya, 2000; Cardin, Raksin and Schmidt, 2005; Otchy *et al.*, 2015) raising the question of how much of the control of syntax relies on the internal dynamics of HVC, versus activity relayed to HVC by its inputs (Okanoya and Yamaguchi, 1997).

In this study, we were especially interested in the possibility that mMAN's inputs to HVC might play an analogous role in modulating syllable sequencing to the role of lMAN's inputs to RA in modulating syllable acoustic structure. In adult birds, lMAN's inputs to RA have been shown to contribute variability to the pitch of individual syllables (Kao, Doupe and Brainard, 2005; Kao and Brainard, 2006; Stepanek and Doupe, 2010), regulate changes to pitch variability between different social contexts (Kao and Brainard, 2006; Hampton, Sakata and Brainard, 2009), and bias pitch in the direction of improved performance during learning experiments (Andalman and Fee, 2009; Warren *et al.*, 2011; Tian and Brainard, 2017). The variability contributed by lMAN may facilitate motor exploration which forms the basis of pitch adaptations during learning (Andalman and Fee, 2009; Sober and Brainard, 2012; Dhawale, Smith and Ölveczky, 2017). These effects are mediated by top-down influences of lMAN on RA via a recurrent loop that passes through the songbird basal ganglia (Kao, Doupe and Brainard, 2005; Andalman and Fee, 2009; Kojima *et al.*, 2018; Tian, Warren and Brainard, 2022), and are consistent with a role of RA in controlling aspects of syllable acoustic structure. In contrast to these effects on syllable structure, lesions of lMAN in adults have not been found to appreciably influence syllable sequencing (Hampton, Sakata and Brainard, 2009). However, because mMAN forms the output of a parallel recurrent loop that passes through the basal ganglia and projects to HVC, we hypothesized that analogous modulation of syllable sequencing during production and learning of song might be mediated by mMAN. Our finding that lesions of mMAN alter multiple aspects of syllable sequencing are consistent with such an influence on song syntax.

The finding that mMAN can influence song syntax raises the question of whether and how this pathway modulates syllable sequencing in the service of motor exploration, contextual modulation, and learning, in a fashion that is analogous to lMAN's modulation of syllable acoustic

structure. That mMAN lesions consistently altered sequence variability indicates a possible role in regulating variability for motor exploration. However, whereas IMAN lesions consistently *decrease* the variability of acoustic structure, we found that mMAN lesions generally *increase* the variability of syllable sequencing, suggesting that top-down contributions to the control of syntax may be different or more distributed than for a syllable feature such as pitch. One possibility is that several inputs to HVC could work synergistically to modulate the syntactical structure of Bengalese finch songs. Consistent with this view, a prior study in Bengalese finches found that lesions of Nif decreased sequence variability (Hosino and Okanoya, 2000 [DOI](#)), suggesting that Nif and mMAN could be playing opposing roles in regulating variability. Our findings also raise the question of whether the medial anterior forebrain pathway including mMAN, plays a similar role in learning to change syllable sequences during development (Foster and Bottjer, 2001 [DOI](#); Lipkind *et al.*, 2013 [DOI](#)) or adulthood (Warren *et al.*, 2012 [DOI](#); Veit *et al.*, 2021 [DOI](#)). In our study, we only recorded song sequencing of male Bengalese finches singing in isolation. Social context, such as female-directed song, can also change song sequencing (Hampton, Sakata and Brainard, 2009 [DOI](#); Chen, Matheson and Sakata, 2016 [DOI](#)). It would be interesting to test whether mMAN plays a role in the social context-modulated changes in sequencing (Jarvis *et al.*, 1998 [DOI](#)), similar to how IMAN contributes to social context-modulated changes in syllable structure (Sakata, Hampton and Brainard, 2008 [DOI](#)). Further studies with careful manipulations of the basal ganglia-mMAN-HVC circuit will be required to elucidate the mechanism with which this circuit affects sequencing in Bengalese finch song and contributes to these behaviors.

Bengalese finch syllable sequences contain structure at multiple levels, mainly the organization of individual stereotyped chunks, resembling the stereotyped motifs of zebra finch song, and the variable branch points between chunks, repeat phrases, and other individual syllables. One possibility to reconcile findings that both dynamics within HVC and its recurrent inputs affect sequencing could be that the sequence within chunks is encoded predominantly in the synaptic connections of HVC neurons (Jin, 2009 [DOI](#)), while the branching transitions found in Bengalese finch song are influenced by input from other nuclei (Hosino and Okanoya, 2000 [DOI](#)). We here show that mMAN lesions affect both transition entropy at previously variable branch points and transitions within previously stereotyped chunks, indicating that such a simple split of function between mMAN and HVC is not the case.

Besides these, other parts of the song circuit have been shown to affect syllable sequencing. The basal ganglia are involved in selection and sequencing of motor elements in mammals, and manipulations of songbird basal ganglia Area X have previously been shown to affect syllable repetitions in zebra finches and Bengalese finches (Kobayashi, Uno and Okanoya, 2001 [DOI](#); Kubikova *et al.*, 2014 [DOI](#); Tanaka *et al.*, 2016 [DOI](#); Xiao *et al.*, 2021 [DOI](#)). These effects might be partly mediated via connections through mMAN to HVC (Kubikova, Turner and Jarvis, 2007 [DOI](#)). Additionally, the control of syllable sequencing must include bilateral coordination of the motor pathways in the two hemispheres (Schmidt, Ashmore and Vu, 2004 [DOI](#); Wang *et al.*, 2008 [DOI](#)). As the two hemispheres in birds are not connected by a corpus callosum, the only pathways mediating interhemispheric synchronization of the two HVCs are by bilateral connections through thalamic nuclei. One such pathway involves a projection from thalamic nucleus Uva to HVC (Coleman and Vu, 2005 [DOI](#); Hamaguchi, Tanaka and Mooney, 2016 [DOI](#); Moll *et al.*, 2023 [DOI](#)), which is required for song production and contributes to syllable initiation and timing. Another is through thalamic nucleus DMP, which projects bilaterally through mMAN to HVC (Schmidt, Ashmore and Vu, 2004 [DOI](#); Seki and Okanoya, 2008 [DOI](#)). mMAN might therefore be well situated to coordinate syllable sequencing in the two hemispheres. Future studies will need to determine whether the effects we observed are the results of removing neuronal processing within mMAN itself, or whether mMAN is conveying signals to HVC from other brain areas, including the contralateral hemisphere, which contribute to the control of song syntax.

Acknowledgements

We thank Jacqueline Göbl, Lioba Fortkord, Yarden Cohen, Tim Gardner, Felix Moll for helpful discussions and comments on earlier versions of this manuscript. We thank David Nicholson and Peter Pilz for help with technical questions. Michael Brainard was supported by the Howard Hughes Medical Institute. Lena Veit was supported as a Howard Hughes Medical Institute Fellow of the Life Sciences Research Foundation and by a Daimler Benz postdoctoral fellowship. Avani Koparkar was supported by the Studienstiftung des deutschen Volkes and International Max Planck Research School for Mechanisms of Mental Function and Dysfunction (IMPRS-MMFD). Jonathan Charlesworth and Timothy Warren were supported by the National Science Foundation (NSF) Graduate Research Fellowship. We acknowledge support from the Open Access Publication Fund of the University of Tübingen.

Author contributions

T.L.W., J.C., S.S. and L.V. gathered the data. A.K. and L.V. analyzed the data, A.K., L.V. and

M.S.B. wrote the paper.

Competing interests

The authors declare no competing interests.

Supplementary Figs

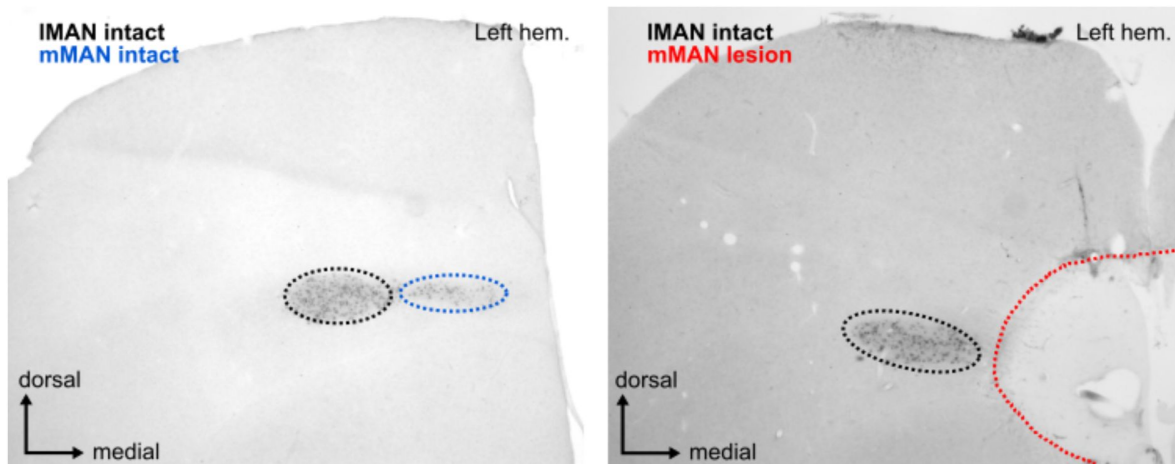


Figure 1 - figure supplement 1

Image of calcitonin gene-related peptide (CGRP)-stained frontal section (left) control and (right) bird 5. CGRP labels cells in both IMAN (seen in black to the left of the lesion) and mMAN (blue, intact; red, completely destroyed).

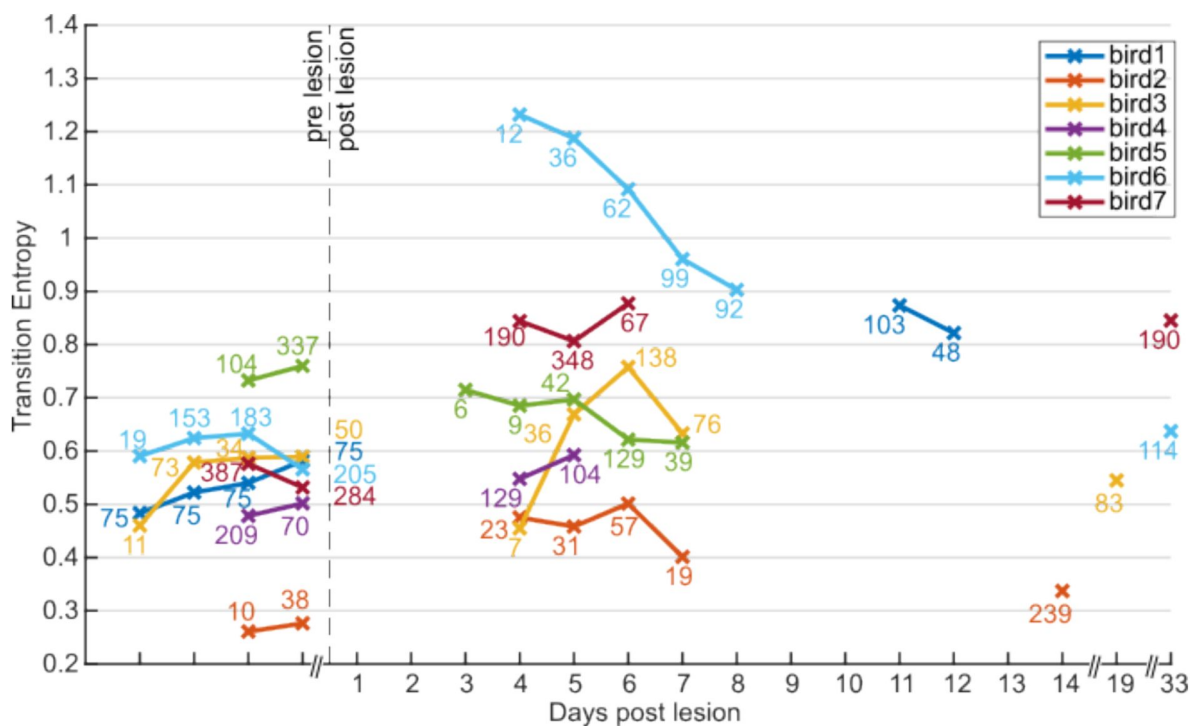


Figure 1 - figure supplement 2

Transition entropy after mMAN lesions remains elevated over several days. Transition entropy did not change in a consistent way after recovery from the lesion for the followed time period. Numbers near the datapoints indicate number of song bouts recorded and analyzed for that day. While the persistence of the effects we observed is longer than transient effects such as those following Nif lesion in zebra finches (~ 2 days by Otchy *et al.*, 2015) we cannot rule out either recovery or further deterioration following lesions on much longer time scales, such as those reported by Kubikova *et al.*, 2007 (X lesions, 6 months) (Kubikova, Turner and Jarvis, 2007). Additional non-connected datapoints at 14, 19, 33 and 33 days post lesion were gathered after additional behavior manipulation and were added here to analyze specifically the persistence of effects after the lesion. For three of these birds, transition entropy remained elevated above the baseline values for 14, 33 and 33 days, respectively.

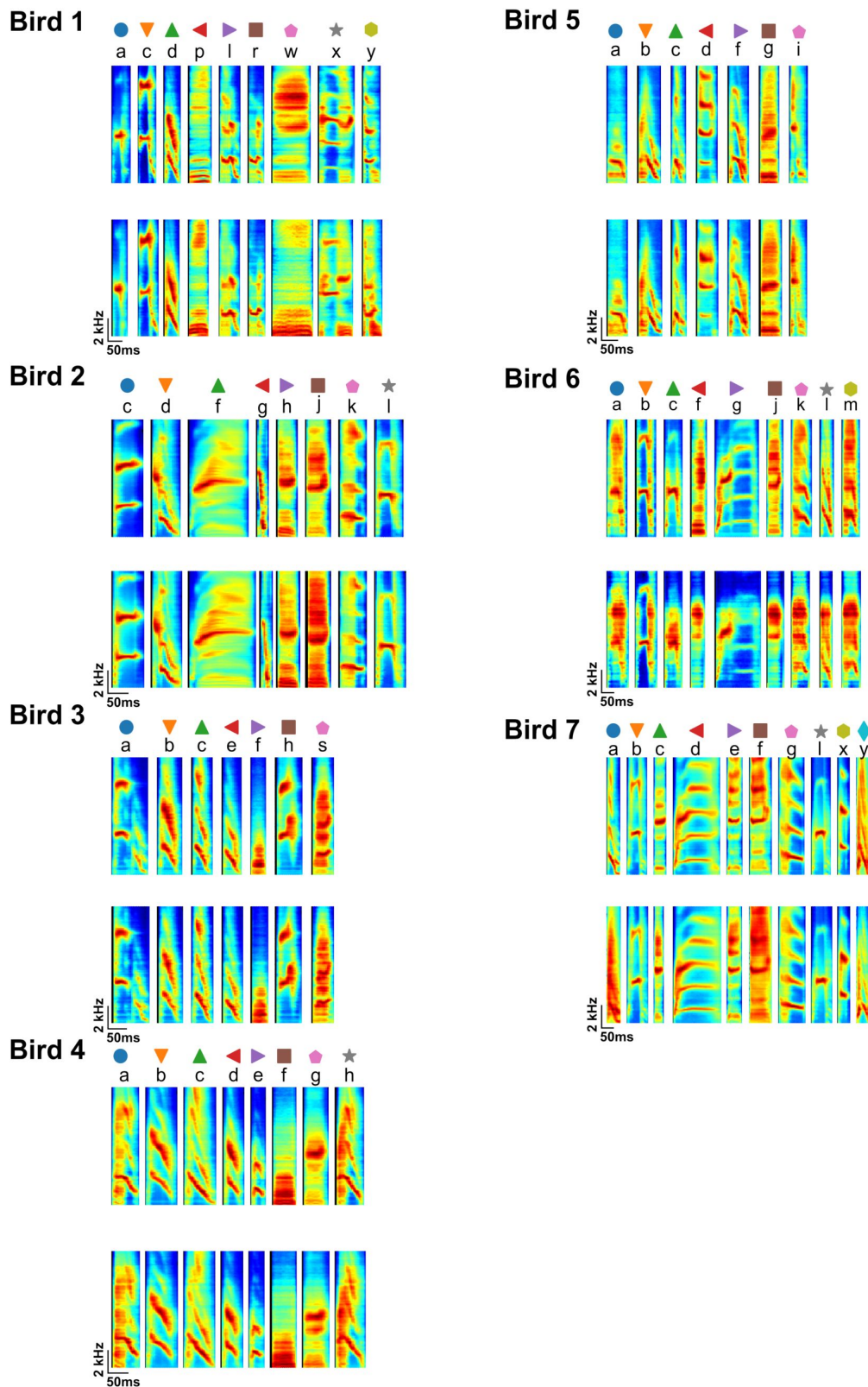


Figure 2 - figure supplement 1

Average spectrograms of 200 instances of all syllable types for all birds before and after mMAN lesions. Letters over the spectrograms indicate syllable labels and symbols over the spectrograms correspond to the symbols used in **Fig. 2** – figure supplement 4.

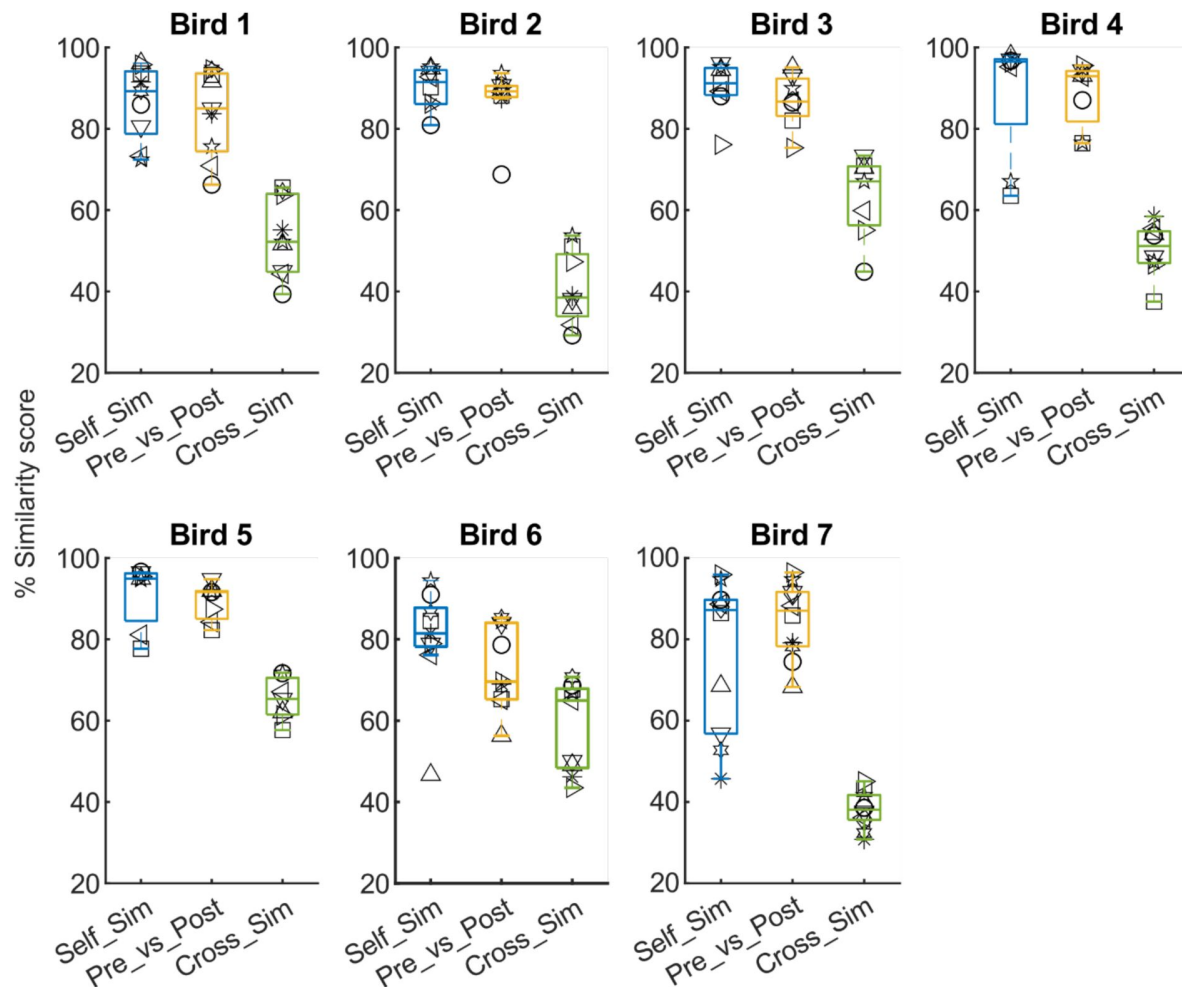


Figure 2 - figure supplement 2

Syllable similarity calculated using Sound Analysis Pro (SAP). 'Self Similarity' = Similarity comparison of syllables before mMAN lesions to syllables of the same type, taken from two separate control recordings before the lesions, 'Pre vs Post' = Similarity comparison of the same syllable types before and after mMAN lesions, 'Cross Similarity' = Similarity comparison of each syllable type to other syllable types. For Birds 1-2 and 4-7, 'Self Similarity' was not significantly different from 'Pre vs Post' Similarity ($p > 0.05$, Wilcoxon sign rank test), while for Bird 3, there was a significant difference ($p = 0.03$, Wilcoxon sign rank test). For all birds 'Pre vs Post' was significantly different from 'Cross Similarity' ($p < 0.05$, Wilcoxon sign rank test). On average, 'Pre vs Post' was 4.8 % less than 'Self Similarity' (range 0.2%-14%) while 'Cross Similarity' was 40% less than 'Self Similarity' (range 20.2%-56.3%). These measures confirm the qualitative impression from [Figure 2](#) - figure supplement 1 that for most birds and syllables there were no greater changes to syllable structure following lesions than was present across control recordings, and that pre-post similarity remained higher than cross-similarity, i.e. syllables remained clearly identifiable.

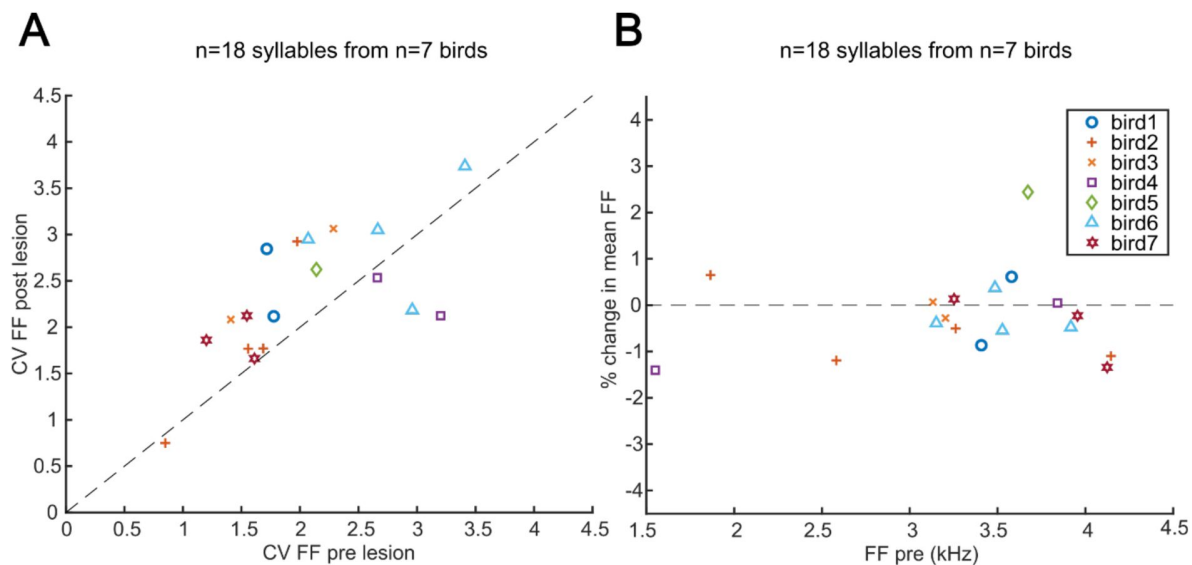


Figure 2 - figure supplement 3

(A) CV of fundamental frequency (FF) of select syllables before and after mMAN lesions. In the Bengalese finch and zebra finch, lesions of IMAN, which sits immediately lateral to mMAN, cause a consistent reduction in the coefficient of variation (CV) of fundamental frequency across repeated renditions of a given syllable (Sakata, Hampton, Brainard 2008, Andalman, Fee 2009, Warren et al. 2011). We therefore supposed that unintended damage to IMAN or its projections to RA might have resulted in a reduction in the CV of syllables following mMAN lesions. Instead we saw a modest increase in the CV of fundamental frequency ($p < 0.05$, Wilcoxon sign rank test; mean across birds of +20%; range -19 to +43%). These data suggest that it is unlikely that changes to syllable structure might have arisen due to accidental damage to IMAN. (B) Percent change in mean fundamental frequency after mMAN lesions vs mean fundamental frequency before mMAN lesions.

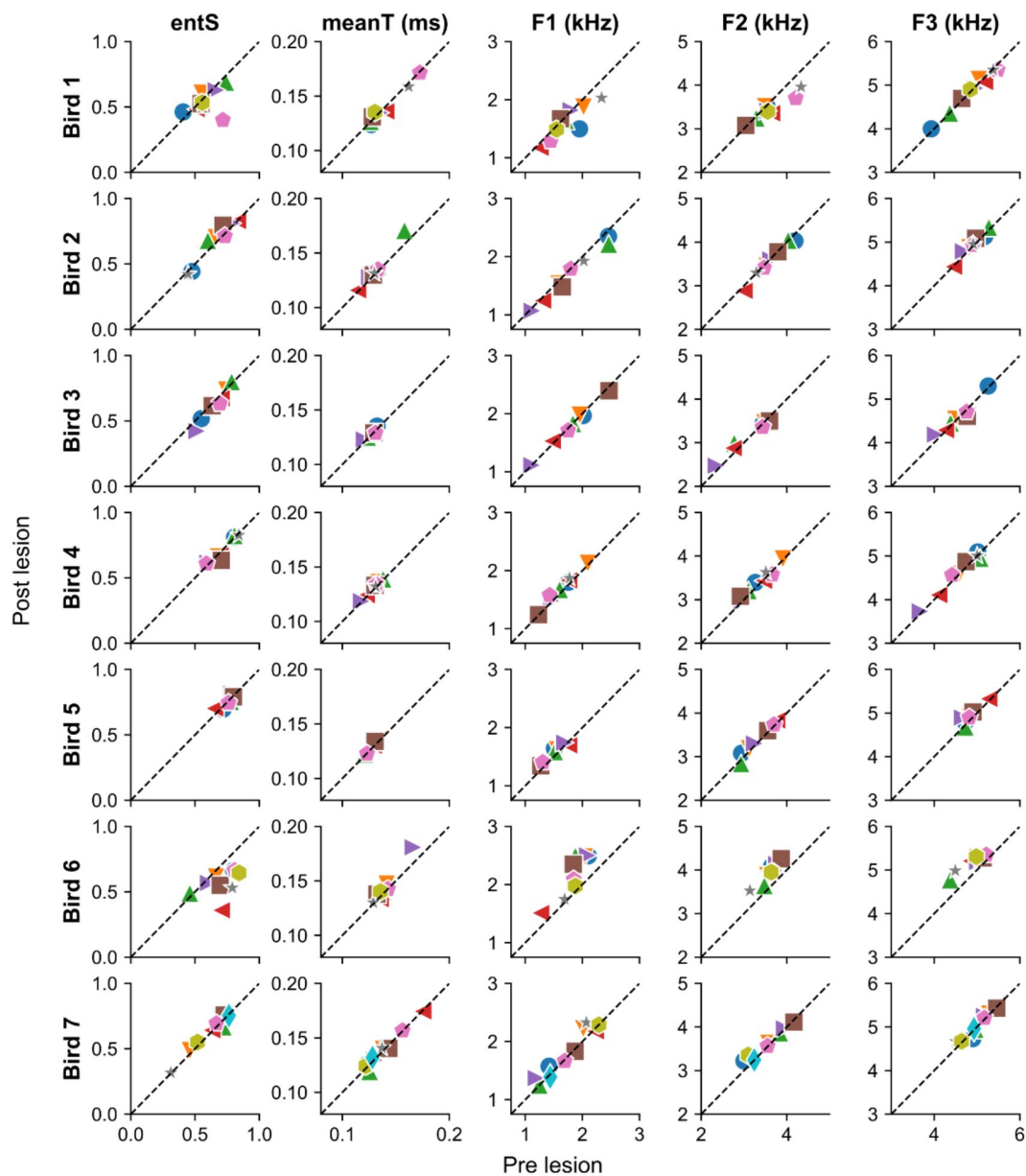


Figure 2 – figure supplement 4

Selected acoustic features for all syllables in all birds before and after mMAN lesions. Different colors represent different syllable types per bird. 'entS' = Entropy of spectral envelope, 'meanT' = Temporal centroid for temporal envelope, 'F1' = First formant, 'F2' = Second formant, 'F3' = Third formant. Acoustic features generally showed little change between pre and post lesion songs. They highlight as relative outliers the same individual examples that stand out in the average spectrograms in [Figure 2](#) – figure supplement 1.

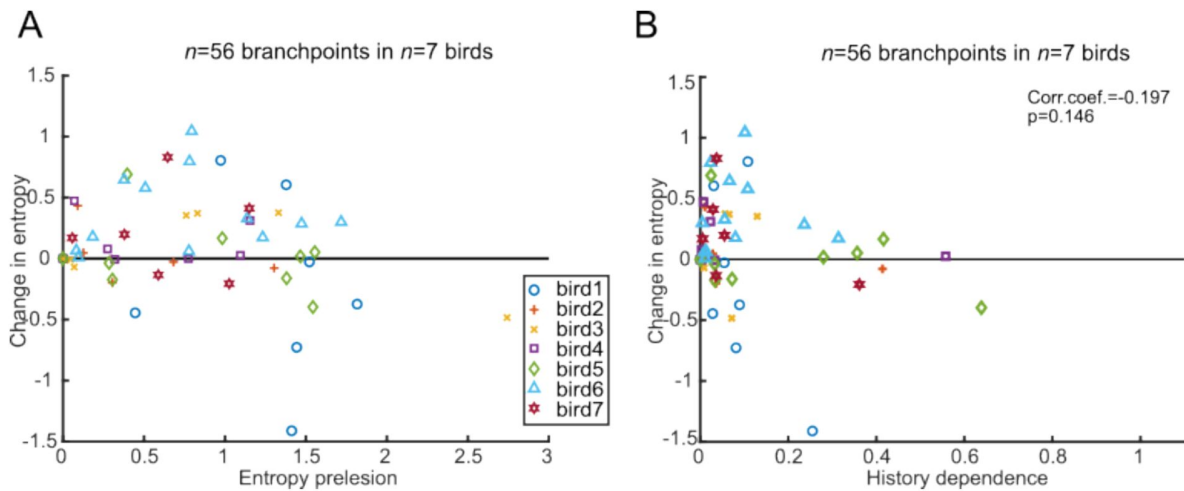


Figure 2 - figure supplement 5

Change in transition entropy for branchpoints (A) Transition entropy for all individual branchpoints after mMAN lesions. (B) Branchpoints with low history dependence before mMAN lesions showed a nonsignificant trend towards higher changes in transition entropy.

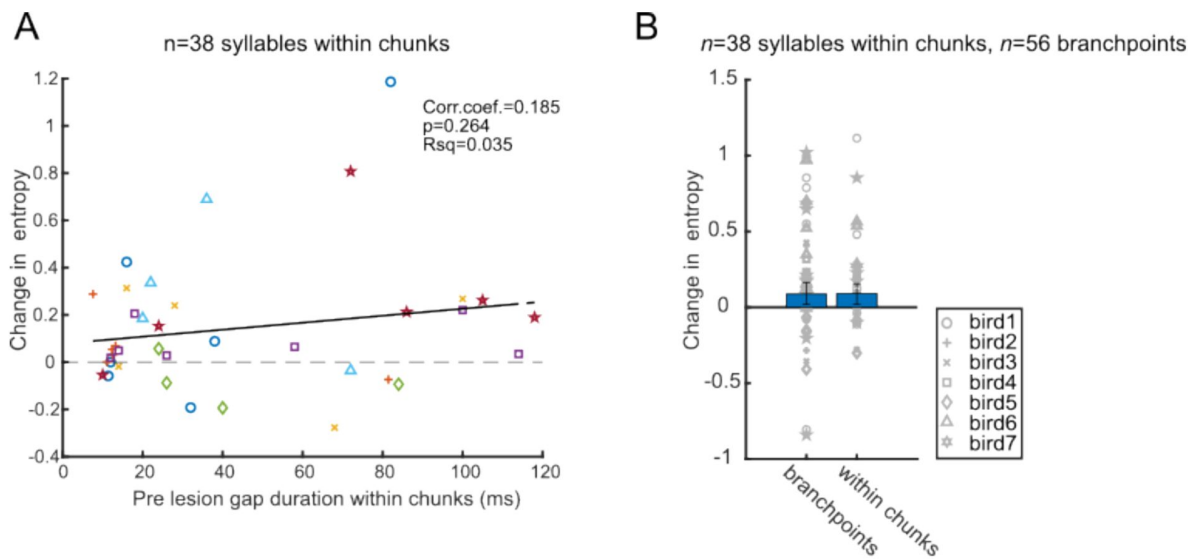


Figure 3 - figure supplement 1

Change in transition entropy for transitions within chunks vs. branch points (A) Duration of gaps within chunks before mMAN lesions was not significantly correlated with entropy change at the corresponding transition ($p > 0.05$, Wilcoxon rank sum test). (B) Change in transition entropy was not significantly different for transitions within chunks and at branchpoints ($p > 0.05$, Wilcoxon rank sum test).

References

- Aldridge J. W., Berridge K. C (2003) **'Basal Ganglia Neural Coding of Natural Action Sequences', in The Basal Ganglia VI** *Advances in Behavioral Biology* :279–287 https://doi.org/10.1007/978-1-4615-0179-4_28
- Ali F., et al. (2013) **The Basal Ganglia Is Necessary for Learning Spectral, but Not Temporal, Features of Birdsong** *Neuron* **80**:494–506 <https://doi.org/10.1016/j.neuron.2013.07.049>
- Andalman A. S., Fee M. S (2009) **A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors** *Proceedings of the National Academy of Sciences* **106**:12518–12523 <https://doi.org/10.1073/pnas.0903214106>
- Aronov D., et al. (2011) **Two Distinct Modes of Forebrain Circuit Dynamics Underlie Temporal Patterning in the Vocalizations of Young Songbirds** *Journal of Neuroscience* **31**:16353–16368 <https://doi.org/10.1523/JNEUROSCI.3009-11.2011>
- Berwick R. C., et al. (2011) **Songs to syntax: the linguistics of birdsong** *Trends in Cognitive Sciences* **15**:113–21 <https://doi.org/10.1016/j.tics.2011.01.002>
- Cardin J. A., Raksin J. N., Schmidt M. F (2005) **Sensorimotor Nucleus Nif Is Necessary for Auditory Processing But Not Vocal Motor Output in the Avian Song System** *Journal of Neurophysiology* **93**:2157–2166 <https://doi.org/10.1152/jn.01001.2004>
- Chatfield C., Lemon R. E (1970) **Analysing sequences of behavioural events** *Journal of Theoretical Biology* **29**:427–445 [https://doi.org/10.1016/0022-5193\(70\)90107-4](https://doi.org/10.1016/0022-5193(70)90107-4)
- Chen Y., Matheson L. E., Sakata J. T (2016) **Mechanisms underlying the social enhancement of vocal learning in songbirds** *Proceedings of the National Academy of Sciences* **113**:6641–6646 <https://doi.org/10.1073/pnas.1522306113>
- Cohen Y., et al. (2020) **Hidden neural states underlie canary song syntax** *Nature* **582**:539–544 <https://doi.org/10.1038/s41586-020-2397-3>
- Cohen Y., et al. (2022) **Automated annotation of birdsong with a neural network that segments spectrograms** *eLife* **11** <https://doi.org/10.7554/eLife.63853>
- Coleman M. J., Vu E. T (2005) **Recovery of impaired songs following unilateral but not bilateral lesions of nucleus uvaeformis of adult zebra finches** *Journal of Neurobiology* **63**:70–89 <https://doi.org/10.1002/neu.20122>
- Costalunga G., et al. (2023) **Wild nightingales flexibly match whistle pitch in real time** *Current Biology* :1–10 <https://doi.org/10.1016/j.cub.2023.06.044>
- Dhawale A. K., Smith M. A., Ölveczky B. P (2017) **The Role of Variability in Motor Learning** *Annual Review of Neuroscience* **40**:479–498 <https://doi.org/10.1146/annurev-neuro-072116-031548>

- Elie J. E., Theunissen F. E (2016) **The vocal repertoire of the domesticated zebra finch: a data-driven approach to decipher the information-bearing acoustic features of communication signals** *Animal Cognition* **19**:285–315 <https://doi.org/10.1007/s10071-015-0933-6>
- Foster E. F., Bottjer S. W (2001) **Lesions of a telencephalic nucleus in male zebra finches: Influences on vocal behavior in juveniles and adults** *Journal of Neurobiology* **46**:142–165 [https://doi.org/10.1002/1097-4695\(20010205\)46:2<142::AID-NEU60>3.0.CO;2-R](https://doi.org/10.1002/1097-4695(20010205)46:2<142::AID-NEU60>3.0.CO;2-R)
- Fujimoto H., Hasegawa T., Watanabe D (2011) **Neural coding of syntactic structure in learned vocalizations in the songbird** *Journal of Neuroscience* **31**:10023–10033 <https://doi.org/10.1523/JNEUROSCI.1606-11.2011>
- Hahnloser R. H. R., Kozhevnikov A. A., Fee M. S (2002) **An ultra-sparse code underlies the generation of neural sequences in a songbird** *Nature* **419**:65–70 <https://doi.org/10.1038/nature00974>
- Hamaguchi K., Tanaka M., Mooney R (2016) **A Distributed Recurrent Network Contributes to Temporally Precise Vocalizations** *Neuron* **91**:680–693 <https://doi.org/10.1016/j.neuron.2016.06.019>
- Hampton C. M., Sakata J. T., Brainard M. S (2009) **An avian basal ganglia-forebrain circuit contributes differentially to syllable versus sequence variability of adult Bengalese finch song** *Journal of Neurophysiology* **101**:3235–3245 <https://doi.org/10.1152/jn.91089.2008>
- Honda E., Okanoya K (1999) **Acoustical and Syntactical Comparisons between Songs of the White-backed Munia (*Lonchura striata*) and Its Domesticated Strain, the Bengalese Finch (*Lonchura striata* var. *domestica*)** *Zoological Science* **16**:319–326 <https://doi.org/10.2108/zsj.16.319>
- Horita H., Wada K., Jarvis E. D (2008) **Early onset of deafening-induced song deterioration and differential requirements of the pallial-basal ganglia vocal pathway** *European Journal of Neuroscience* **28**:2519–2532 <https://doi.org/10.1111/j.1460-9568.2008.06535.x>
- Hosino T., Okanoya K (2000) **Lesion of a higher-order song nucleus disrupts phrase level complexity in Bengalese finches** *NeuroReport* **11**:2091–5 <https://doi.org/10.1097/00001756-200007140-00007>
- Hultsch H., et al. (2004) **Long-term persistence of song performance rules in nightingales (*Luscinia megarhynchos*): a longitudinal field study on repertoire size and composition** *Behaviour* **141**:371–390 <https://doi.org/10.1163/156853904322981914>
- Isola G. R., Vochin A., Sakata J. T (2020) **Manipulations of inhibition in cortical circuitry differentially affect spectral and temporal features of Bengalese finch song** *Journal of Neurophysiology* **123**:815–830 <https://doi.org/10.1152/jn.00142.2019>
- Ivanitskii V. V., Marova I. M (2022) **The Syntactic Organization of Bird Song** *Biology Bulletin* **49**:1158–1170 <https://doi.org/10.1134/S1062359022080076>
- Jaffe P. I., Brainard M. S (2020) **Acetylcholine acts on songbird premotor circuitry to invigorate vocal output** *eLife* **9** <https://doi.org/10.7554/eLife.53288>
- Jarvis E. D., et al. (1998) **For Whom The Bird Sings: context-dependent gene expression** *Neuron* **21**:775–788 [https://doi.org/10.1016/S0896-6273\(00\)80594-2](https://doi.org/10.1016/S0896-6273(00)80594-2)

- Jin D. Z (2009) **Generating variable birdsong syllable sequences with branching chain networks in avian premotor nucleus HVC** *Physical Review E* **80** <https://doi.org/10.1103/PhysRevE.80.051902>
- Jin D. Z., Kozhevnikov A. A (2011) **A Compact Statistical Model of the Song Syntax in Bengalese Finch** *PLOS Computational Biology* **7** <https://doi.org/10.1371/journal.pcbi.1001108>
- Kao M. H., Brainard M. S (2006) **Lesions of an Avian Basal Ganglia Circuit Prevent Context-Dependent Changes to Song Variability** *Journal of Neurophysiology* **96**:1441–1455 <https://doi.org/10.1152/jn.01138.2005>
- Kao M. H., Doupe A. J., Brainard M. S (2005) **Contribution of an avian basal ganglia-forebrain circuit to real-time modulation of song** *Nature* **433**:638–643 <https://doi.org/10.1038/nature03127>
- Katahira K., et al. (2011) **Complex Sequencing Rules of Birdsong Can be Explained by Simple Hidden Markov Processes** *PLOS ONE* **6** <https://doi.org/10.1371/journal.pone.0024516>
- Katahira K., et al. (2013) **A simple explanation for the evolution of complex song syntax in Bengalese finches** *Biology Letters* **9**:3–7 <https://doi.org/10.1098/rsbl.2013.0842>
- Kobayashi K., Uno H., Okanoya K (2001) **Partial lesions in the anterior forebrain pathway affect song production in adult Bengalese finches** *NeuroReport* **12**:353–358 <https://doi.org/10.1097/00001756-200102120-00034>
- Kojima S., et al. (2018) **The Avian Basal Ganglia Are a Source of Rapid Behavioral Variation That Enables Vocal Motor Exploration** *Journal of Neuroscience* **38**:9635–9647 <https://doi.org/10.1523/JNEUROSCI.2915-17.2018>
- Kubikova L., et al. (2014) **Basal ganglia function, stuttering, sequencing, and repair in adult songbirds** *Scientific Reports* **4** <https://doi.org/10.1038/srep06590>
- Kubikova L., Turner E. A., Jarvis E. D (2007) **The pallial basal ganglia pathway modulates the behaviorally driven gene expression of the motor pathway** *European Journal of Neuroscience* **25**:2145–2160 <https://doi.org/10.1111/j.1460-9568.2007.05368.x>
- Lashley K. S (1951) **‘The problem of serial order in behavior.’**, in *Cerebral Mechanisms in Behavior* :112–146
- Lipkind D., et al. (2013) **Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants** *Nature* **498**:104–108 <https://doi.org/10.1038/nature12173>
- Long M. A., Fee M. S (2008) **Using temperature to analyse temporal dynamics in the songbird motor pathway** *Nature* **456**:189–194 <https://doi.org/10.1038/nature07448>
- Lynch G. F., et al. (2016) **Rhythmic Continuous-Time Coding in the Songbird Analog of Vocal Motor Cortex** *Neuron* **90**:877–892 <https://doi.org/10.1016/j.neuron.2016.04.021>
- McCasland J (1987) **Neuronal control of bird song production** *Journal of Neuroscience* **7**:23–39 <https://doi.org/10.1523/JNEUROSCI.07-01-00023.1987>
- McInnes L., et al. (2018) **UMAP: Uniform Manifold Approximation and Projection** *Journal of Open Source Software* **3** <https://doi.org/10.21105/joss.00861>

- Miller M. N., Cheung C. Y. J., Brainard M. S (2017) **Vocal learning promotes patterned inhibitory connectivity** *Nature Communications* **8** <https://doi.org/10.1038/s41467-017-01914-5>
- Moll F. W., et al. (2023) **Thalamus drives vocal onsets in the zebra finch courtship song** *Nature* **616** <https://doi.org/10.1038/s41586-023-05818-x>
- Morita T., et al. (2021) **Measuring context dependency in birdsong using artificial neural networks** *PLOS Computational Biology* **17** <https://doi.org/10.1371/journal.pcbi.1009707>
- Nottebohm F., Stokes T. M., Leonard C. M (1976) **Central control of song in the canary, *Serinus canarius*** *Journal of Comparative Neurology* **165**:457–486 <https://doi.org/10.1002/cne.901650405>
- Okanoya K., Yamaguchi A (1997) **Adult bengalese finches (*Lonchura striata* var. *domestica*) require real-time auditory feedback to produce normal song syntax** *Journal of Neurobiology* **33**:343–356 [https://doi.org/10.1002/\(SICI\)1097-4695\(199710\)33:4<343::AID-NEU1>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1097-4695(199710)33:4<343::AID-NEU1>3.0.CO;2-A)
- Ölveczky B. P., et al. (2011) **Changes in the neural control of a complex motor sequence during learning** *Journal of Neurophysiology* **106**:386–397 <https://doi.org/10.1152/jn.00018.2011>
- Otchy T. M., et al. (2015) **Acute off-target effects of neural circuit manipulations** *Nature* **528**:358–363 <https://doi.org/10.1038/nature16442>
- Picardo M. A., et al. (2016) **Population-Level Representation of a Temporal Sequence Underlying Song Production in the Zebra Finch** *Neuron* **90**:866–876 <https://doi.org/10.1016/j.neuron.2016.02.016>
- Rajan R., Doupe A. J (2013) **Behavioral and Neural Signatures of Readiness to Initiate a Learned Motor Sequence** *Current Biology* **23**:87–93 <https://doi.org/10.1016/j.cub.2012.11.040>
- Sainburg T., Thielk M., Gentner T. Q (2020) **Finding, visualizing, and quantifying latent structure across diverse animal vocal repertoires** *PLOS Computational Biology* **16**:1–48 <https://doi.org/10.1371/journal.pcbi.1008228>
- Sakata J. T., Brainard M. S (2006) **Real-time contributions of auditory feedback to avian vocal motor control** *Journal of Neuroscience* **26**:9619–9628 <https://doi.org/10.1523/JNEUROSCI.2027-06.2006>
- Sakata J. T., Brainard M. S (2008) **Online contributions of auditory feedback to neural activity in avian song control circuitry** *Journal of Neuroscience* **28**:11378–11390 <https://doi.org/10.1523/JNEUROSCI.3254-08.2008>
- Sakata J. T., Hampton C. M., Brainard M. S (2008) **Social Modulation of Sequence and Syllable Variability in Adult Birdsong** *Journal of Neurophysiology* **99**:1700–1711 <https://doi.org/10.1152/jn.01296.2007>
- 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** *Journal of Neuroscience* **11**:2896–913 <https://doi.org/10.1523/JNEUROSCI.11-09-02896.1991>
- Schmidt M. F., Ashmore R. C., Vu E. T (2004) **‘Bilateral control and interhemispheric coordination in the avian song motor system’, in Annals of the New York Academy of Sciences** *New York Academy of Sciences* :171–186 <https://doi.org/10.1196/annals.1298.014>

- Seki Y., Okanoya K (2008) **Functional evidence for internal feedback in the songbird brain nucleus HVC** *NeuroReport* **19**:679–682 <https://doi.org/10.1097/WNR.0b013e3282fb7dae>
- Simpson H., Vicario D (1990) **Brain pathways for learned and unlearned vocalizations differ in zebra finches** *Journal of Neuroscience* **10**:1541–1556 <https://doi.org/10.1523/JNEUROSCI.10-05-01541.1990>
- Sober S. J., Brainard M. S (2012) **Vocal learning is constrained by the statistics of sensorimotor experience** *Proceedings of the National Academy of Sciences* **109**:21099–21103 <https://doi.org/10.1073/pnas.1213622109>
- Sober S. J., Wohlgemuth M. J., Brainard M. S (2008) **Central contributions to acoustic variation in birdsong** *Journal of Neuroscience* **28**:10370–10379 <https://doi.org/10.1523/JNEUROSCI.2448-08.2008>
- Stepanek L., Doupe A. J (2010) **Activity in a Cortical-Basal Ganglia Circuit for Song Is Required for Social Context-Dependent Vocal Variability** *Journal of Neurophysiology* **104**:2474–2486 <https://doi.org/10.1152/jn.00977.2009>
- Suge R., Okanoya K (2010) **Perceptual chunking in the self-produced songs of Bengalese finches (*Lonchura striata* var. *domestica*)** *Animal Cognition* **13**:515–523 <https://doi.org/10.1007/s10071-009-0302-4>
- Tanaka M., et al. (2016) **Focal expression of mutant huntingtin in the songbird basal ganglia disrupts cortico-basal ganglia networks and vocal sequences** *Proceedings of the National Academy of Sciences* **113**:E1720–E1727 <https://doi.org/10.1073/pnas.1523754113>
- Tanji J (2001) **Sequential Organization of Multiple Movements: Involvement of Cortical Motor Areas** *Annual Review of Neuroscience* **24**:631–651 <https://doi.org/10.1146/annurev.neuro.24.1.631>
- Tchernichovski O., et al. (2000) **A procedure for an automated measurement of song similarity** *Animal Behaviour* **59**:1167–1176 <https://doi.org/10.1006/anbe.1999.1416>
- Tian L. Y., Brainard M. S (2017) **Discrete Circuits Support Generalized versus Context-Specific Vocal Learning in the Songbird** *Neuron* **96**:1168–1177 <https://doi.org/10.1016/j.neuron.2017.10.019>
- Tian L. Y., Warren T. L., Brainard M. S. (2022) **Dynamic top-down biasing implements rapid adaptive changes to individual movements** *bioRxiv* <https://doi.org/10.1101/2022.06.16.496455>
- Tumer E. C., Brainard M. S (2007) **Performance variability enables adaptive plasticity of “crystallized” adult birdsong** *Nature* **450**:1240–4 <https://doi.org/10.1038/nature06390>
- Vates G. E., Vicario D. S., Nottebohm F (1997) **Reafferent thalamo- “cortical” loops in the song system of oscine songbirds** *Journal of Comparative Neurology* **380**:275–90
- Veit L., et al. (2021) **Songbirds can learn flexible contextual control over syllable sequencing** *eLife* **10**:1–19 <https://doi.org/10.7554/eLife.61610>
- Vu E. T., Mazurek M. E., Kuo Y. C (1994) **Identification of a forebrain motor programming network for the learned song of zebra finches** *Journal of Neuroscience* <https://doi.org/10.1523/jneurosci.14-11-06924.1994>

- Vyssotski A. L., et al. (2016) **A Neural Code That Is Isometric to Vocal Output and Correlates with Its Sensory Consequences** *PLOS Biology* **14**:1–21 <https://doi.org/10.1371/journal.pbio.2000317>
- Wang C. Z. H., et al. (2008) **Rapid interhemispheric switching during vocal production in a songbird** *PLOS Biology* **6**:2154–2162 <https://doi.org/10.1371/journal.pbio.0060250>
- Wang H., et al. (2019) **Transcriptional regulatory divergence underpinning species-specific learned vocalization in songbirds** *PLOS Biology* **17** <https://doi.org/10.1371/journal.pbio.3000476>
- Warren T. L., et al. (2011) **Mechanisms and time course of vocal learning and consolidation in the adult songbird** *Journal of Neurophysiology* **106**:1806–1821 <https://doi.org/10.1152/jn.00311.2011>
- Warren T. L., et al. (2012) **Variable sequencing is actively maintained in a well learned motor skill** *Journal of Neuroscience* **32**:15414–15425 <https://doi.org/10.1523/JNEUROSCI.1254-12.2012>
- Williams S. M., Nast A., Coleman M. J (2012) **Characterization of Synaptically Connected Nuclei in a Potential Sensorimotor Feedback Pathway in the Zebra Finch Song System** *PLOS ONE* **7** <https://doi.org/10.1371/journal.pone.0032178>
- Wiltschko A. B., et al. (2015) **Mapping Sub-Second Structure in Mouse Behavior** *Neuron* **88**:1121–1135 <https://doi.org/10.1016/J.NEURON.2015.11.031>
- Wittenbach J. D., et al. (2015) **An Adapting Auditory-motor Feedback Loop Can Contribute to Generating Vocal Repetition** *PLOS Computational Biology* **11** <https://doi.org/10.1371/journal.pcbi.1004471>
- Wohlgemuth M. J., Sober S. J., Brainard M. S (2010) **Linked control of syllable sequence and phonology in birdsong** *Journal of Neuroscience* **30**:12936–12949 <https://doi.org/10.1523/JNEUROSCI.2690-10.2010>
- Xiao L., et al. (2021) **Expression of FoxP2 in the basal ganglia regulates vocal motor sequences in the adult songbird** *Nature Communications* **12** <https://doi.org/10.1038/s41467-021-22918-2>
- Zann R. A. (1996) **The Zebra Finch: A Synthesis of Field and Laboratory Studies**
- Zhang Y. S., et al. (2017) **Temperature manipulation in songbird brain implicates the premotor nucleus HVC in birdsong syntax** *Journal of Neuroscience* **37**:2600–2611 <https://doi.org/10.1523/JNEUROSCI.1827-16.2017>

Article and author information

Avani Koparkar

Neurobiology of Vocal Communication, Institute for Neurobiology, University of Tübingen, Germany

ORCID iD: [0009-0004-3462-9453](https://orcid.org/0009-0004-3462-9453)

Timothy L. Warren

Center for Integrative Neuroscience, University of California, San Francisco, USA and Howard Hughes Medical Institute, Departments of Horticulture and Integrative Biology, Oregon State University, Corvallis, OR

ORCID iD: [0000-0002-4429-4106](https://orcid.org/0000-0002-4429-4106)

Jonathan D. Charlesworth

Center for Integrative Neuroscience, University of California, San Francisco, USA and Howard Hughes Medical Institute, Department of Clinical Research, Noctrix Health, Inc., Pleasanton, CA

ORCID iD: [0000-0002-6651-0009](https://orcid.org/0000-0002-6651-0009)

Sooyoon Shin

Center for Integrative Neuroscience, University of California, San Francisco, USA and Howard Hughes Medical Institute

ORCID iD: [0000-0002-0339-4856](https://orcid.org/0000-0002-0339-4856)

Michael S. Brainard

Center for Integrative Neuroscience, University of California, San Francisco, USA and Howard Hughes Medical Institute

ORCID iD: [0000-0002-9425-9907](https://orcid.org/0000-0002-9425-9907)

Lena Veit

Neurobiology of Vocal Communication, Institute for Neurobiology, University of Tübingen, Germany

For correspondence: lena.veit@uni-tuebingen.de

ORCID iD: [0000-0002-9566-5253](https://orcid.org/0000-0002-9566-5253)

Copyright

© 2024, Koparkar et al.

This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Editors

Reviewing Editor

Jesse Goldberg

Cornell University, Ithaca, United States of America

Senior Editor

Andrew King

University of Oxford, Oxford, United Kingdom

Reviewer #1 (Public Review):

Summary:

Songbirds provide a tractable system to examine neural mechanisms of sequence generation and variability. In past work, the projection from LMAN to RA (output of the anterior forebrain pathway) was shown to be critical for driving vocal variability during babbling, learning, and adulthood. LMAN is immediately adjacent to MMAN, which projects to HVC.

MMAN is less well understood but, anatomically, appears to resemble LMAN in that it is the cortical output of a BG-thalamocortical loop. Because it projects to HVC, a major sequence generator for both syllable phonology and sequence, a strong prediction would be that MMAN drives sequence variability in the same way that LMAN drives phonological variability. This hypothesis predicts that MMAN lesions in a Bengalese finch would reduce sequence variability. Here, the authors test this hypothesis. They provide a surprising and important result that is well motivated and well analyzed: MMAN lesions increase sequence variability - this is exactly the opposite result from what would be predicted based on the functions of LMAN.

Strengths:

- (1) A very important and surprising result shows that lesions of a frontal projection from MMAN to HVC, a sequence generator for birdsong, increase syntactical variability.
- (2) The choice of Bengalese finches, which have complex transition structures, to examine the mechanisms of sequence generation, enabled this important discovery.
- (3) The idea that frontal outputs of BG-cortical loops can generate vocal variability comes from lesions/inactivations of a parallel pathway from LMAN to RA. The difference between MMAN and LMAN functions is striking and important.

<https://doi.org/10.7554/eLife.93272.2.sa1>

Reviewer #2 (Public Review):

Summary:

This study investigates the neural substrates of syntax variation in Bengalese finch song. Here, the authors tested the effects of bilateral lesions of mMAN, a brain area with inputs to HVC, a premotor area required for song production. Lesions in mMAN induce variability in syntactic elements of song specifically through increased transition entropy, variability within stereotyped song elements known as chunks and increases in the repeat number of individual syllables. These results suggest that mMAN projections to HVC contribute to multiple aspects of song syntax in the Bengalese finch. Overall the experiments are well-designed, the analysis excellent, and the results are of high interest.

Strengths:

The study identifies a novel role for mMAN, medial magnocellular nucleus of the anterior nidopallium, in the control of syntactic variation within adult Bengalese finch song. This is of particular interest as multiple studies previously demonstrated that mMAN lesions do not effect song structure in zebra finches. The study undertakes a thorough analysis to characterise specific aspects of variability within the song of lesioned animals. The conclusions are well supported by the data.

<https://doi.org/10.7554/eLife.93272.2.sa0>

Author Response

The following is the authors' response to the original reviews.

Public reviews:

Reviewer #1 (Public Review):

Summary:

Songbirds provide a tractable system to examine neural mechanisms of sequence generation and variability. In past work, the projection from LMAN to RA (output of the anterior forebrain pathway) was shown to be critical for driving vocal variability during babbling, learning, and adulthood. LMAN is immediately adjacent to MMAN, which projects to HVC. MMAN is less well understood but, anatomically, appears to resemble LMAN in that it is the cortical output of a BG-thalamocortical loop. Because it projects to HVC, a major sequence generator for both syllable phonology and sequence, a strong prediction would be that MMAN drives sequence variability in the same way that LMAN drives phonological variability. This hypothesis predicts that MMAN lesions in a Bengalese finch would reduce sequence variability. Here, the authors test this hypothesis. They provide a surprising and important result that is well motivated and well analyzed: MMAN lesions increase sequence variability - this is exactly the opposite result from what would be predicted based on the functions of LMAN.

Strengths:

- (1) A very important and surprising result shows that lesions of a frontal projection from MMAN to HVC, a sequence generator for birdsong, increase syntactical variability.*
- (2) The choice of Bengalese finches, which have complex transition structures, to examine the mechanisms of sequence generation, enabled this important discovery.*
- (3) The idea that frontal outputs of BG-cortical loops can generate vocal variability comes from lesions/inactivations of a parallel pathway from LMAN to RA. The difference between MMAN and LMAN functions is striking and important.*

Weaknesses:

- (1) If more attention was paid to how syllable phonology was (or was not) affected by MMAN lesions then the claims could be stronger around the specific effects on sequence.*

Reviewer #2 (Public Review):

Summary:

This study investigates the neural substrates of syntax variation in Bengalese finch songs. Here, the authors tested the effects of bilateral lesions of mMAN, a brain area with inputs to HVC, a premotor area required for song production. Lesions in mMAN induce variability in syntactic elements of song specifically through increased transition entropy, variability within stereotyped song elements known as chunks, and increases in the repeat number of individual syllables. These results suggest that mMAN projections to HVC contribute to multiple aspects of song syntax in the Bengalese finch. Overall the experiments are well-designed, the analysis excellent, and the results are of high interest.

Strengths:

The study identifies a novel role for mMAN, the medial magnocellular nucleus of the anterior nidopallium, in the control of syntactic variation within adult Bengalese finch song. This is of particular interest as multiple studies previously demonstrated that mMAN lesions do not affect song structure in zebra finches. The study undertakes a thorough analysis to characterise specific aspects of variability within the song of lesioned animals. The conclusions are well supported by the data.

Weaknesses:

The study would benefit from additional mechanistic information. A more fine-grained or reversible manipulation, such as brain cooling, might allow additional insights into how mMAN influences specific aspects of syntax structure. Are repeat number increases and transition entropy resulting from shared mechanisms within mMAN, or perhaps arising from differential output to downstream pathways (i.e. projections to HVC)? Similarly, unilateral manipulations would allow the authors to further test the hypothesis that mMAN is involved in inter-hemispheric synchronization.

We thank the reviewers and editor for their encouraging and helpful comments and suggestions. We have revised the previous submission with new analyses and discussion to address points raised by the reviewers.

Following the suggestion of Reviewer 1 we have added an analysis of the effects of mMAN lesions on syllable phonology, using a variety of measures. We have included 3 new Figure Supplements that detail our analyses and elaborate on these points.

We agree with Reviewer 2 that reversible and unilateral manipulations would be interesting and potentially enable additional insights into the mechanisms by which mMAN influences song sequencing, and we are planning to perform such experiments in future studies.

We made additional minor changes throughout the manuscript to address other points raised by the reviewers, and we thank them again for their time and effort in providing constructive feedback to improve our study.

A complete point by point detailing of these changes is included below, interspersed with the reviewer comments.

Reviewer #1 (Recommendations For The Authors):

The opposite result from what would be predicted based on the functions of LMAN.

Shoring up the paper's claims and ruling out alternative interpretations will require attention to the following issues:

Major comments

(1) Acoustic structure of syllables

Line 294 & Sup. Figure 2, in some birds the syllable acoustic structures seem to be significantly different between the pre- and post-lesion condition, e.g. 'w' in Bird 1, 'g' in Bird 2, 'blm' in Bird 6. This observation seems to contradict the claim that acoustic structures are not affected by MMAN lesions.

Related to the previous point, a more detailed analysis is needed to quantify the extent of acoustic changes caused by MMAN lesions. For example, do these pre- and post-lesion syllables form distinct clusters if embedded in a UMAP? Do more standard measures of syllable phonology (e.g. SAP similarity scores or feature distributions) show differences in pre- and post-MMAN lesion?

We agree with the reviewer that there were individual syllables as illustrated in the average spectrograms of Figure 2 – figure supplement 1 that qualitatively differed between pre- and post-lesion recordings. We have followed the reviewer's suggestion to quantify changes to syllable phonology using both similarity scores by Sound Analysis Pro (SAP) and a variety of identified acoustic features.

In brief, these measures largely corroborate the conclusion that for most birds and syllables there was little or no difference in phonology between pre- and post-lesion songs, but that in

a minority of cases syllables were altered noticeably (further detail below). In those cases where syllable phonology was altered, changes were not consistent across birds, and we cannot rule out off-target effects due to damage to structures or fibers of passage neighboring mMAN, so that it is unclear whether some subtle changes to syllable phonology can be attributed to mMAN lesions versus other causes. Future studies could more specifically examine whether damage to mMAN alone is sufficient in some cases to degrade syllable structure by using viral or other approaches that enable the more specific disruption of mMAN projection neurons.

In practice, almost all syllables were identifiable in post-lesion songs so that we could unambiguously assign identity for purposes of evaluating effects of lesions on sequencing. Moreover, in any individual cases where there was ambiguity in syllable identity, we used the sequential context to assign the most likely label. Thus, any errors in assignment in such cases would have tended to reduce rather than accentuate the magnitude of reported sequencing effects. Lastly, each of the reported effects of mMAN lesions on sequencing were observed in multiple birds for which we detected no significant changes to syllable similarity.

Further details of the analyses of syllable structure are detailed below, and have been added as new figure supplements:

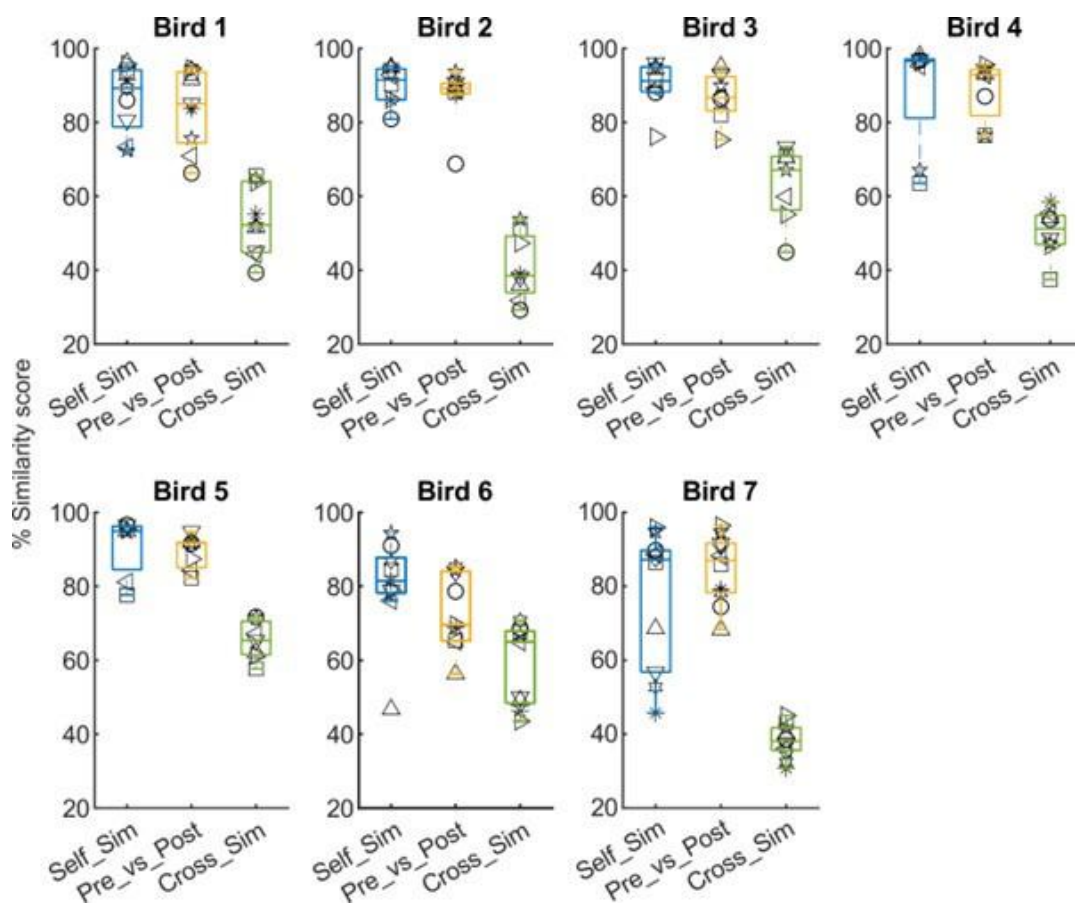
(1) Syllable similarity scores calculated using SAP (Sound Analysis Pro) (new Figure 2 – figure supplement 2). We compared pre-post lesion similarity scores for each syllable with selfsimilarity measures for the same syllables taken from separate control recordings before lesions. For comparison, we also included a cross-similarity score for syllables of different types. These measures confirmed the qualitative impression from spectrograms that for most birds there were no greater changes to syllable structure following lesions than was present across control recordings. For one bird, pre-post changes were significantly larger than changes across control recordings, but pre-post similarity remained higher than crosssimilarity.

(2) Analysis of fundamental frequency and coefficient of variation (CV) of fundamental frequency of select syllables for each bird before and after mMAN lesions (new Figure 2-figure supplement 3). This analysis is directly comparable with the same analysis performed on LMAN lesions in Sakata, Hampton, Brainard (2008). We carried out this analysis in part to address changes to syllable structure that might have inadvertently arisen due to damage to LMAN, which sits immediately lateral to mMAN. In the Bengalese finch and zebra finch, lesions of LMAN cause little change to the mean fundamental frequency of individual syllables but cause a consistent reduction in the coefficient of variation (CV) of fundamental frequency across repeated renditions of a given syllable (Sakata, Hampton, Brainard 2008, Andalman, Fee 2009, Warren et al. 2011.). We therefore supposed that unintended damage to LMAN or its projections to RA might have resulted in a reduction in the CV of syllables following mMAN lesions. Instead, we saw a modest increase in the CV of fundamental frequency (mean across birds of +20%; range -19 to +43%). These data suggest that off target effects on LMAN were largely absent in our experiments (consistent with histology, e.g. Figure 1 - figure supplement 1).

(3) Comparison of Entropy of spectral envelope (entS), Temporal centroid for the temporal envelope (meanT), First, second and third formants (F1, F2, F3), before and after lesions (calculated using the python SoundSig toolbox (Elie and Theunissen 2016) (new Figure 2-figure supplement 4). Acoustic features generally showed little change between pre and post lesion songs. They highlight as relative outliers the same individual examples that stand out in the average spectrograms in Figure 2 – figure supplement 1.

Author response image 1.

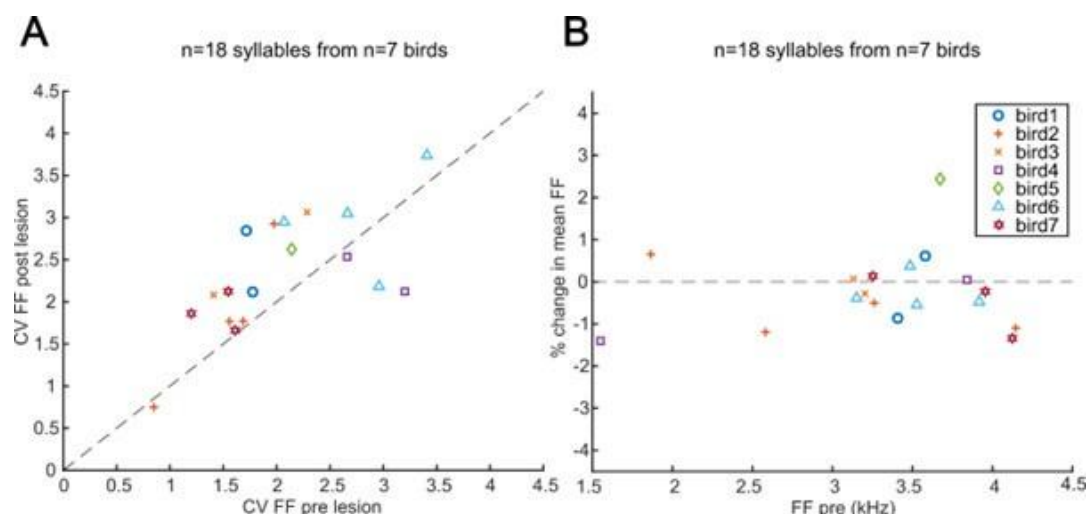
Syllable similarity calculated using Sound Analysis Pro (SAP). ‘Self Similarity’ = Similarity comparison of syllables before mMAN lesions to syllables of the same type, taken from two separate control recordings before the lesions, ‘Pre vs Post’ = Similarity comparison of the same syllable types before and after mMAN lesions, ‘Cross Similarity’ = Similarity comparison of each syllable type to other syllable types. For Birds 1-2 and 4-7, ‘Self Similarity’ was not significantly different from ‘Pre vs Post’ Similarity ($p > 0.05$, Wilcoxon sign rank test), while for Bird 3, there was a significant difference ($p = 0.03$, Wilcoxon sign rank test). For all birds ‘Pre vs Post’ was significantly different from ‘Cross Similarity’ ($p < 0.05$, Wilcoxon sign rank test). On average, ‘Pre vs Post’ was 4.8 % less than ‘Self Similarity’ (range 0.2%-14%) while ‘Cross Similarity’ was 40% less than ‘Self Similarity’ (range 20.2%-56.3%). These measures confirm the qualitative impression from Figure 2- figure supplement 1 that for most birds and syllables there were no greater changes to syllable structure following lesions than was present across control recordings, and that pre-post similarity remained higher than cross-similarity, i.e. syllables remained clearly identifiable.



Author response image 2.

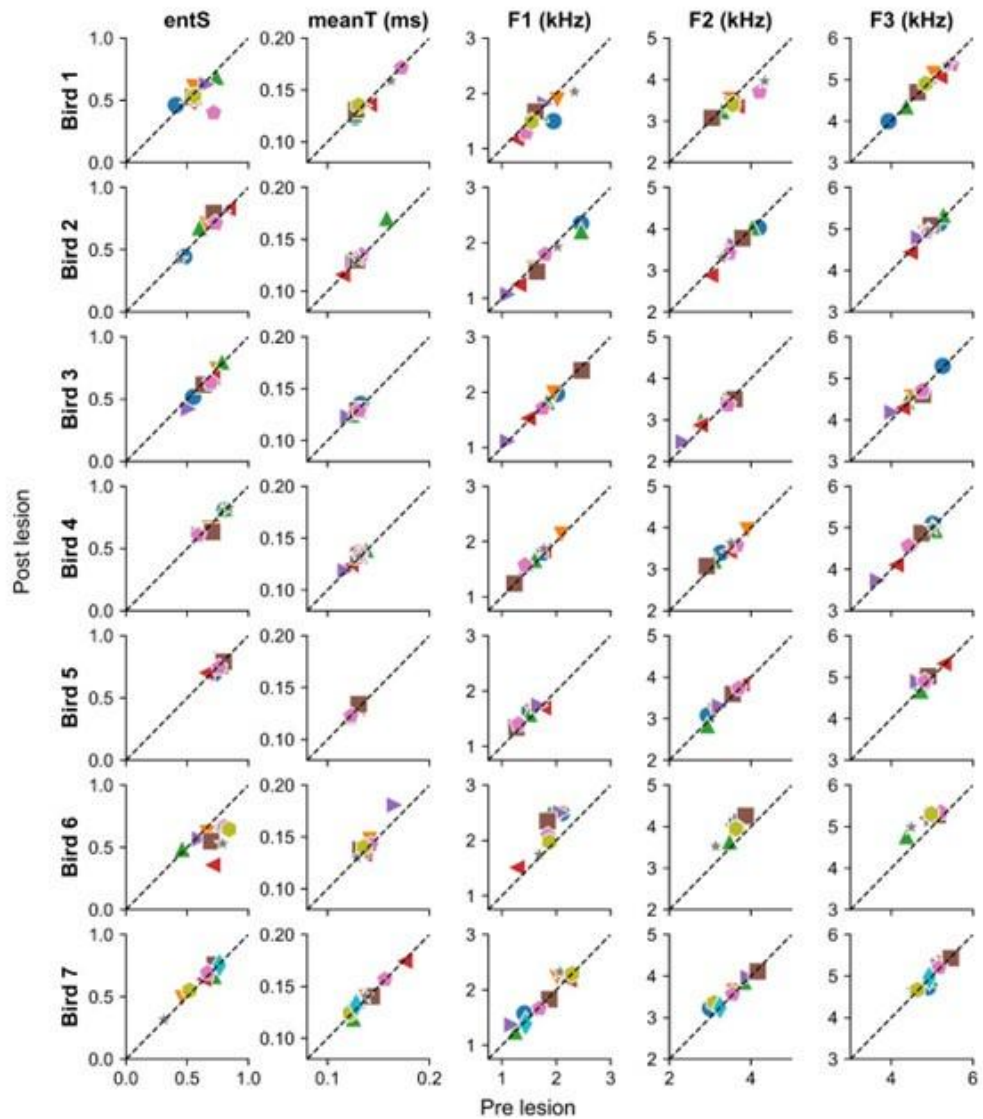
(A) CV of fundamental frequency (FF) of select syllables before and after mMAN lesions. In the Bengalese finch and zebra finch, lesions of lMAN, which sits immediately lateral to mMAN, cause a consistent reduction in the coefficient of variation (CV) of fundamental frequency across repeated renditions of a given syllable (Sakata, Hampton, Brainard 2008, Andalman, Fee 2009, Warren et al. 2011). We therefore supposed that unintended damage to lMAN or its projections to RA might have resulted in a reduction in the CV of syllables following mMAN lesions. Instead we saw a modest increase in the CV of fundamental

frequency ($p < 0.05$, Wilcoxon sign rank test; mean across birds of +20%; range -19 to +43%). These data suggest that it is unlikely that changes to syllable structure might have arisen due to accidental damage to IMAN. (B) Percent change in mean fundamental frequency after mMAN lesions vs mean fundamental frequency before mMAN lesions.



Author response image 3.

Selected acoustic features for all syllables in all birds before and after mMAN lesions. Different colors represent different syllable types per bird. 'entS' = Entropy of spectral envelope, 'meanT' = Temporal centroid for temporal envelope, 'F1' = First formant, 'F2' = Second formant, 'F3' = Third formant. Acoustic features generally showed little change between pre and post lesion songs. They highlight as relative outliers the same individual examples that stand out in the average spectrograms in Figure 2 – figure supplement 1.



(2) Shoring up claims of increased transitional variability

Line 301 & Sup. Figure 1, in several birds (1, 2, 5, 6), seems that there is a downward trend for postlesion, i.e. the transition entropy gradually decreases with time. How to exclude the possibility that the increased variability is a transient effect, e.g. caused by surgery side effects or destabilization of circuits, which may eventually recover to normal?

Transition entropy remains elevated for as long as the birds were followed in this study. While the persistence of the effects we observed is longer than transient effects such as those following Nif lesion in zebra finches (Otchy et al., 2015 ~2 days), we cannot rule out either recovery or further deterioration following lesions on much longer time scales, such as those reported by Kubikova et al., 2007 (X lesions, 6 months). We have now added data points for 4

birds where we had songs from later timepoints following lesions; for three of these birds, transition entropy remained elevated above the baseline values for 14 and 33 days, respectively (Figure 1 - figure supplement 2).

Line 313 & Sup. Figure 4, the claim that "transitions that had low history dependence tended to show larger changes after mMAN lesions" needs better statistical support, because in Sup. Figure 4, the correlation is not significant.

We apologize for the phrasing. We have changed the sentence to: "Consistent with the first possibility, we observed that there was a nonsignificant trend toward larger changes after mMAN lesion for transitions with low history dependence."

Figure 4C-D, only data from 5 out of 7 birds was included, did the other two birds not have repeats? If so, the authors need to be explicit on data exclusion.

The reviewer's inference is correct that in our dataset only 5 out of 7 birds had songs which contained repeat phrases. We have added the following sentence to state that explicitly: "In our dataset of 7 birds, only 5 birds had songs which contained repeat phrases."

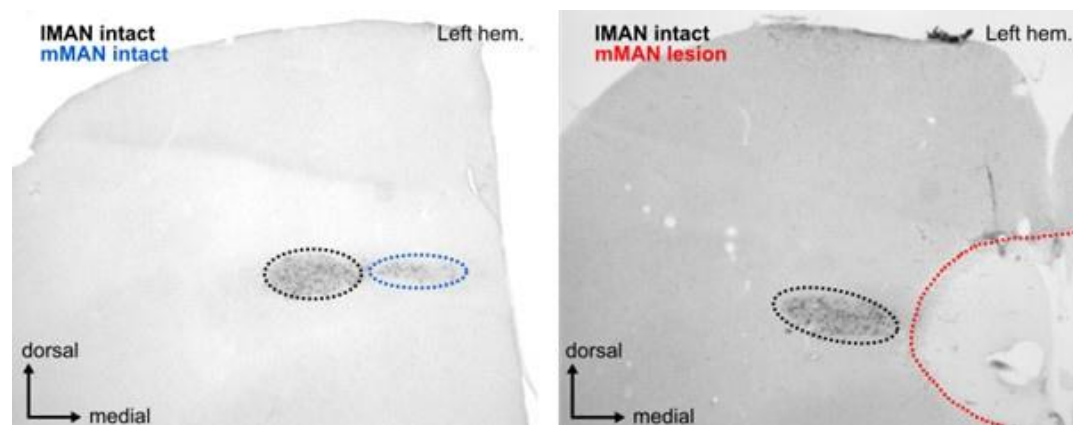
Minor comments

Sup. Figure 3, to help readers understand, 1) add symbols and arrows to point to the structures; 2) indicate the orientation of the slide, e.g. which direction is medial/lateral; 3) a negative control without lesion needs to be shown for comparison.

We have made the suggested changes and updated new Figure 1- figure supplement 1.

Author response image 4.

Image of calcitonin gene-related peptide (CGRP)-stained frontal section (left) control and (right) bird 5. CGRP labels cells in both IMAN (seen in black to the left of the lesion) and mMAN (blue, intact; red, completely destroyed).



A statistical test is needed for Sup. Figure 5B.

We have modified the Figure legend for Figure 3 – figure supplement 1 as follows:

"Change in transition entropy was not significantly different for transitions within chunks and at branchpoints ($p > 0.05$, Wilcoxon rank sum test)"

Line 363, these can be moved to the Introduction, so readers have a better sense of what's already known about mMAN lesion.

We have moved the sentence to Introduction.

Fig 1e. RA also projects to DLM.

Our intention was to focus on the connections involving mMAN; we have now added the connection in Figure 1E.

Reviewer #2 (Recommendations For The Authors):

Please address this issue in the discussion (no new experiments required): It would be interesting to consider how social context modulates the variability of the song. In these experiments, Bengalese finches were singing in isolation. How might changes in syntax be modulated by the presence of a female in directed song and in other social contexts?

Thank you for your suggestion. One study by Jarvis, et al., (Jarvis E., et al., 1998) shows that ZENK expression in mMAN after singing does not differ between female-directed singing, undirected singing and singing in presence of a male conspecific. This suggests that activity in mMAN might not be modulated by social context. But we agree that it would be interesting to test how a change in social context (which typically leads to reduced transition entropy) interacts with the increased variability we see after mMAN lesions. We have added the following sentences to the discussion:

“In our study, we only recorded song sequencing of male Bengalese finches singing in isolation. Social context, such as female-directed song, can also change song sequencing (Hampton, Sakata and Brainard, 2009; Chen, Matheson and Sakata, 2016). It would be interesting to test whether mMAN plays a role in the social context-modulated changes in sequencing (Jarvis et al., 1998), similar to how lMAN contributes to social context-modulated changes in syllable structure (Sakata, Hampton and Brainard, 2008).”