An anterior forebrain pathway in parrots is necessary for individual signatures of learned vocalizations

Zhilei Zhao[1], Han Kheng Teoh[2], Julie Carpenter[1], Frieda Nemon[1], Brian Kardon[1], Itai Cohen[2], Jesse H. Goldberg[1,*]

[1]*Department of Neurobiology and Behavior, Cornell University, Ithaca, NY 14853, USA*

[2]*Department of Physics, Cornell University, Ithaca, NY 14853, USA.*

**Corresponding author*

Abstract

Parrots have enormous vocal imitation capacities and produce individually unique vocal signatures. Like songbirds, parrots have a nucleated neural song system, but because the two sister groups diverged over 50 million years ago it remains unknown if the song systems operate according to similar principles. Here we established a neuroscience-compatible call-and-response behavioral paradigm to elicit learned contact calls in budgerigars (*Melopsittacus undulatus*). Using variational autoencoder-based machine learning methods, we show that contact calls within affiliated groups converge but that individuals maintain unique acoustic features, or vocal signatures, even after call convergence. Next, we transiently and reversibly inactivated the outputs of the anterior forebrain pathway (AFP) as animals engaged in call-and-response interactions. Contact call pitch modulation, stereotypy and acoustic structure were dramatically degraded, resulting in the loss of individually unique features of the vocal signature. These results are in stark contrast to results from songbirds, where AFP inactivations with identical methods have the opposite result of increasing the stereotypy of learned vocalizations and retaining acoustic structure. Together, our results show that the anterior pathways of parrots and songbirds have different functional organization and that parrot frontal cortical circuits are necessary for individually unique features of vocal communication.

Introduction

The ability to learn vocalizations and use them for social communication was a key milestone in the evolution of human language [1]. This trait, rare among non-human animals, is observed in three clades of birds: oscine songbirds, parrots, and hummingbirds [2–4]. Interestingly, all of these clades evolved nucleated forebrain circuits, or 'song systems' with similar-appearing basal ganglia thalamocortical pathways. But because these clades diverged over 50 million years ago and have non-vocal learners in their lineages, it remains unknown if the principles of song system function are conserved [1,4].

The neural mechanisms of song production and learning have been extensively studied in three songbird species: zebra finches, bengalese finches, and canaries [5–8]. Conceptually similar results across these species support a conserved logic to song system architecture where anatomy maps

clearly to function (Figure 1A). The output of the forebrain song system, RA (robust nucleus of the arcopallium), is analogous to mammalian primary motor cortex and has only two main inputs: HVC (used as a proper name) and lateral magnocellular nucleus of the anterior nidopallium (LMAN) [9,10]. HVC is part of the posterior motor pathway that is necessary for adult song production but not for vocal babbling in juveniles. HVC lesions or inactivations degrade acoustic structure in adult song but have little effect on normal vocal babbling in juveniles [8,11–13]. During singing, HVC neurons in adult birds exhibit stereotyped, temporally precise activity that drives correspondingly stereotyped syllable sequences in the adult song [10,14,15]. These basic observations of posterior pathway function were additionally supported in more exotic dueting songbirds, including white-browed sparrow weavers and plain-tailed wrens, supporting the idea of conserved posterior song system function across all songbird species studied [16,17].

In contrast to HVC, LMAN is the output of the songbird anterior forebrain pathway (AFP)—a basal ganglia (BG) thalamocortical loop—necessary for song learning but not for adult song production. Lesion or inactivation LMAN or its thalamic input [11,12], has little effect on adult song but a devastating effect on vocal babbling and learning [18–21]. With LMAN inactivated, the normal variability of juvenile vocal babbling is instantly lost and the song becomes abnormally stereotyped. Consistent with its role in driving vocal variation, LMAN neurons exhibit variable, bursty activity during singing that drives the correspondingly variable song components, including the vocal babbling of juveniles [13,20,21].

The discovery that the AFP functions as a variability-generator necessary for learning is one of the most important contributions of the songbird field to neuroscience. It shows that variability characteristic of yet-to-be learned behaviors is not the *de facto* output of an immature motor system [22], but is instead actively generated by neural circuits for the purpose of learning. Though a conceptually similar 'variability-generating' circuit has yet to be identified in mammals, frontal BG-cortical loops may generate variability since catecholamines reduce variability in both humans and in songbirds [23,24].

To test if frontal cortical variability-generating functions generalize to distantly related vocal learners or if they are an idiosyncratic feature of songbirds, we studied vocal production in the budgerigars (*Melopsittacus undulatus*), a member of the parrot family. Anatomical studies show that the parrot song system resembles the songbird one (Figure 1B), and as a result has been proposed to have a similar functional organization [25–31]. First, the central nucleus of the anterior arcopallium (AAC) may be analogous to RA because it is similarly positioned in the posterior forebrain, sends axons directly to motor neurons in the brainstem, and receives inputs from both a posterior and frontal pathway. Lesioning AAC or its input from the posteriorly-located nucleus NLC (central nucleus of the lateral nidopallium) degrades the acoustic structure of learned vocalizations [28,32,33], supporting the possibility that NLC may be analogous to HVC [25,31]. The function of the parrot AFP, on the other hand, remains largely unexplored. In contrast to the songbird, the parrot has two nuclei that constitute AFP outputs: MO (oval nucleus of the mesopallium) and NAO (oval nucleus of the anterior nidopallium). Like LMAN, they are frontal cortical nuclei that project to both the basal ganglia and AAC (the RA analogue) [25–31]. Yet to our knowledge only one study conducted lesions of MO and found that frequency modulation was gradually lost over a course of several days [28], but only a small number of calls were examined

from one bird and findings were based on visual inspections of calls. Thus it remains unclear if the frontal song system has LMAN-like functions. Here we revisited the function of AFP in parrots with reversible brain inactivations and machine learning-enabled quantification of large datasets of budgerigar vocalizations.

Results

A behavioral paradigm to elicit thousands of learned contact calls per bird

Studying the detailed acoustic structure of vocalizations and how they may change with neural manipulations requires analysis of large datasets, ideally thousands of calls per bird [34–36]. We set out to establish a behavioral paradigm that would provide such a dataset even in single-housed birds with pristine acoustic isolation. We focused on contact calls, a major type of learned vocalizations in budgerigars, that are used to coordinate flock behavior during flight and to localize individuals in a colony [37]. Past work showed that contact calls are learned, continuously adjusted even in adulthood, and may contain unique signatures that allow for individual recognition [38–42]. To sample contact calls efficiently, we simulated its natural use with a trial-by-trial behavioral paradigm. We first recorded colony noise from a social group (5 males, 3 females) and divided the recording into two-minute segments. We then isolated individuals in a sound-proof box and used directional microphones to record their vocalizations (Figure 2A). In a given trial, one segment of colony noise was randomly selected and played to the bird through a speaker, followed by a silent inter-trial interval (average 10 mins). Isolated birds responded to colony noise with the production of contact calls so reliably that we were able to obtain hundreds to thousands of contact calls per day with this method (average counts range: 180-1612, n=8 birds, Figure 2B).

A variational autoencoder identifies individual signatures even within shared call types

To compare contact calls between individuals, previous studies relied on visual inspection or descriptive acoustic features [39–42]. However, acoustic features such as pitch, pitch goodness and entropy can miss important acoustic details [36]. We therefore adopted a recently developed method [36], variational auto-encoder (VAE), to analyze contact calls in an unsupervised manner (Figure 2C). Briefly, vocalizations were first converted to spectrograms, then an encoder-decoder network was trained to reconstruct the original spectrograms from low-dimensional latent features. We next examined these latent features for downstream visualization using UMAP (uniform manifold approximation and projection) for dimensionality reduction and with a SVM (support vector machine) for quantitative classification of uniqueness.

Using these methods, we quantitatively confirmed and extended three previous observations on budgerigar contact calls. First, individuals produced both individually unique contact calls as well as shared calls that exhibited highly similar acoustic structures [39–41] (Figure 2B). Second, quantitative analysis revealed that despite this group convergence individuals maintained unique signatures even within shared calls [38,42]. This is visualized in the UMAP projection of the latent features, in which the shared group call appears as a large single cluster, while calls from specific individuals occupy distinct parts of the large cluster (Figure 2D-E). To quantify the uniqueness of individual signatures, we trained SVM classifiers to decode individual identity from call latent features [43]. Unsurprisingly, as we increased the dimension of principal inputs, i.e. used more

information from the spectrogram, the decoding accuracy gradually improved (Figure 2F). With eight or more principal inputs, the SVM correctly predicted bird identity more than 90% of the time, even though birds shared highly similar calls. We further restricted the analysis to the two major call types (Figure 2B, 2E, marked by asterisk and plus symbols) that were shared across individuals, and found that the SVM could still decode bird identity ~90% of the time (Figure 2G). This accuracy was significantly higher than a previous study based on summary acoustic metrics [42] (~40%), suggesting, unsurprisingly, that spectrograms contained extra information for discrimination not captured by acoustic feature vectors [36].

Group vocalizations dynamically change over time

A proposed function of contact calls is to signal affiliation among birds, which, in the wild, can change over time [37,44–46]. Call matching would be a more honest signal of affiliation if calls dynamically changed over time, requiring ongoing social interactions [40,44]. To test if our analyses could capture this dynamism of budgerigar contact calls, we compared contact calls of each individual in the group produced at week 3 (Figure 2D) and week 15 (Figure 2E) following the group assimilation. We found that the group convergence and individual signatures were not static but exhibited dynamic drift, as shown by past qualitative description [39]. We trained a single VAE on the group calls produced 12 weeks apart and visualized them in the same UMAP space (Figure 2D-E). We then quantified time-dependent changes in call structures by training an SVM classifier on calls at week 3, then decode caller identity at week 15, or vice versa. Decoding accuracy dropped significantly from ~90% to ~40% (Figure 2H).

Although these behavioral phenomena have been described before, here we provided quantitative characterization under a unified framework. We show that living in a social group, budgerigars need to constantly monitor contact calls of others, and modify their own calls accordingly to preserve both convergence and individual signature. Our ability to confirm past findings with new analytical methods also lays the foundation for studying how brain inactivations affect call acoustic structure.

AFP-inactivated budgerigars maintain the ability to respond to stimulus calls

We next examined the role of the AFP circuits in contact call production. Since birds do not readily vocalize in the days after surgery and permanent lesions are associated with experience-independent circuit remodeling [47], lesions can fail to test the effect of immediate loss of AFP contribution to vocalizations. We thus used reverse microdialysis to inactivate the output of the AFP. Reverse microdialysis is an established method in songbirds to transiently and reversibly inactivate brain regions over hour timescales and is particularly well suited for the study of nucleated systems [8,48,49]. In our past work in zebra finches using identical homemade probes, we carefully calibrated inactivated tissue volume as a function of drug concentration and replicated the observation that LMAN inactivations cause an immediate and reversible loss of vocal variability [8,12,13,48–50].

Based on our past calibrations, we were not confident that we could specifically target MO or NAO separately because they are immediately neighboring structures with a shared border (Fig. 1B, Methods). Thus, we implanted custom microdialysis probes bilaterally at the center of the MO/NAO complex to inactivate both structures at once (Figure 3A). After birds recovered from

159 surgery, we infused saline (PBS) or tetrodotoxin (TTX 50 μM) into the probe on alternating days
160 to compare their effects on contact calls in the call-and-response paradigm. We verified our
161 targeting with histology and confirmed diffusion from the probes into both MO and NAO but not
162 the underlying basal ganglia nucleus MMSt (Figure 3B).

163 To sample contact calls, we used the same call-and-respond behavioral paradigm (Figure 2A),
164 but played individual calls instead of colony noise to better examine the temporal relationship
165 between stimulus and response (Figure 3C-D). To test the role of the frontal system in producing
166 calls with clearly distinct acoustic structures, we chose five birds unknown to one another with
167 clearly distinguishable individual signatures (Figure 3E-I).

168 With MO/NAO inactivated, all birds were still able to reliably respond to call stimuli (Figure 3C-I).
169 Response rate significantly decreased in three birds, but not the other two (Figure 4A). Similarly,
170 we didn't observe a consistent change of response latency across birds (Figure 4B). Together
171 these data show that the AFP is not necessary for responding to a stimulus call with appropriate
172 timing.

173 **Inactivation of budgerigar AFP reduces call acoustic structure and increases call**
174 **variability**

175 We next examined the acoustic structure of contact calls before, during, and after MO/NAO
176 inactivation. The pitch of normal budgerigar contact calls ranges from 2 to 4 kHz, centered near
177 3 kHz [37] (Figure 4C). We found that when AFP was inactivated, despite a slight decrease the
178 mean pitch of responded calls still fell into this range (Figure 4C), suggesting that the birds were
179 able to produce contact-call-like vocalizations. However, two aspects of call acoustic structure
180 were dramatically disrupted. First, calls produced under the TTX condition exhibited significantly
181 less pitch modulation compared to vehicle, and virtually all of the rapid pitch modulations were
182 eliminated (Figure 3E-I), as quantified by coefficient of variation (Figure 4D, PBS:TTX, 1.80±0.58)
183 and pitch modulation depth (Figure 4E, PBS:TTX, 1.99±0.64, Methods). Second, call durations
184 were affected. Normal contact calls, and those observed with vehicle in the probes, last between
185 0.15 to 0.25 sec with a narrow distribution (Figure 4F). When AFP was inactivated, this temporal
186 control of call duration was reduced, causing the production of abnormally short or long calls
187 (Figure 3E-I, Figure 4F).

188 We then quantified call stereotypy using identical statistical methods previously used to analyze
189 the change in zebra finch syllables following inactivation of LMAN [12,21]. In zebra finches, LMAN
190 inactivation causes an immediate reduction in syllable duration variability, quantified as the
191 coefficient of variation of syllable duration distributions [11,12]. Further, the phonological variability
192 of syllables is also dramatically reduced, as quantified with acoustic feature vectors derived from
193 the Sound Analysis Pro software package [21]. Using these identical analyses, we found that
194 inactivation of MO/NAO in budgerigars had the opposite effects of increasing the variability of
195 syllable duration and phonology (Figure 4G-H).

196 Notably, budgerigars exhibited defects in pitch and duration in the first calls they produced
197 immediately after TTX infusion (Figure 3E-I, green boxes). Furthermore, the effects were
198 reversible: when the drug was washed out, contact calls recovered to normal within hours (Figure

S1). In one bird we also used muscimol to inactivate AFP and observed similar changes in contact calls (Figure S2), suggesting that the call defects under TTX were caused by inactivation of MO/NAO, rather than blocking of fibers that pass through this region.

Inactivation of budgerigar AFP reduces the individuality of contact calls

The significant loss of pitch modulation and increase of phonological variability with AFP inactivation may reduce individual signatures in contact calls. We first checked this possibility by training a single VAE for calls produced under the PBS and TTX condition of all birds, then embedding their latent features in the same UMAP space. Since normal calls of these stranger birds are visually distinct (Figure 3E-I), they should form well-separated clusters in the UMAP. Changes in acoustic structure induced by AFP inactivation will drive TTX calls to deviate from the normal clusters. However, the deviation could happen in two ways that lead to different outcomes on the individual signatures (Figure 5A). If AFP inactivation reduces individuality, calls from different birds should move away from their normal clusters and converge into one large and centered cluster (Figure 5A, hypothesis 1). Alternatively, if AFP inactivation affected each bird in unique ways, individuality may be maintained, in which case the UMAP embeddings of the AFP-inactivated calls would still reveal clearly distinct clusters associated with different individuals (Figure 5A, hypothesis 2). Figure 5B shows that our data are more consistent with hypothesis 1: calls converged to a centered cluster following AFP inactivation, suggesting the loss of individual identity information.

However, since UMAP is a non-linear method, large distances in UMAP space do not directly translate to large differences in acoustic space. To more rigorously test the loss of individual signatures with AFP inactivation, we asked if a potential listener in the social group is still able to decode caller identity when listening to calls produced under the TTX condition. We simulated this situation by training a SVM classifier on the normal calls, then using it to decode the caller identity at the TTX condition. We found that the decoding accuracy dropped significantly from ~95% to ~50% (Figure 5C), demonstrating a dramatic reduction in individual identity information.

Finally, we reasoned that although the original individual signatures in contact calls were greatly reduced with AFP inactivation, different birds may not produce identical sounds, due to individual differences in the biomechanics of the vocal organ or individually learned components that may be driven by the NLC-AAC pathway. This could be seen in the UMAP embedding: although calls produced under the TTX condition converge into one large central cluster, some level of separation between individuals is preserved (Figure 5B). For quantification, we trained SVM classifiers separately on the PBS and TTX datasets. We found that with enough information from the spectrograms, the SVM could still decode caller identity on the TTX dataset with a decent level of accuracy (Figure 5D). However, the decoding accuracy at the TTX condition was significantly lower than the PBS condition across all ranges of principal dimensions (Figure 5D), suggesting that AFP-inactivated calls were less discernible. We further quantified the separability of calls produced by different individuals using a Silhouette score that contrasts intra- and inter-cluster distances (Methods). We found that Silhouette score was significantly smaller on the TTX dataset (Figure 5E), again suggesting that AFP-inactivated calls were less separable across individuals.

Discussion

One of the most important contributions of the songbird field to neuroscience is the discovery that frontal cortical circuits, outputs of a basal ganglia-thalamocortical loop, actively inject variability into behavioral output for the purpose of learning [13,18,20,48]. Inactivation of LMAN transforms the variable song of juvenile birds into stereotyped syllables produced in a stereotyped sequence. And in adults, LMAN inactivations similarly reduce natural trial-by-trial variability of learned vocalizations. Here we tested if this result generalizes to budgerigars, a parrot species with a song system that, based on anatomy alone, has been hypothesized to have a similar functional organization [25,31].

Surprisingly, we found that inactivation of the output of the AFP, the NAO/MO complex, in parrots had the opposite effect as observed in songbirds: the acoustic structure of learned calls degraded and became more variable. These findings raise the possibility that the parrot song system functions according to different principles.

In light of these results, we discuss the substantial differences that exist between songbirds and parrots at behavioral, vocal mechanical, and neural circuit levels. Behaviorally, parrot vocalizations retain outsized plasticity even in adulthood and may even convey semantic information [42,51–54]. For example, individuals learn each other's calls during courtship and group assimilation and budgerigars can even be trained to use their calls to label a color of an upcoming choice in a delayed-match-to-sample task [55]. Budgerigars can also, in their warble vocalizations not studied in this paper, learn to produce hundreds to thousands of specific English words [56], and can pronounce specific words differently across trials (e.g. more slowly or quickly, "Hello", versus "Heeeellllooooooo"). Such trial-by-trail flexibility in a word/syllable's pronunciation is commonplace in human speech but not observed in songbirds such as the zebra finch.

Parrots also differ at the level of vocal mechanics. Like humans, parrots vocalize with a single, bilaterally innervated, centered larynx and use their tongues to pattern sounds [28,57]. In contrast, songbirds pattern their sounds with bilaterally symmetric, unilaterally innervated syringes which are located at the top of each bronchus [58,59]. Finally, though the parrot and songbird song systems are both nucleated and contain distinct frontal and posterior pathways, there are also dramatic differences. First, each 'nucleus' of the parrot song system is recurrently interconnected with surrounding 'shell' tissue [27,31]. In songbirds, shells exist around nuclei, but are not substantially interconnected with the nuclei. Their roles in song learning are unclear, as they appear to be activated not by singing behavior but by general movement, for example, in trained hopping tasks [60,61]. Second, while RA axons specifically target vocalization-associated structures of the brainstem (e.g. respiratory nuclei and syringeal motor neurons) [62], AAC axons in the brainstem are more diffuse, raising the possibility that they control more than just vocalizations [28]. Third, recurrent long range feedback projections (most notably from AAC back to its inputs) abound in parrots (Figure 1B), while songbirds instead have primarily feedforward cortical pathways with recurrence only through a loop with the brainstem and thalamus (Figure 1A) [63]. This omnipresence of both short-range and long-range recurrent cortical feedback loops, a defining feature of the mammalian cortex [64,65], is absent in songbirds. Decades of theoretical work and more recent advances in machine learning suggest that, whereas feedforward networks are good at pattern

recognition and production, recurrent neural nets (RNNs) are suited for short-term memory storage and complex dynamics [64,65], raising the possibility that the human-like recurrent network architecture of the parrot vocal system is associated with parrots' outsized vocal capabilities. An important caveat of our results is that our methods only enabled us to inactivate MO and NAO at once. It remains possible that inactivation of only one of these structures would, like LMAN inactivation, preserve the learned acoustic structure of the call and reduce variability. Future studies with more specific viral or highly targeted lesion methods could in principle address these possibilities, but it's noteworthy that our experimental approach, including our drug concentrations, were the same as our past work in finches and yet yielded conceptually distinct results.

Going forward, these experiments set the stage for two key future directions. First, because frontal inactivations resulted in the loss of the temporally precise acoustic elements that define a bird's unique call, our findings suggest that at least one of the frontal nuclei exhibits temporally precise, call-locked discharge, as has been reported in songbird HVC but not LMAN [10,20]. Future electrophysiological recordings targeted to either MO or NAO will test if synfire chain-like activity important for songbird song temporal structure also exists in parrots—but in a distinct brain area with distinct connectivity. Second, though our study only examined the vocalizations of birds in isolation, it will be interesting to study the social dynamics of small groups of birds in which some birds have their vocal signatures degraded with frontal inactivations described here. This would test the ultimate function of these calls, for example if they are necessary for bond formation, maintenance, coordinated flight, and overall social cohesion.

Acknowledgements

We thank Michael Sheehan, Lindy McBride, Yumeng Zhang, and members of the Goldberg lab for comments. We thank Jack Bradbury and Sandra Vehrencamp for inspiration for the study.

Author contributions

Z.Z. and J.G. designed the study. Z.Z. collected the data with help from F.N. and B.K.. Z.Z., J.C., F.N., and H.K.T. analyzed the data. H.K.T. received supervision from I.C.. Z.Z. and J.G. wrote the manuscript with help from all other authors.

Declaration of interests

The authors declare no competing interests.

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be answered by the Lead Contact, Jesse H. Goldberg (jessehgoldberg@gmail.com).

Materials availability

This study did not generate any unique biological materials.

Data and code availability

Data and code generated in this study can be obtained through the Lead Contact.

Experimental model and subject details

Animal care and experiments were carried out in accordance with NIH guidelines and were approved by the Cornell Institutional Animal Care and Use Committee. Subjects used in this study were 10 male and 6 female budgerigars (*Melopsittacus undulatus*) that aged between 1 to 3 years. Birds were bred in our aviary or obtained from outside breeders (Magnolia Bird Farm, California; Petsmart, New York; OK Birds, Oklahoma). We used 8 birds for contact call quantification, 3 for determining anatomical coordinates, and 5 for chemical inactivation. Birds were group-housed in flight cages or individually housed in wire cages with a 12-12 light cycle. Food and water were provided *ad libitum*.

Method details

Contact call playback

We conducted two types of playback experiments. For the first type (Figure 2), we recorded the call repertoire of 8 birds (5 males and 3 females) in the same cage by playing colony noise on-and-off in a trial structure. Colony noise was first recorded with a portable recorder (V618, EVIDA) inside their flight cage for 2-3 days. We removed the silent portion of the recording and chopped it into 2-min segments. We then isolated individual birds to a wire cage inside a sound attenuating box (Figure 2A). A flat-spectrum speaker (iLoud Micro Monitor, IK Multimedia) was placed outside the box (~10 cm) and pointed toward an air-exchange hole on the box. We adjusted the speaker

volume so the playback sound measured 53-55 dB inside the box. For each trial, we randomly chose a colony noise segment and played it to the bird through the speaker. The inter-trial interval followed an exponential distribution (mean: 10 mins, range: 7.5-12 mins). Playback and response signals were recorded with directional microphones (Pro 35, Audio-Technica), bandpass-filtered (30-15,000 Hz) and amplified with a preamplifier (XL48, Midas), and digitized through a DAQ (USB6210, National Instruments) before sending to the computer. We wrote custom LabView codes to control the playback and recording. Each bird was recorded for 2 days.

For the second type of playback (Figure 3-4), we recorded responses to individual contact calls with the same hardware setup. We first constructed the call repertoire of each bird from its colony-noise responses and normalized the amplitude of all calls to the same value. For each trial, we randomly selected one caller and one call from its repertoire, then played it three times to the experimental bird with intervals uniformly sampled between 2-5 sec. The inter-trial interval followed an exponential distribution (mean: 5 mins, range: 4-6 mins).

Surgery and histology

To perform surgery and histology in the budgerigar, we adopted procedures that are commonly used in the zebra finch. All surgeries were performed under isoflurane anesthesia. Origin for the anterior/posterior and medial/lateral axis was established relative to the bifurcate blood vessel *lambda*. Brain surface after removing dura mater was used to zero the dorsal/ventral axis. We adjusted the head angle of the bird until the anterior skull was horizontal, which was referred to as the zero-degree angle. To access a brain region, we removed the first layer of skull with a cranial drill, the second layer of skull with a scalpel, and the dura mater with sharp forceps. After injection or implant, the exposed brain surface was sealed with biocompatible adhesive (World Precision Instruments, KWIK-CAST). Birds were allowed to recover for at least one week before any data collection. For histology, brains were fixed in 4% paraformaldehyde solution for 2-3 days, then cut into 100-μm coronal slices with a vibratome and mounted on slides in polyvinyl alcohol (Sigma-Aldrich, 10981-100ML). Dark-field and fluorescent images were acquired using a fluorescence scope (Leica DM4000 B).

Tracer injection

To determine the stereotactic coordinates of MO/NAO, we first calculated a rough estimate based on the published budgerigar brain atlas (http://www.brauthlab.umd.edu/atlas.htm). We then injected 100 nL fluorescent cholera toxin subunit B (CTB, C34775, Invitrogen) tracers into the brain and measured the targeting accuracy with histology after 7-9 days. The coordinates were iteratively improved with 6 injections in 3 birds.

Microdialysis probe implant

We made custom microdialysis probes according to the method in Aronov et al 2008. To target MO/NAO, the inflow polyimide tubing is approximately 2.25 mm long. The cellulose membrane measures roughly 0.2 mm in width and 1 mm in length, with 0.7 mm exposed to deliver drugs. The probes were implanted bilaterally at 8.80 AP, 2.90 ML and 1.45 DV (unit: mm), with a 30° AP angle and 20° ML angle. To secure the implants, we inserted 4 insect pins into the skull and covered the base of probes with dental cement (C&B Metabond, Parkell, S380). The probes were

flushed regularly with PBS to prevent clogging. After data collection, we infused fluorescein (10 mM) into the probe to visualize implant location and drug diffusion.

Pharmacological inactivation

Compared to direct injection, a higher concentration of drug is needed with the microdialysis probe, since the solute needs to diffuse passively through the cellulose membrane (Aronov et al 2008). Following previous studies in the zebra finch (Aronov et al 2008), we diluted tetrodotoxin citrate (Abcam, ab120055-1MG) in PBS to 50 μM. One week after surgery, we started infusing PBS into probes to acclimate the bird to the procedure. Infusion was conducted roughly 1 hour before the cage light turned on. For each probe, we infused 1.0 mL solution at a rate of ~0.2 mL/min. For TTX, we left the drug in the probe for 15 mins then washed it out with PBS. The bird was then placed back to the cage to rest before the light turned on and behavioral trials began. We used the same procedure for TTX and PBS, and conditions were alternated daily, so a direct comparison is meaningful. We had 2-3 infusion replicates for each bird under each condition. For experiments using muscimol, we diluted it to 1.5 mg/mL in PBS (Aronov et al 2008). Since muscimol degrades faster than TTX, to sample enough contact calls, we kept muscimol in the probes during data collection for about 6 hours, then washed it out with PBS.

Quantification and statistical analysis

Contact call identification

We wrote custom Matlab scripts to identify contact calls from audio recordings. Noise at low and high frequency was first removed with a bandpass Butterworth filter (500-10,000 Hz). We then calculated the spectral flatness (Wiener energy) and set a threshold (0.65) to distinguish vocal signal from broadband noise. Segments that passed the threshold were considered candidates and exported for visual inspection. Contact call is characteristic in having a major modulated frequency band between 2-4 kHz. Cage noise, squawks, harmonic calls, and warble were removed from downstream analysis.

Acoustic feature analysis

We used the SAT Matlab package from Sound Analysis Pro 2011 to calculate acoustic features of identified calls (http://soundanalysispro.com/matlab-sat). Pitch modulation depth was defined as (max - min) / mean of pitch. Spectrograms were calculated using the Short-Time Fourier Transform (STFT) algorithm in Matlab (window size 512, nfft 512). Variability of acoustic features was calculated based on the method in Olveczky et al 2005, to facilitate direct comparison to results in zebra finches. Briefly, we randomly sampled 50 pairs of calls from each bird under each condition. Since subject F6 had two normal call types (Figure 5B), we only sampled call pairs that belonged to the same type. We then used the ‘Similarity batch’ function in Sound Analysis Pro 2011 to calculate pairwise similarity scores with the default parameters except for the ‘min interval’, which was set to 25 ms. Following Olveczky et al 2005, values of ‘accuracy’ were used as the metric for similarity and further converted to variability scores.

414 **Variational auto-encoder (VAE)**

415 We followed the method in Goffinet et al 2021 to train the VAE network for contact call analysis.
416 Briefly, spectrograms of amplitude-normalized calls were centered in a 128*128 matrix and used
417 as the inputs. We performed resampling so each individual had the same number of calls, then
418 randomly split the dataset for training (80%) and testing (20%). The VAE architecture was adapted
419 from Goffinet et al 2021 with dimension of latent features fixed to 32. The network was trained for
420 100 epochs at learning rate 0.001 and batch size 32 with the Adam optimizer implemented in
421 Pytorch 1.12.1. We then calculated the latent features of each call by feeding it into the encoder
422 network. Separate networks were trained for the behavioral (Figure 2) and inactivation
423 experiments (Figure 3-5).

424 **Support vector machine (SVM)**

425 In past work, individual signatures in vocal signals were quantified with two major categories of
426 approaches: 1) methods based on information theory [66]; 2) classification models such as support
427 vector machine (SVM) [43,67]. Although information theoretic approaches allow comparison across
428 different systems and quantification of features' contribution to individuality, here our primary goal
429 was to explicitly compare discriminability of the same system across conditions, e.g. train a model
430 on the PBS dataset then check performance on the TTX dataset. We therefore adopted the
431 classification approach to quantify individuality in contact calls. Specifically, we built SVM
432 classifiers to decode individual identity from contact calls using the scikit-learn module in Python.
433 We first conducted principal component analysis (PCA) on the VAE outputs to obtain the most
434 informative features, then used them as inputs for SVM. We varied the dimension of principal
435 inputs and calculated how the decoding accuracy changed accordingly. For each model, we
436 sampled an equal number of calls from each individual to balance the dataset, then split it for
437 training (80%) and testing (20%). Accuracy was calculated only on the testing dataset. The
438 sampling procedure was repeated 50 times for each model. The mean accuracy and standard
439 deviation were plotted in the figures.

440 **Silhouette score calculation**

441 We calculated the Silhouette score to quantify the separability of calls produced by different
442 individuals. We first trained all calls produced on the PBS and TTX condition in a single VAE
443 network, then reduced the dimension of latent features using PCA. We focused on the first 10
444 principal components, since together they could explain 97.8% of the total variance. We
445 calculated the Silhouette score separately for the PBS and TTX dataset. We sampled 50 calls
446 from each individual, then used the scikit-learn module in Python to calculate the Silhouette score
447 on the first 10 principal components. The sampling was repeated 100 times to obtain a distribution
448 of Silhouette scores. Shuffling of labels was also performed at each sampling to obtain a null
449 distribution.

450

451

References

1. Jarvis, E.D. (2019). Evolution of vocal learning and spoken language. Science *366*, 50–54.

2. Petkov, C.I., and Jarvis, E.D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. Front. Evol. Neurosci. *4*, 12.

3. Tyack, P.L. (2020). A taxonomy for vocal learning. Philos. Trans. R. Soc. Lond. B Biol. Sci. *375*, 20180406.

4. ten Cate, C. (2021). Re-evaluating vocal production learning in non-oscine birds. Philos. Trans. R. Soc. Lond. B Biol. Sci. *376*, 20200249.

5. Sober, S.J., Wohlgemuth, M.J., and Brainard, M.S. (2008). Central contributions to acoustic variation in birdsong. J. Neurosci. *28*, 10370–10379.

6. Warren, T.L., Tumer, E.C., Charlesworth, J.D., and Brainard, M.S. (2011). Mechanisms and time course of vocal learning and consolidation in the adult songbird. J. Neurophysiol. *106*, 1806–1821.

7. Nottebohm, F., Stokes, T.M., and Leonard, C.M. (1976). Central control of song in the canary, Serinus canarius. J. Comp. Neurol. *165*, 457–486.

8. Aronov, D., Veit, L., Goldberg, J.H., and Fee, M.S. (2011). Two distinct modes of forebrain circuit dynamics underlie temporal patterning in the vocalizations of young songbirds. Journal of Neuroscience *31*, 16353–16368.

9. Dugas-Ford, J., Rowell, J.J., and Ragsdale, C.W. (2012). Cell-type homologies and the origins of the neocortex. Proc. Natl. Acad. Sci. U. S. A. *109*, 16974–16979.

10. Fee, M.S., Kozhevnikov, A.A., and Hahnloser, R.H. (2004). Neural mechanisms of vocal sequence generation in the songbird. Ann. N. Y. Acad. Sci. *1016*, 153–170.

11. Chen, J.R., Stepanek, L., and Doupe, A.J. (2014). Differential contributions of basal ganglia and thalamus to song initiation, tempo, and structure. J. Neurophysiol. *111*, 248–257.

12. Goldberg, J.H., and Fee, M.S. (2011). Vocal babbling in songbirds requires the basal ganglia-recipient motor thalamus but not the basal ganglia. J. Neurophysiol. *105*, 2729–2739.

13. Aronov, D., Andalman, A.S., and Fee, M.S. (2008). A specialized forebrain circuit for vocal babbling in the juvenile songbird. Science *320*, 630–634.

14. Picardo, M.A., Merel, J., Katlowitz, K.A., Vallentin, D., Okobi, D.E., Benezra, S.E., Clary, R.C., Pnevmatikakis, E.A., Paninski, L., and Long, M.A. (2016). Population-Level Representation of a Temporal Sequence Underlying Song Production in the Zebra Finch. Neuron *90*, 866–876.

15. Hamaguchi, K., Tschida, K.A., Yoon, I., Donald, B.R., and Mooney, R. (2014). Auditory synapses to song premotor neurons are gated off during vocalization in zebra finches. Elife *3*, e01833.

16. Fortune, E.S., Rodríguez, C., Li, D., Ball, G.F., and Coleman, M.J. (2011). Neural

489 mechanisms for the coordination of duet singing in wrens. Science *334*, 666–670.

490 17. Hoffmann, S., Trost, L., Voigt, C., Leitner, S., Lemazina, A., Sagunsky, H., Abels, M.,
491 Kollmansperger, S., Maat, A.T., and Gahr, M. (2019). Duets recorded in the wild reveal that
492 interindividually coordinated motor control enables cooperative behavior. Nat. Commun. *10*,
493 2577.

494 18. Bottjer, S.W., Miesner, E.A., and Arnold, A.P. (1984). Forebrain lesions disrupt
495 development but not maintenance of song in passerine birds. Science *224*, 901–903.

496 19. Scharff, C., and Nottebohm, F. (1991). A comparative study of the behavioral deficits
497 following lesions of various parts of the zebra finch song system: implications for vocal
498 learning. J. Neurosci. *11*, 2896–2913.

499 20. Kao, M.H., Doupe, A.J., and Brainard, M.S. (2005). Contributions of an avian basal ganglia-
500 forebrain circuit to real-time modulation of song. Nature *433*, 638–643.

501 21. Olveczky, B.P., Andalman, A.S., and Fee, M.S. (2005). Vocal experimentation in the
502 juvenile songbird requires a basal ganglia circuit. PLoS Biol. *3*, e153.

503 22. Edelman, G.M. (1993). Neural Darwinism: selection and reentrant signaling in higher brain
504 function. Neuron *10*, 115–125.

505 23. Ben-Pazi, H., Gross-Tsur, V., Bergman, H., and Shalev, R.S. (2003). Abnormal rhythmic
506 motor response in children with attention-deficit-hyperactivity disorder. Dev. Med. Child
507 Neurol. *45*, 743–745.

508 24. Singh Alvarado, J., Goffinet, J., Michael, V., Liberti, W., III, Hatfield, J., Gardner, T.,
509 Pearson, J., and Mooney, R. (2021). Neural dynamics underlying birdsong practice and
510 performance. Nature *599*, 635–639.

511 25. Paton, J.A., Manogue, K.R., and Nottebohm, F. (1981). Bilateral organization of the vocal
512 control pathway in the budgerigar, Melopsittacus undulatus. J. Neurosci. *1*, 1279–1288.

513 26. Brauth, S.E., Heaton, J.T., Durand, S.E., Liang, W., and Hall, W.S. (1994). Functional
514 anatomy of forebrain auditory pathways in the budgerigar (Melopsittacus undulatus). Brain
515 Behav. Evol. *44*, 210–233.

516 27. Striedter, G.F. (1994). The vocal control pathways in budgerigars differ from those in
517 songbirds. J. Comp. Neurol. *343*, 35–56.

518 28. Brauth, S.E., Heaton, J.T., Shea, S.D., Durand, S.E., and Hall, W.S. (1997). Functional
519 anatomy of forebrain vocal control pathways in the budgerigar (Melopsittacus undulatus).
520 Ann. N. Y. Acad. Sci. *807*, 368–385.

521 29. Durand, S.E., Heaton, J.T., Amateau, S.K., and Brauth, S.E. (1997). Vocal control
522 pathways through the anterior forebrain of a parrot (Melopsittacus undulatus). J. Comp.
523 Neurol. *377*, 179–206.

524 30. Jarvis, E.D., and Mello, C.V. (2000). Molecular mapping of brain areas involved in parrot
525 vocal communication. J. Comp. Neurol. *419*, 1–31.

526 31. Chakraborty, M., Walløe, S., Nedergaard, S., Fridel, E.E., Dabelsteen, T., Pakkenberg, B.,

Bertelsen, M.F., Dorrestein, G.M., Brauth, S.E., Durand, S.E., et al. (2015). Core and Shell Song Systems Unique to the Parrot Brain. PLoS One *10*, e0118496.

32. Heaton, J.T., and Brauth, S.E. (2000). Effects of lesions of the central nucleus of the anterior archistriatum on contact call and warble song production in the budgerigar (Melopsittacus undulatus). Neurobiol. Learn. Mem. *73*, 207–242.

33. Lavenex, P.B. (2000). Lesions in the budgerigar vocal control nucleus NLc affect production, but not memory, of english words and natural vocalizations. J. Comp. Neurol. *421*, 437–460.

34. Tchernichovski, O., Lints, T.J., Deregnaucourt, S., Cimenser, A., and Mitra, P.P. (2004). Studying the song development process: rationale and methods. Ann. N. Y. Acad. Sci. *1016*, 348–363.

35. Kollmorgen, S., Hahnloser, R.H.R., and Mante, V. (2020). Nearest neighbours reveal fast and slow components of motor learning. Nature *577*, 526–530.

36. Goffinet, J., Brudner, S., Mooney, R., and Pearson, J. (2021). Low-dimensional learned feature spaces quantify individual and group differences in vocal repertoires. Elife *10*. 10.7554/eLife.67855.

37. Farabaugh, S.M., and Dooling, R.J. (1996). Acoustic Communication in Parrots: Laboratory and Field Studies of Budgerigars, *Melopsittacus undulatus*. In Ecology and Evolution of Acoustic Communication in Birds.

38. Brown, S.D., Dooling, R.J., and O'Grady, K.E. (1988). Perceptual organization of acoustic stimuli by budgerigars (Melopsittacus undulatus): III. Contact calls. J. Comp. Psychol. *102*, 236–247.

39. Farabaugh, S.M., Linzenbold, A., and Dooling, R.J. (1994). Vocal plasticity in budgerigars (Melopsittacus undulatus): evidence for social factors in the learning of contact calls. J. Comp. Psychol. *108*, 81–92.

40. Hile, A.G., Plummer, T.K., and Striedter, G.F. (2000). Male vocal imitation produces call convergence during pair bonding in budgerigars, Melopsittacus undulatus. Anim. Behav. *59*, 1209–1218.

41. Hile, A.G., and Striedter, G.F. (2000). Call Convergence within Groups of Female Budgerigars (Melopsittacus undulatus). Ethology *106*, 1105–1114.

42. Dahlin, C.R., Young, A.M., Cordier, B., Mundry, R., and Wright, T.F. (2014). A test of multiple hypotheses for the function of call sharing in female budgerigars, Melopsittacus undulatus. Behav. Ecol. Sociobiol. *68*, 145–161.

43. Tachibana, R.O., Oosugi, N., and Okanoya, K. (2014). Semi-automatic classification of birdsong elements using a linear support vector machine. PLoS One *9*, e92584.

44. Brockway, B.F. (1964). Ethological studies of the budgerigar (Melopsittacus undulatus): Non-reproductive behavior. Behaviour *22*, 193–222.

45. Brockway, B.F. (1964). Ethological studies of the budgerigar: reproductive behavior. Behaviour *23*, 294–323.

46. Wyndham, E. (1980). Diurnal cycle, behaviour and social organization of the budgerigar Melopsittacus undulatus. Emu - Austral Ornithology *80*, 25–33.

47. Otchy, T.M., Wolff, S.B., Rhee, J.Y., Pehlevan, C., Kawai, R., Kempf, A., Gobes, S.M., and Olveczky, B.P. (2015). Acute off-target effects of neural circuit manipulations. Nature *528*, 358–363.

48. Andalman, A.S., and Fee, M.S. (2009). A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. Proc. Natl. Acad. Sci. U. S. A. *106*, 12518–12523.

49. Goldberg, J.H., and Fee, M.S. (2012). A cortical motor nucleus drives the basal ganglia-recipient thalamus in singing birds. Nat. Neurosci. *15*, 620–627.

50. Ölveczky, B.P., Otchy, T.M., Goldberg, J.H., Aronov, D., and Fee, M.S. (2011). Changes in the neural control of a complex motor sequence during learning. J. Neurophysiol. *106*, 386–397.

51. Hobson, E.A., Avery, M.L., and Wright, T.F. (2014). The socioecology of Monk Parakeets: Insights into parrot social complexity. The Auk: Ornithological Advances *131*, 756–775.

52. Ikkatai, Y., Watanabe, S., and Izawa, E.-I. (2016). Reconciliation and third-party affiliation in pair-bond budgerigars (Melopsittacus undulatus). Behaviour *153*, 1173–1193.

53. Berg, K.S., Delgado, S., Cortopassi, K.A., Beissinger, S.R., and Bradbury, J.W. (2012). Vertical transmission of learned signatures in a wild parrot. Proc. Biol. Sci. *279*, 585–591.

54. Balsby, T.J.S., Momberg, J.V., and Dabelsteen, T. (2012). Vocal imitation in parrots allows addressing of specific individuals in a dynamic communication network. PLoS One *7*.

55. Manabe, K., Kawashima, T., and Staddon, J.E.R. (1995). Differential vocalization in budgerigars: Towards an experimental analysis of naming. J. Exp. Anal. Behav. *63*, 111–126.

56. Words, Puck the Parakeet speaks 1728 English (1995). Guinness world records (Guinness World Records).

57. Beckers, G.J.L., Nelson, B.S., and Suthers, R.A. (2004). Vocal-tract filtering by lingual articulation in a parrot. Curr. Biol. *14*, 1592–1597.

58. Goller, F., and Cooper, B.G. (2004). Peripheral motor dynamics of song production in the zebra finch. Ann. N. Y. Acad. Sci. *1016*, 130–152.

59. During, D.N., Ziegler, A., Thompson, C.K., Faber, C., Muller, J., Scharff, C., and Elemans, C.P. (2013). The songbird syrinx morphome: a three-dimensional, high-resolution, interactive morphological map of the zebra finch vocal organ. BMC Biol. *11*, 1.

60. Bottjer, S.W., Brady, J.D., and Cribbs, B. (2000). Connections of a motor cortical region in zebra finches: relation to pathways for vocal learning. J. Comp. Neurol. *420*, 244–260.

61. Feenders, G., Liedvogel, M., Rivas, M., Zapka, M., Horita, H., Hara, E., Wada, K., Mouritsen, H., and Jarvis, E.D. (2008). Molecular mapping of movement-associated areas in the avian brain: a motor theory for vocal learning origin. PLoS One *3*, e1768.

604 62. Wild, J.M. (1993). Descending projections of the songbird nucleus robustus archistriatalis.
605 J. Comp. Neurol. *338*, 225–241.

606 63. Schmidt, M.F., Ashmore, R.C., and Vu, E.T. (2004). Bilateral control and interhemispheric
607 coordination in the avian song motor system. Ann. N. Y. Acad. Sci. *1016*, 171–186.

608 64. Sussillo, D., and Abbott, L.F. (2009). Generating coherent patterns of activity from chaotic
609 neural networks. Neuron *63*, 544–557.

610 65. Sanchez-Vives, M.V., and McCormick, D.A. (2000). Cellular and network mechanisms of
611 rhythmic recurrent activity in neocortex. Nat. Neurosci. *3*, 1027–1034.

612 66. Beecher, M.D. (1989). Signaling systems for individual recognition: an information theory
613 approach. Anim. Behav. *38*, 248–261.

614 67. Gelfand, D.L., and McCracken, G.F. (1986). Individual variation in the isolation calls of
615 Mexican free-tailed bat pups (Tadarida brasiliensis mexicana). Anim. Behav. *34*, 1078–
616 1086.

617

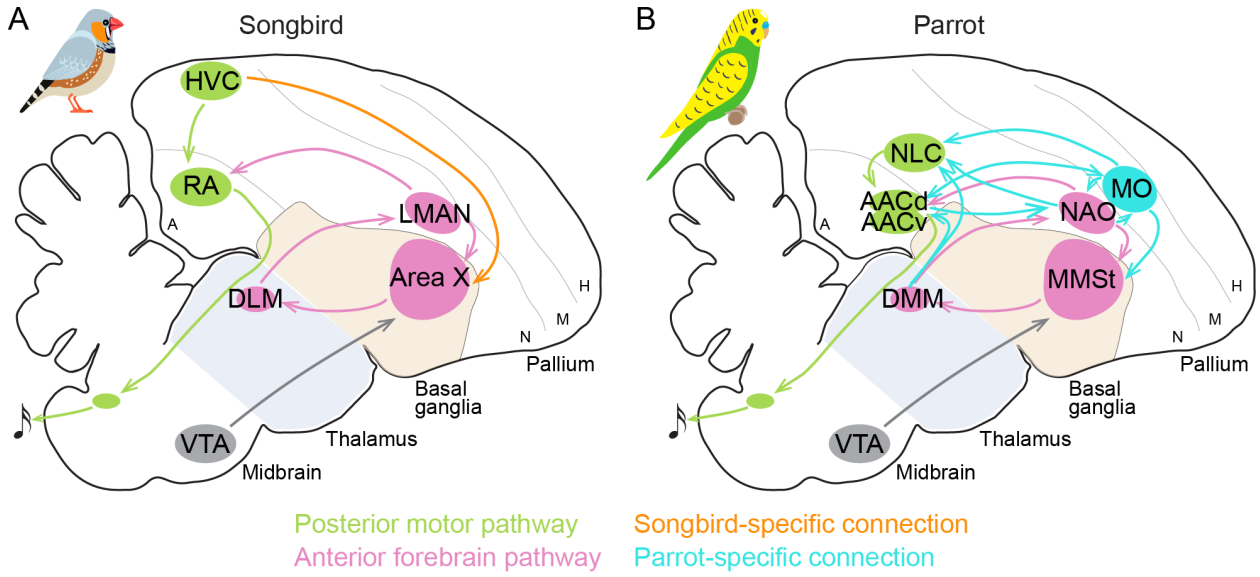

Figure 1. Vocal learning circuits in songbirds and parrots

(A) Sagittal view showing the posterior motor pathway (green) and anterior forebrain pathway (pink) in the zebra finch. Orange line indicates the songbird-specific connection from HVC to Area X. Abbreviations: H, hyperpallium; M, mesopallium; N, nidopallium; A, arcopallium; RA, robust nucleus of the arcopallium; HVC, proper name; LMAN, lateral magnocellular nucleus of the anterior nidopallium; DLM, medial portion of the dorsolateral thalamus; VTA, ventral tegmental area.

(B) Vocal learning circuits in budgerigars also have the posterior (green) and anterior (pink) pathways. Cyan circles and lines indicate the parrot-specific vocal nuclei and connections. The shell structures that surround the vocal nuclei are not depicted. Abbreviations: AAC, central nucleus of the anterior arcopallium; NLC, central nucleus of the lateral nidopallium; MO, oval nucleus of the anterior mesopallium; NAO, oval nucleus of the anterior nidopallium; MMSt, magnocellular nucleus of the medial striatum; DMM, magnocellular nucleus of the dorsal thalamus.

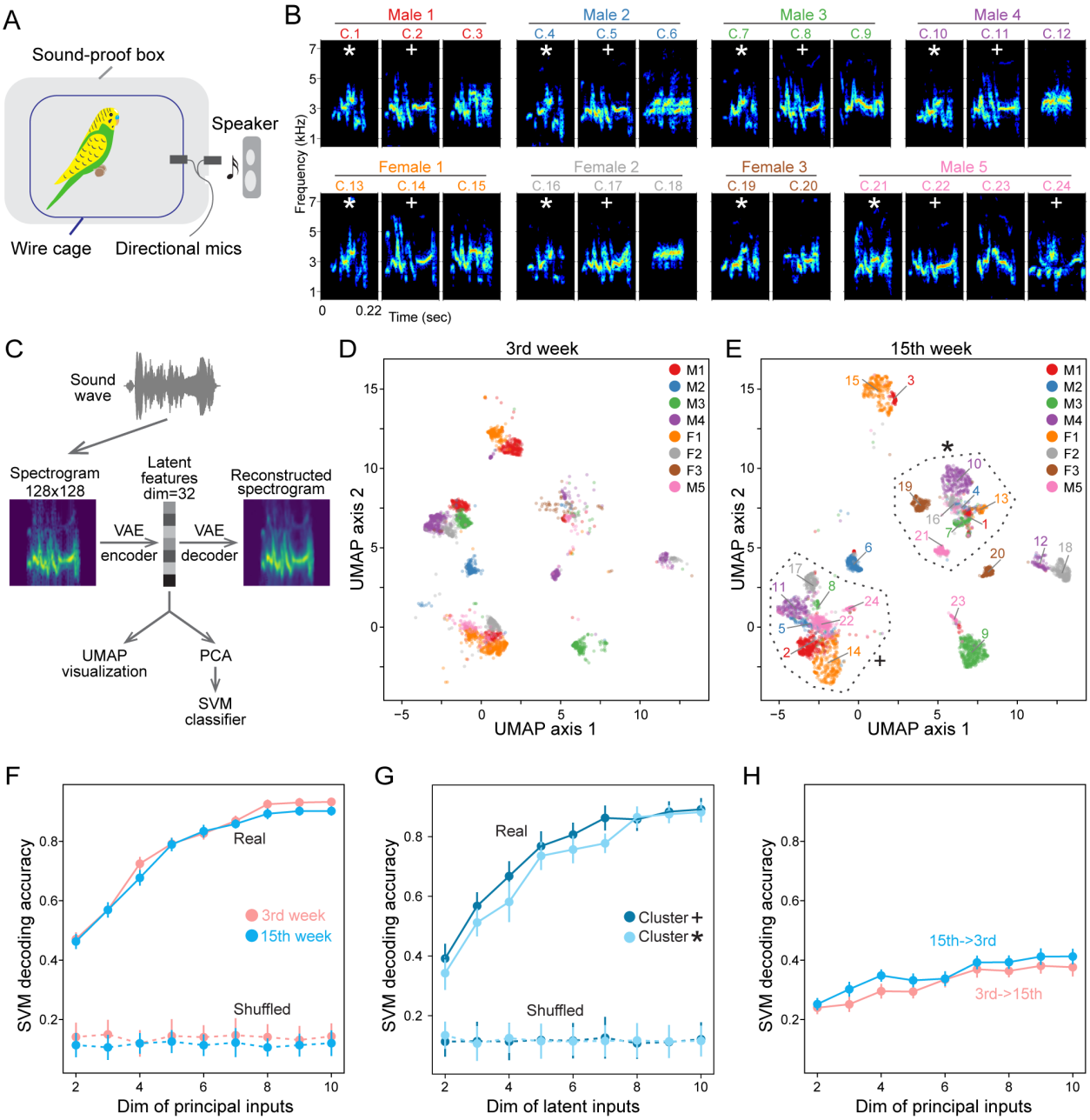

Figure 2. Budgerigar contact calls contain individual signature

(A) Schematics showing the recording setup. Two directional microphones point in the opposite direction to capture sound from inside or outside the box.

(B) Spectrograms of example contact calls from a group of 8 budgerigars (5 males, 3 females) at week 15 since assimilation. All plots have the same scale on the x- and y- axis. Animal and cluster labels above each spectrogram correspond to labels in panel *E*. Asterisk symbol marks a call type that was shared across all 8 individuals, while plus symbol marks one that was shared across 7 individuals.

(C) Pipeline for analyzing contact calls.

(D-E) UMAP embedding of contact calls at week 3 (D) and week 15 (E) since group assimilation. Two datasets were trained through the same VAE and embedded in the same UMAP space, but shown separately for clarity. Each dot is a call, colored by caller identity. A maximum of 600 calls were shown for each individual. Animal and cluster labels in (E) correspond to spectrograms in panel *B*. Dashed lines in (E) circled the two major clusters that were shared across individuals and marked by an asterisk or plus symbol in panel *B*.

(F) Accuracy of SVM classifiers that decode individual identity from call latent features as a function of input dimensions. Solid lines are models trained on real data, while dashed lines are models trained on shuffled data.

(G) Accuracy of SVM classifiers that were trained and tested solely on the calls that belong to the two large shared clusters (marked by asterisk and plus symbol in panel B and E).

(H) Accuracy of SVM classifiers that were trained on one time point and used to decode individual identity at a different time point.

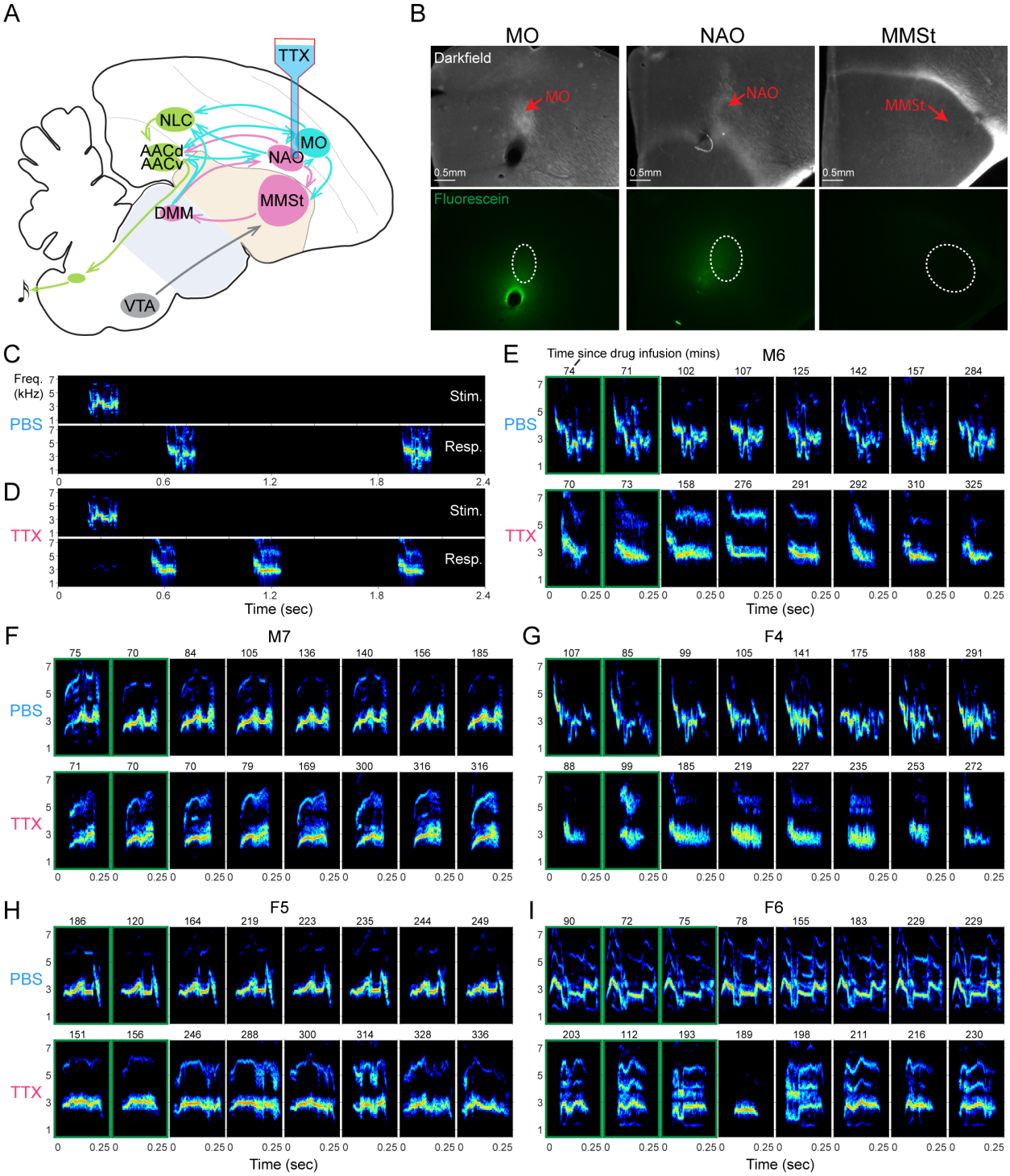

Figure 3. Frontal cortical inactivation abolishes pitch modulation in contact calls

(A) Schematics of the inactivation experiments. Reverse microdialysis probes were implanted in the center of the MO/NAO complex and infused with either 50 micromolar TTX or PBS.

(B) Histological confirmation of implant site in one representative bird. At the end of experiments fluorescein was infused into the probes to visualize diffusion and probe placement. Fluorescein

covered both MO and NAO. Dashed circles at the bottom row indicate the boundaries of corresponding nuclei.

(C-D) Example trials from one bird under the PBS (C) and TTX (D) condition. Top rows show spectrograms of the stimuli, bottom rows show time-aligned responses. Y-axis is frequency in kHz.

(E-I) Example spectrograms from 5 birds under the PBS (top row) and TTX (bottom row) condition. The first 2-3 spectrograms at each row in the green boxes are the first calls produced during each PBS or TTX infusion replicate. Other spectrograms were selected randomly. M6-M7, males; F4-F6, females. Y-axis is frequency in kHz. Numbers above each spectrogram indicate minutes since drug infusion.

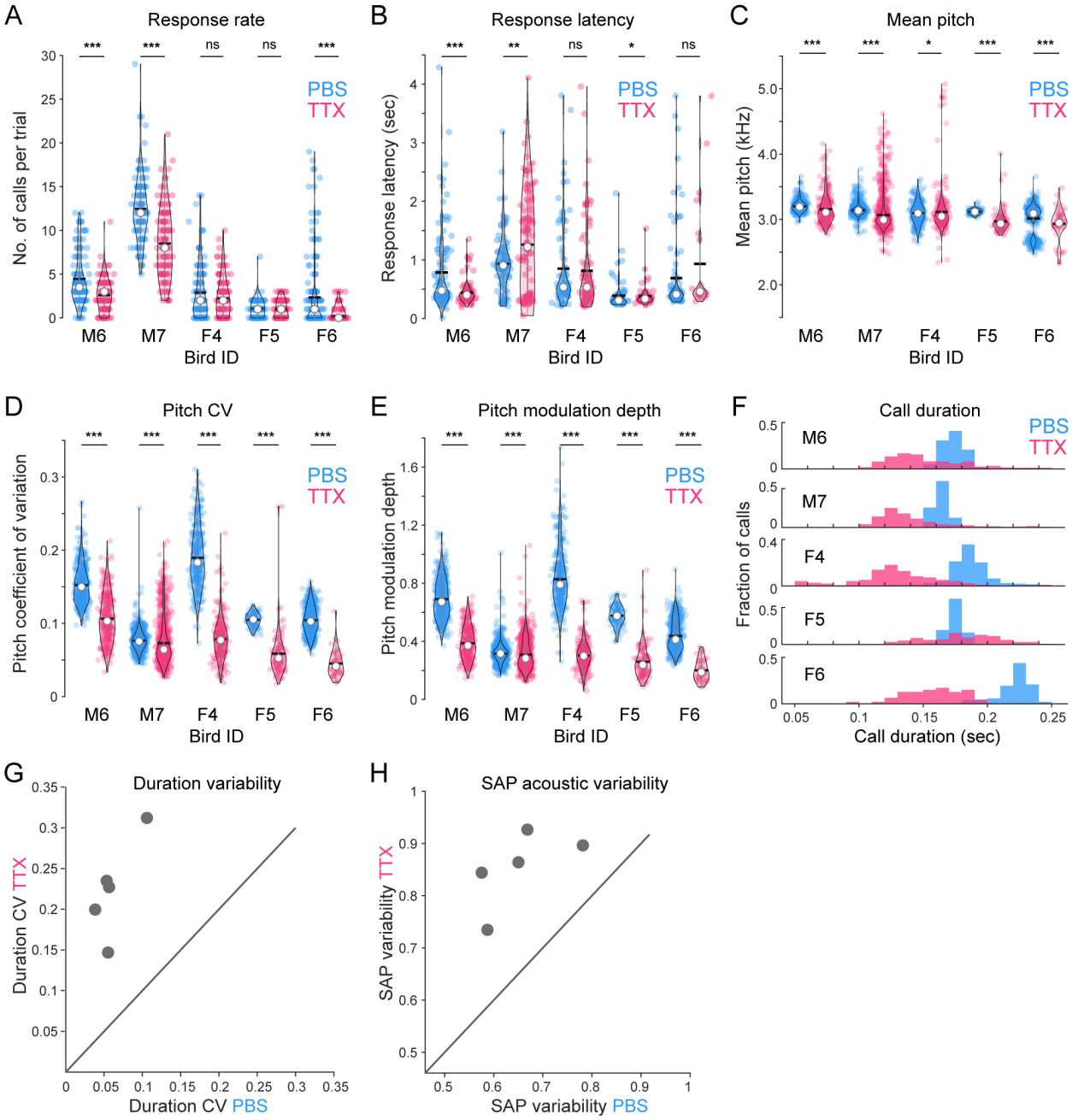

Figure 4. Frontal cortical inactivation increases variability of contact calls

(A-E) Compare behavioral trial statistics (A-B) and acoustic features (C-E) between the PBS (blue) and TTX (red) condition. Each colored dot is one trial (A-B) or one call (C-E). The violins indicate the shape of distribution. The black line and white dot indicate the mean and median. P-value: *<0.05, **<0.01, ***<0.001, ns not significant, with Mann-Whitney U tests.

(F) Histograms of contact call duration under the PBS (blue) and TTX (red) condition.

(G) Variability of call duration quantified by coefficient of variation (CV). Each dot indicates data from one bird.

684 (H) Variability of acoustic features calculated by Sound Analysis Pro (SAP). Each dot indicates
685 the averaged dissimilarity of 50 randomly selected pairs of calls in a bird under the PBS or TTX
686 condition.

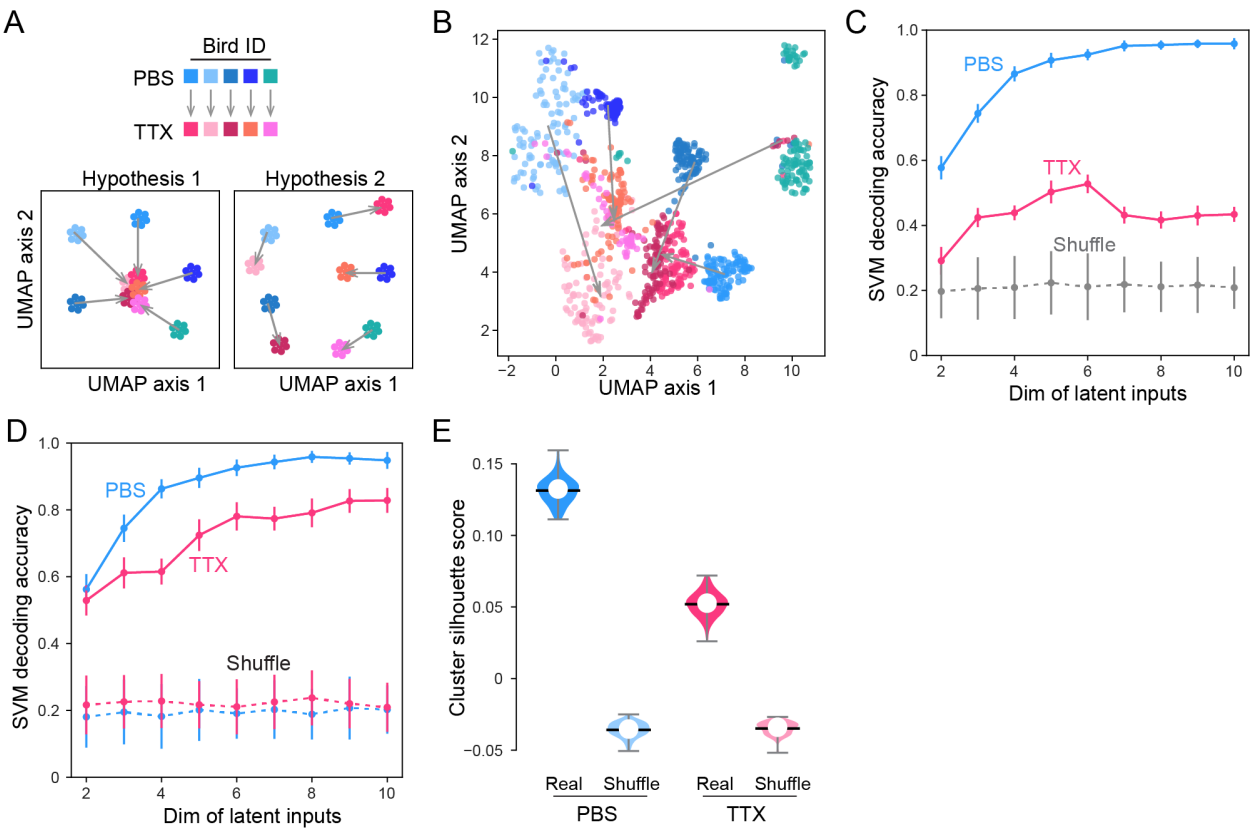

Figure 5. Frontal cortical inactivation reduces individual signatures in contact calls.

(A) Hypotheses on how AFP inactivation would change the embedding of contact calls in the UMAP space. Clusters are colored according to individual identity and experimental condition. Gray arrows point from PBS to TTX condition in the same birds.

(B) UMAP embedding of contact calls. Each dot indicates a call, colored by caller identity and experiment condition. Gray arrows point from the centroid of PBS cluster to the centroid of TTX cluster for each bird.

(C) SVM classifier accuracy as a function of input dimensions. The model was trained on the PBS condition, then used to decode individual identity at the TTX condition. Dots indicate the mean accuracy, vertical lines indicate the standard deviation.

(D) Same as (C), but two models were trained separately for PBS and TTX.

(E) Separability of contact calls produced by different individuals at the PBS and TTX condition, as quantified by Silhouette score (Methods). Equal number of calls were sampled from each individual for the analysis. The results of 100 sampling are plotted. Colored violins indicate the shape of Silhouette score distribution. Black lines and white dots indicate the mean and median values, respectively.

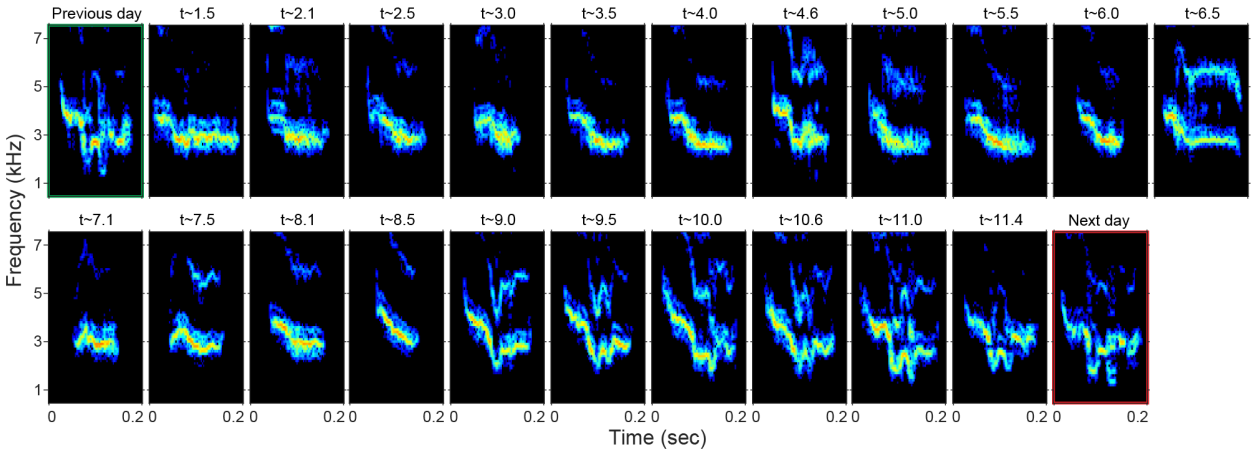

Figure S1. Contact calls gradually recover from AFP inactivation

Spectrograms of contact calls from one example TTX infusion trial. The green and red boxes show one call spectrogram from the same bird the day before and after TTX infusion, respectively. Numbers above other spectrograms indicate the time since TTX infusion in the unit of hour. The contact calls began to recover gradually to normal shape after 9-10 hours.

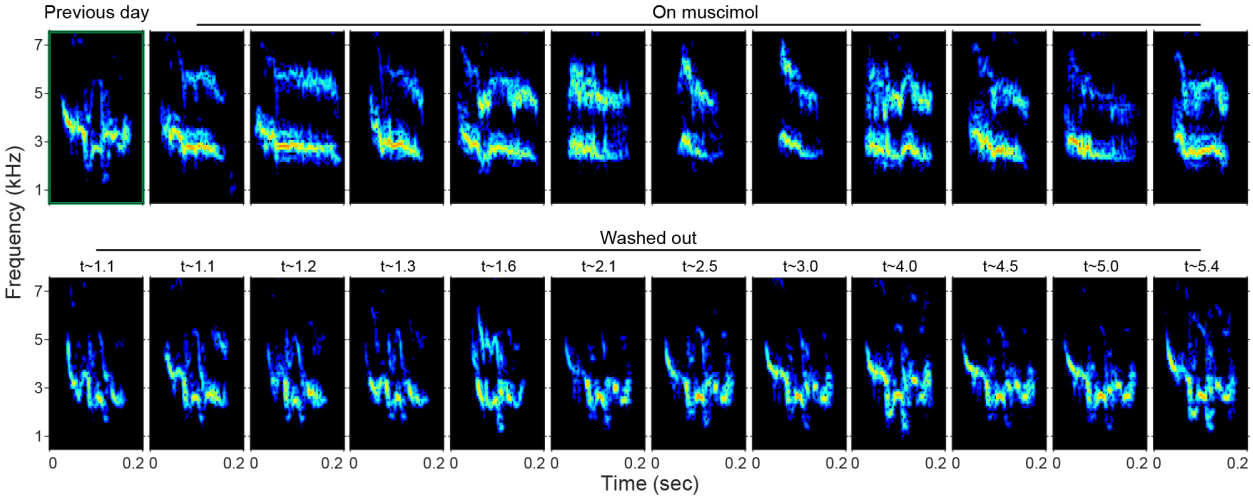

Figure S2. Inactivation by muscimol had a similar effect on contact call acoustic structure

Spectrograms of contact calls from one example muscimol infusion trial. The green box shows one call spectrogram from the same bird the day before muscimol infusion. Numbers above spectrograms at the bottom row indicate the time in hours since muscimol was washed out with saline.