A neural clock underlying the temporal dynamics of an auditory memory

by

Andrew H. Bahle

Submitted to the Department of Brain and Cognitive Sciences
in partial fulfillment of the requirements for the degree of

PHD IN NEUROSCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2024

Authored by: Andrew H. Bahle
Department of Brain and Cognitive Sciences
May 3, 2024

Certified by: Michale S. Fee
Professor of Brain and Cognitive Sciences, Thesis Supervisor

Accepted by: Mark Harnett
Professor of Brain and Cognitive Sciences
Graduate Officer, Department of Brain and Cognitive Sciences

A neural clock underlying the temporal dynamics of an auditory memory

by

Andrew H. Bahle

Submitted to the Department of Brain and Cognitive Sciences
on May 3, 2024 in partial fulfillment of the requirements for the degree of

PHD IN NEUROSCIENCE

ABSTRACT

Imitation is an essential hallmark of intelligent systems. Children imitate the speech, body language and expressions of adults, eventually graduating to creative expressions of their own individual thoughts and ideas. In machine intelligence and A.I., large language models have recently demonstrated a striking ability to convincingly imitate written forms of human language, from observation of massive corpora of text. A fundamental question is how these varied intelligent systems achieve such robust imitation. In animals, imitation is accomplished by complex neural circuits in the brain. To perform imitation, animals must first represent the sensory consequences of the action to be imitated and store this representation as a memory. Next, they must recall this sensory memory, evaluating their imitation attempts until a satisfactory match is achieved. In this thesis, I study the neural control of vocal imitation in the songbird *Taengiopia guttata*, focusing on the first stage of imitation when animals must form a temporally structured sensory memory, or template, of the action to be imitated. In the first chapter, I present work attempting to localize the brain regions involved in the formation of the sensory memory used in imitation. We provide evidence that HVC, a pre-motor region that controls the timing of adult song, is involved in storing the timing of the tutor memory. This works shows how focal cooling can be used to study the formation of temporally structured memories even in the absence of overt behavior. In chapter 2, we ask what neural dynamics support the observed effect of cooling on the imitation. Using freely moving calcium imaging and head-fixed high-throughput electrophysiology, we show that tutoring evokes sparse sequential activity in HVC, reminiscent of its activity during adult production of the vocal imitation. This activity was present as early as the very first day of tutoring, perhaps indicating that HVC connectivity is innately predisposed to produce sparse sequential representations of song. In the final chapter, we explore changes in the representation of the tutor song before and after tutoring. We observe the emergence of tutor selective neural responses in HVC after tutoring, and quantify this selectivity at the population level, in different cell-types. We further show that this tutor song selectivity is stronger in HVC than any of its auditory inputs, suggesting that tutor song selectivity results from the storage of a tutor memory in HVC itself. Together this work shows how HVC neural dynamics can act as a clock for the storage and recall of an auditory memory and gives insight into how memories containing temporal structure might be stored more broadly.

Thesis supervisor: Michale S. Fee
Title: Professor of Brain and Cognitive Sciences

Acknowledgments

I would like to start by acknowledging the many people in my life that have contributed directly and indirectly to my time in graduate school. First to my family who were so supportive of my ever changing interests and encouraged me to pursue whatever I was most curious about. My mother for encouraging me to make my own path and providing an environment where I felt free to explore and be creative. My father for showing me how to be gentle and obsessive. Growing up in a very small town, I had lots of nearby family. Thanks to the aunts and uncles who always let me borrow as many books as I wanted, and took a deep interest in all that I did.

To my partner Sophia. Thank you for being so incredibly easy to be with, and for laughing with me so much during the hard and annoying parts of life. These difficult last years of grad school were possible thanks to you. You often joke that I am the smart one and you are the funny one. But I actually think I am quite funny... and you are very often smarter than me... May I eventually beat you in scrabble twice in a row.

Thanks to the many mentors I have had scientifically, from elementary school to now. I am especially grateful to my first research mentor Orie Shafer, who trusted a music performance student to do research, and let me attempt any project that interested me. Thanks also to my Fulbright mentors Edvard and May-Britt Moser for their genuine care and guidance. They showed me that caring deeply about science doesn't mean giving up on caring just as deeply about the people doing it.

Of course I would also like to thank my mentor Michale Fee. When I first met him I immediately knew Michale and I would have fun doing science together. He embodies so many traits I hope to emulate: he demands clarity, specificity and rigour and only works on problems he sees as absolutely fundamental and deeply important. He taught me so much about how to think, and talk about science and it is clear why he is revered in the neuroscience community as much for his style of thinking as for his empirical findings. Nonetheless Michale is kind, funny and warm, traits which are too often absent in science. Perhaps we shouldn't choose presidents based on who you would like to have a beer with, but in scientific mentorship this might be ok.

Thanks so much to all of my wonderful and smart committee members.

I felt very lucky when I found out Ila was joining the department and even luckier when she agreed to be on my committee before even arriving in Cambridge. During my brief stint as a grid-cell jockey, I of course developed a healthy reverence for her thoughtful theoretical work. Thank you so much Ila for chairing my committee and for all of your great advice.

During the early days of my PhD I always wanted to an excuse to talk to Josh McDermott but I didn't have much of a reason, being so solidly on team systems neuroscience.

Fortunately, I came up with something. During the interview weekend I attended at MIT, I saw a talk on the schedule (amidst all the neural circuits and optogenetics stuff) that I thought probably wouldn't interest me much. That talk was Josh's and it was best talk of the day. Thanks for helping me to think broadly about the questions of audition, perception and computation, even in our modest little bird.

My first year of graduate school, Vivek gave a truly electrifying talk on the fly central complex and I was lucky enough to be one of the students selected to join him for dinner afterwards. On the way to dinner Vivek imparted a friendly challenge that has stuck with me throughout my PhD. He said, "Oh so you want to work with Michale Fee. Are you guys going to finally figure our where in the Bird's song is kept?". Vivek, I present our humble attempt.

To the Fee lab, thanks to all the current and past members. More than colleagues, you have all been an incredible group of friends and a great academic family. I have learned so much from you all and feel very grateful that I landed amongst such a kind, and smart bunch. Its hard to overstate how much my conversations with Fee lab members have shaped all the work in this thesis. A special thank you to Emily Mackevicius for our many discussions about models of tutor memory formation and to Galen Lynch for patiently helping me with my many technical questions. I am immensely grateful to you both for being so welcoming to me early on, even when I was an annoying and overzealous first year graduate student (an unfortunate occupational hazard).

When I moved to Cambridge, I was unsure how to find friends that shared my non scientific interests. Luckily I found many fascinating and fun people in Cambridge with incredibly diverse interests and delightfully strong opinions. I learned at least as much from talking with the brilliant friends I made here as I did in my coursework and thesis studies. Cambridge is unique, sometimes bizarre, often insular and elitist, but it has the effect of concentrating so many intensely curious minds in one place that you cant help but feel lucky to spend time in it. Thanks to the many wonderful, kind and chaotic friends that helped me weather grad school as well as I did.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Motor behaviors must be adapted to the world in which an animal lives. Innate behaviors have been adapted over millions of years of evolution, and are produced by neural circuits assembled using genetic and developmental programs. In contrast, learned behaviors are flexible and adapted to the distinctive experiences of individual animals during their lifetimes. In social species, imitation is a powerful strategy that allows animals to learn not only from their own experience, but the collective experiences of other individuals. Indeed, many of our most sophisticated and complex behaviors such as speech, language, cultural practices, music and athletic performance are learned by imitating others.

The songbird—in particular, the zebra finch (Taeniopygia guttata)—is a powerful model system in which to develop a mechanistic understanding of how neural circuits learn and produce a sophisticated imitated behavior. Male zebra finches sing a complex vocalization—a stereotyped sequence of 3-7 song syllables comprising an approximately one second long motif—used primarily to attract a mate [1, 2, 3, 4]. This song is learned, and emerges through a series of stages: first, a young male bird hears the song of a tutor, usually the father, and stores a memory of this song (called a template); the young bird then generates highly variable babbling vocalizations, called subsong; finally, by comparing its own variable practice vocalizations to the memorized tutor song, the young bird gradually refines its

vocalizations, frequently producing a close match to the tutor song (figure 1.1) [1, 3]. This learning process is highly structured and has often been compared to speech acquisition in humans [5, 6, 7, 8, 9].

When I began my PhD, it was clear that the songbird field was at a remarkably exciting stage. While many missing details and areas of contention remain, there now exist broadly plausible circuit level mechanistic hypotheses for most aspects of song learning. At the level of song production, neural sequences models are an excellent candidate mechanism with strong experimental and theoretical support. There is also a developing understanding of circuits for variability generation and dopaminergic evaluation signals, long hypothesized to implement reinforcement learning during juvenile vocal practice [10, 11, 12, 13, 14, 15, 16, 17, 18]. Finally, a descending pathway from the auditory cortex to the dopaminergic system has been shown to convey error signals during singing, which provides an important step towards identifying where and how song evaluation occurs [19]. While song-learning may not be solved *in toto* for many years, there is a sense that many of the big picture questions now have tentative answers and/or plausible theoretical explanations. This progress allows for the exciting possibility of proposing more or less "complete" circuit hypotheses for how the whole system might work. One key remaining problem however, is the song template; how is the auditory memory of the tutor song stored and recalled during learning.

One of the most remarkable recent findings in the songbird field is that HVC, a vocal motor center that acts as a clock to control song timing, is also critical for tutor memory storage [20, 21, 22]. This work falls on the heels of a long literature describing birds-own-song (BOS) and tutor selective auditory responses in HVC [23, 25, 26, 27, 28, 29, 30, 31, 32, 24]. But what is the connection between a motor timing circuit and the storage of an auditory memory? In this thesis, I attempt to test the underlying hypothesis that, in addition to its role in motor timing, HVC also functions as the 'clock' that stores the timing of the auditory memory. In the following chapters I propose a detailed mechanistic circuit model of how HVC mediates storage of the content of this memory during tutoring, recalls this

content during vocal practice, and the ways in which error signals are computed to guide reinforcement learning. This work is significant because it tests a circuit model that explains the link between two important but distinct bodies of work in the songbird field: the motor and auditory functions of HVC. Given the numerous analogies between avian and human vocal learning circuitry, I believe this work will have broad implications for understanding imitation learning, a key aspect of human cognition.

As described above, this thesis aims to develop an understanding of what happens during the tutoring process by studying tutor memory formation *as it occurs* in juvenile birds. This long standing and difficult problem, sometimes called the "holy grail" of the field [33], has been pursued by generations of neuroscientists. Their immense work on the anatomy, physiology and behavior of songbirds has given us both incredible privilege and necessary constraint, allowing us to focus our attention on only the most important avenues. Here I briefly review key parts of this literature which provide necessary context for understanding my own contribution. More specifically, I review the song system and its relevance for adult song production and juvenile motor learning, basal ganglia cortical-thalamic pathways and their hypothesized role in reinforcement learning, and the auditory system of the songbird especially as it pertains to auditory representation of the tutor song, the birds own song, as well as deviations from these stimuli.

1.0.1 Circuits for song production

In adult songbirds, singing is controlled by a set of brain regions spanning forebrain, midbrain, and thalamic areas specialized for vocal motor control (figure 1.3). Two of these brain areas, HVC (used as a proper noun) and RA (robust nucleus of the arcopallium), are in a region of the avian brain thought to be analogous to mammalian cortex, and are part of a descending song motor pathway (SMP) [34, 35]. HVC contains approximately 60,000 neurons per hemisphere that project to RA [36], which in turn contains roughly 8000 neurons [37] that innervate an area of the brainstem controlling the 8 muscles on each side of the vocal

organ (the syrinx) [38, 39]. Several studies have shown that individual RA-projecting HVC neurons (HVC_{RA}) generate a single brief burst of spikes (5-10ms in duration) at exactly one moment in the song motif. As a population, these bursts are distributed roughly uniformly throughout the motif [40, 41, 42, 43, 44, 45]. In contrast, neurons in the downstream nucleus RA individually generate 8-10 brief bursts of spikes (also of 10ms in duration) during the motif, in a pattern that may be driven directly, at each moment in time, by synaptic input from HVC [46, 47, 48]. Since RA neurons converge onto tracheosyringeal motor neurons that drive the vocal musculature, it has been proposed that the 'musical score' (or vocal content) of the birds song is established by the synapses from HVC to RA and/or RA to the brainstem, while the temporal order and timing of the song is controlled by the sequence of bursts in HVC [13, 49].

What generates the sequence of neural activity in HVC? Is the speed at which these sequences propagate controlled by circuit dynamics within HVC, or is the timing of these sequences controlled by circuit dynamics elsewhere, for example, driven by inputs from the thalamic nucleus Uvaeformis (Uva), or the Nucleus Interface (NIf), that project to HVC [50]? Previous work demonstrated that mild cooling of HVC during singing with a small thermoelectric device slows song production by an average of 2.8% per degree [51]. Notably, all timescales of song are slowed uniformly during cooling: subsyllabic structure, syllable durations, silent gaps, and overall motif duration. One model consistent with this observation, is that HVC sequences are generated by synaptically-connected chains of neurons—i.e. HVC neurons active at time 1 synaptically activate the neurons at time 2, which activate the neurons at time 3, and so on. Indeed it has often been hypothesized that these sequences arise from such direct synaptic connections within HVC [52, 53, 54] and computational models have shown that chains like this could emerge during development via heterosynaptic plasticity rules, such as spike timing dependent plasticity (STDP) [55, 56, 57]. Another possibility is that dynamics arise from activity that propagates around a cortical-midbrain-thalamic loop (HVC-RA-midbrain-Uva-NIf-HVC). However, cooling of RA fails to slow the song [51].

Importantly, recent evidence also strongly argues against this case, as HVC and RA exhibit sleep replay of sparse sequences that persist even after lesions of Uva and NIf, and are slowed by HVC cooling [58]. In this proposal we take inspiration from the idea that HVC serves as a motor clock to test the hypothesis that it additionally serves as a clock to store the timing of the tutor memory.

1.0.2 Circuits for motor learning and song evaluation

Song motor acquisition is thought to proceed by basal-ganglia-mediated reinforcement learning. In the songbird, the part of the basal ganglia dedicated to song learning, Area X, receives synaptic input from HVC, as well as another cortical area LMAN (the lateral magnocellular nucleus of the anterior neostriatum). LMAN generates highly variable patterns of activity during singing, and is thought to inject vocal variability into the motor pathway via its projection to RA [59, 48, 60, 61, 62]. Alternatively, variability may arise in the inputs to LMAN or as the product of recurrent interactions between LMAN and its inputs [63, 64]. Additionally, Area X receives information about song performance via dopaminergic (DA) axons originating in the ventral tegmental area (VTA) [15, 19, 65, 66, 67, 68]. Some have proposed [12] that medium spiny neurons (MSNs) in Area X integrate signals from HVC, LMAN, and VTA to detect the times at which (conveyed by HVC), variations in the song (conveyed by LMAN) lead to improved song performance (conveyed by VTA) [12]. This model posits that convergence of these factors drives synaptic potentiation in HVC synapses onto MSNs, such that HVC reactivates LMAN neurons that improve song performance, via the Area X, to thalamus, to LMAN loop.

In order for song performance to improve in such a model, dopaminergic neurons in VTA that project to Area X must transmit performance evaluation signals similar to reward prediction errors observed in mammalian basal ganglia. Recent evidence suggests this is indeed the case [15, 17]. Specifically, if artificial acoustic noise bursts are targeted to a particular syllable on a random 50% of song renditions, then VTA neurons generate a brief pause on

trials with a noise burst (as if indicating a vocal error), and exhibit a burst of spikes on trials on which the noise burst is absent (as if indicating a better than expected performance). Performance evaluation signals have also been observed during natural singing: VTA spiking increases on trials in which a syllable is very similar to its average rendition, but decreases when syllables most differ acoustically from this average [16]. Furthermore, optogenetic activation of VTA terminals in Area X is sufficient to drive learning when stimulation is made contingent on the pitch of a syllable, for example causing the pitch to increase when VTA is activated on syllable renditions with higher-than-average pitch [68]. Based on work contained in this thesis, we hypothesize that these error signals first arise in the brain due to predictive signaling from HVC to auditory cortical areas.

1.0.3 Vocal error signals in a descending pathway from auditory cortex to the dopaminergic system

Where do song performance evaluation signals in VTA come from? Retrograde tracing from VTA has revealed an auditory cortical region (ventral intermediate arcopallium, AIV), that provides inhibitory drive to X-projecting dopaminergic neurons in VTA [69]. Furthermore, AIV neurons exhibit robust spiking responses to artificial song errors, i.e. when songs are perturbed with artificial acoustic noise bursts, but not when those same noise bursts are played outside of singing [19, 69]. Note that, as expected from this arrangement, AIV and VTA have opposite effects on vocal learning; a recent study found that optogenetic activation of AIV terminals in VTA is sufficient to drive *aversive* song learning in a manner comparable to pitch learning induced by noise bursts (i.e. birds learn to avoid song-contingent stimulation of AIV terminals) [70]. Are the error signals observed in AIV computed locally or trasmitted by upstream regions? Further retrograde tracing from AIV has revealed projections from several cortical auditory areas, including, the L1 portion of field L, HVC shelf, the caudal lateral mesopallium (CLM), and the caudal medial nidopallium (NCM) [19]. Error-related signals similar to those in AIV have been observed during singing in both HVC shelf, CLM

and L1, two of which are reported to receive direct input from HVC [71, 72, 73]. Notably, NCM and CLM also exhibit adaptation to familiar songs and high responses to novel songs; repeated playback of a song causes highly specific and rapid reduction of neuronal responses (called song specific adaptation, SSA) that can last for hours to days [74, 75, 76, 77]. Some have speculated that this adaptation and novelty response may be the same mechanism by which the error responses described above develop [78], but this idea remains untested.

1.0.4 Where and how is the tutor memory stored

A large body of prior literature pursued the hypothesis that higher-order auditory cortical area, NCM, was the primary site of tutor memory storage (for a review see [79]). However, lesion studies suggest NCM is required to express a behavioral preference for the tutor song but is not required for song imitation [79, 80, 81, 82, 83]. This work is consistent with an arrangement in which NCM is important for memory based recognition of conspecific songs (including the tutor song), while other regions are involved in the tutor memory used for motor imitation. If NCM is indeed not required for imitation then CLM is an especially promising candidate for the storage of the tutor memory and computation of song error signals. CLM exhibits both SSA and song error responses, and contains a subregion that receives inputs from multiple song-motor areas [84, 21, 71, 74]. We hypothesize that the tutor memory is stored in the effective synaptic input from HVC to CLM (and, perhaps secondarily HVC shelf, another auditory area that receives HVC input).

HVC is well positioned to bridge auditory and motor domains as it receives input from diverse auditory areas including NIf, CLM, HVC_{shelf} and the primary auditory cortical regions in songbirds, collectively known as field L. In turn, HVC projects to a subregion of CLM with the evocative name avalanche (Av) [26, 29, 34, 84, 85]. Direct disruption of HVC activity during tutoring using optogenetic stimulation or electrical micro-stimulation impairs song imitation as does blockade of NMDA mediated plasticity in HVC [22]. NIf lesions as well as genetic lesions of only HVC projecting NIf neurons prevent imitation and rhythmic

stimulation of NIf terminals in HVC can act as an artificial "tutor song"; birds grow up to produce a rhythmic adult song biased by the period of this stimulation [20]. Finally, genetic ablations of Avalanche-projecting HVC neurons severely impair song imitation and prevent song degradation associated with deafening in tutored birds [21]. These avalanche projecting neurons generate sparse sequential activity during singing similar to the sequences produced by other HVC projection neurons sub-types [21]. Based on these observations, some have proposed that HVC may transmit a motor-based representation of the tutor memory to Avalanche, where it is compared with singing-related auditory feedback to generate error signals important for song learning [22, 28, 86, 87, 88, 89, 90]. However, whether such tutor memory information is in fact conveyed by these neurons during singing or tutoring remains unknown.

1.0.5 Auditory evoked activity in HVC

Given HVC's apparent functional role during tutoring and subsequent vocal practice, a natural question is whether HVC displays auditory activity which may underlie these functionalities. Indeed, in addition to HVC's pre-motor role, a large literature has documented robust auditory responsiveness in HVC in adult and juvenile birds [23, 25, 30, 31, 91, 92, 93, 94, 95]. Several observations stand out as being of particular relevance to the work contained in this thesis. In sleeping or urethane anesthetized adult zebra finches, auditory responses in HVC show strong selectivity for the bird's own song (BOS) over other songs [96, 97, 98, 25]. In other species of songbirds, basal ganglia projecting HVC neurons generate sparse bursts of activity at specific moments of the playback of BOS during wakefulness, and most importantly, the timing of HVC bursts to BOS playback precisely mirrors the timing of those same neurons' bursts during singing [27, 31, 99]. Such mirroring provides strong evidence for precise coordination of auditory and motor signals and perhaps reflects retrieval of song memories that guide song learning and maintenance. Auditory-vocal mirroring is not observed in awake adult zebra finches, but there is at least one report of similar mir-

roring between singing and sleep playback of the birds own song in a small number of basal ganglia projecting HVC neurons [72]. Almost all HVC neurons are silent during auditory playback in awake adult zebra finches, but fire robustly to auditory stimuli during sleep or when inhibition is suppressed pharmacologically [100, 89]. This somewhat puzzling observation may be related to a few things: the simplicity of the song they learn, the fact that they are not open-ended or seasonal learners, or alternatively some yet unknown reason [79]. In awake juvenile zebra finches however, HVC does exhibit auditory responses. During this early stage multi-unit responses are selective to the bird's tutor song [25, 91] and inhibitory neurons are selectively active during learned portions of the tutor song [89]. Tutor selectivity peaks during active song learning and declines in the later stages of development, perhaps due to the maturation of inhibitory circuitry during the process of motor learning.

The experiments described above show that HVC is necessary for tutor memory formation, and exhibits auditory responses during tutoring, but how do these auditory responses support tutor memory formation? Several lines of evidence show that plasticity in HVC occurs during tutoring and is important for song imitation. In particular, song imitation is disrupted by blockade of NMDA-type glutamate receptors [22] as well as pharmacological lesions of dopaminergic terminals in HVC [101]. In addition, two-photon imaging in HVC reveals that synaptic spines undergo a rapid stabilization following a single day of tutoring [88], a property which is absent if NMDA receptors are blocked during tutoring [22]. These changes can also be seen in electron microscopy data comparing HVC ultrastructure in tutored and untutored birds. Indeed, a single day of tutoring leads to a decrease in the number of excitatory synapses, but enlargement of those synapses that remain, consistent with synaptic strengthening and pruning [102]. Notably, in young birds, tutoring leads to a dramatic increase of spontaneous activity during sleep in downstream RA [103] which may reflect the replay of tutor-evoked activity, in a manner similar to the sleep-replay of song motor activity observed in adults [103, 104].

1.0.6 Summary and outlook

As evident from my brief overview of the songbird literature, an immense and detailed experimental literature has guided my thesis. Although the songbird does not offer the same genetic toolkit as rodents and/or some invertebrate species, it does provide the distinct advantage of spatially distinct and sparsely connected, specialized brain structures for song imitation. I cannot overstate how long neuroscientists have been thinking about tutor template formation in the songbird. The specific mechanistic hypothesis we propose and test in this thesis are deeply informed by lines of thinking that date back to the 1960's if not earlier [105, 2, 106]. In the excitement of making progress in scientific understanding it can be easy to forget to acknowledge work from so long ago, often speculative and lacking the tools experimentally or theoretically to precisely pose or test their lines of thinking in our contemporary manner. However these speculations, theoretical musings and premature declarations are an important part of science and support all aspects of the work contained in this thesis. I too hope to have the occasion in the future to speculate (perhaps wildly) in the right direction.

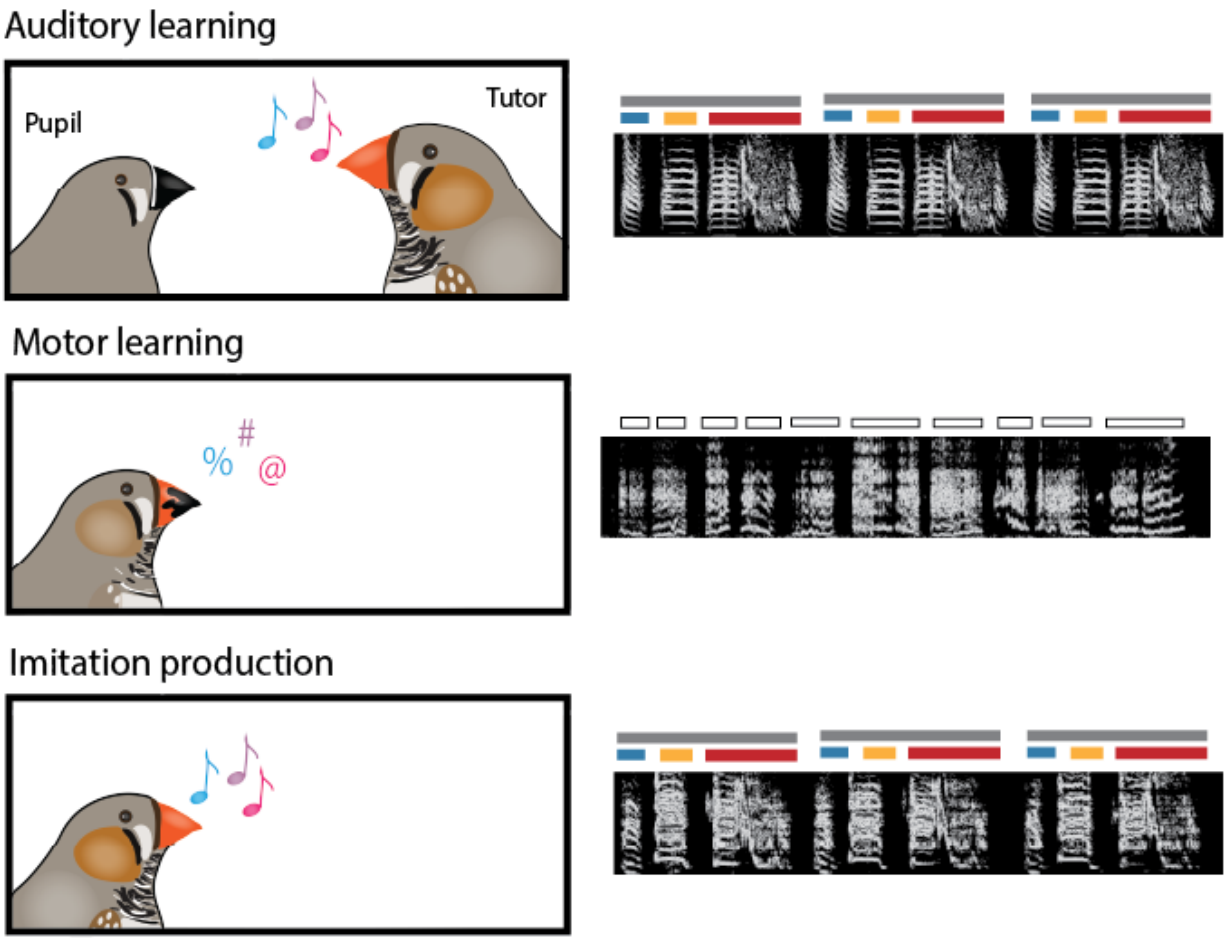


Figure 1.1: Overview of song learning

(a) Overview of the song learning process. Top: during the auditory learning stage juvenile male pupils form a memory of the tutor song of an adult male tutor. A spectrogram of an an adult tutor song is shown on the right, with syllables segments shown with colored boxes and the motif shown in grey boxes. Middle: juvenile male birds undergo motor learning in which they produce immature songs which become more structured throughout the learning process. A spectrogram of a juvenile song in an intermediate stage of song learning is shown on the right. Bottom: Juvenile males reach maturity and produce a "crystallized" song which changes very little throughout adulthood. This song is often quite a good match as can be seen by comparing the spectrogram of the imitation shown on the bottom right, to the tutor song at the top right

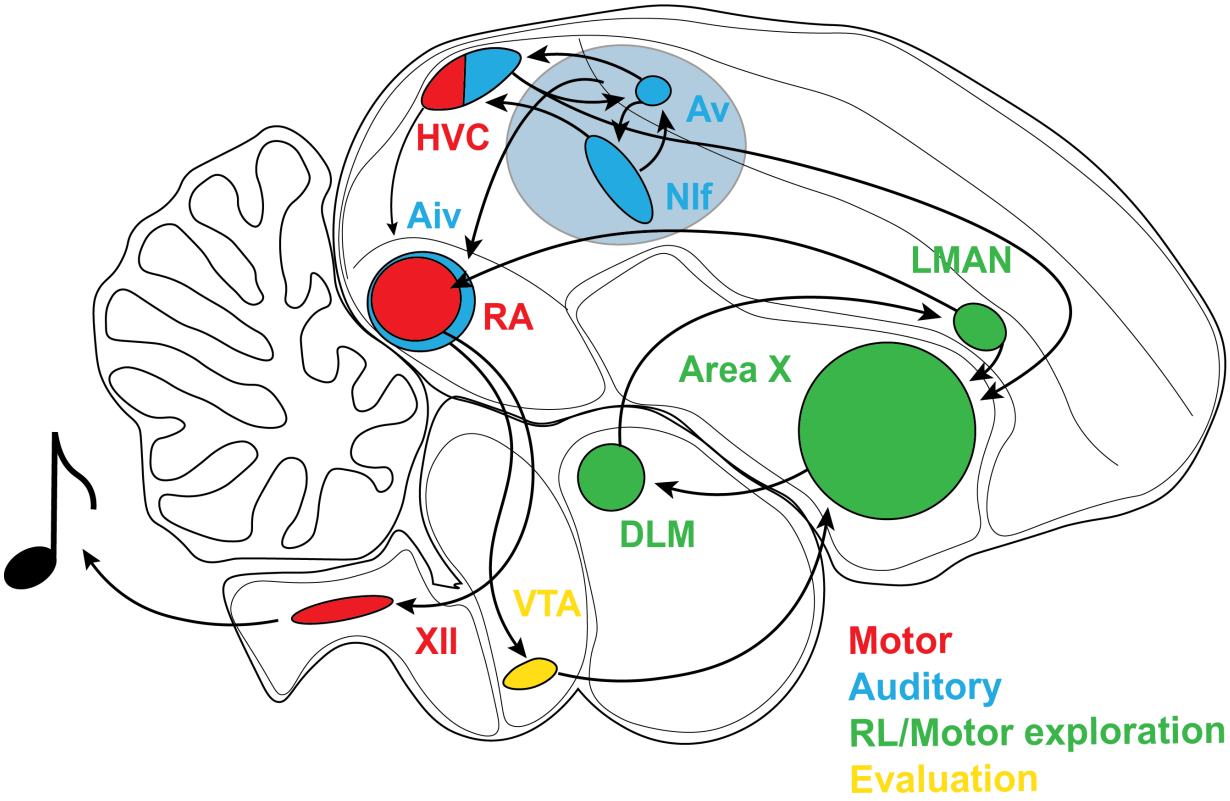


Figure 1.3: Schematic of nuclei and brain regions making up the song system

(a) Overview of brain regions most important for chapters contained in this thesis. Red regions denote the song motor pathway, important for the production of adult song. Blue shows the auditory regions which are interconnected with song nuclei, broad blue transparent region indicate that NIf and Av are embedded in a larger auditory cortical network which and indicates the rough location of thalamic auditory inputs to the cortex. Green regions indicate brain regions implicated in motor exploration and learning also known as the anterior-forebrain-pathway (AFP). This region receives in put from dopaminergic neurons in a region thought to correspond to VTA in mammals, shown in yellow. This VTA-like regions integrates auditory error information from Aiv and other auditory cortical areas including those in close vicinity of Av

Chapter 2

Perturbing HVC dynamics during auditory memory formation using cooling

2.1 Abstract

Songbirds acquire their songs through an imitation process reminiscent of speech learning in humans; juvenile songbirds form a memory an adult male tutor's song and gradually refine their own vocalizations to produce a mature song closely matched to the tutor song. During production of the adult imitation, neurons in the pre-motor nucleus HVC produce a sequence of ultra-sparse activity spanning adult song [42]. Subsequent research demonstrated that dynamics within HVC control the timing of this song; focal cooling of HVC, but not downstream motor regions, slows the production of song uniformly [51]. While this research significantly advanced our understanding of song motor production, the key processes of tutor auditory memory formation and recall that support imitation, are poorly understood. Recent work however, suggests that these processes may be controlled in part by HVC. Disruption during tutoring of HVC itself, its auditory inputs or its auditory outputs impairs birds' ability to imitate [22, 21, 20]. Inspired by these observations, we hypothesized that HVC acts as a clock that stores the timing of the tutor memory. Here we test a critical prediction of this theory: that cooling of HVC during tutoring leads to a sped-up tutor song memory, much like slowing the motor of a tape recorder during recording, and then playing it back at normal speed. To test this, we designed a modular thermoelectric cooling device, and found that transient cooling during tutoring caused birds to produce faster imitations consistent with this model. This work gives insight into the mechanism by which tutor song

memories are formed and recalled, and shows how memory content can be systematically manipulated during memory formation.

2.2 Introduction

In order for animals to make use of their experiences, it is crucial that they not only form memories of individual perceptual events, but also of the structured relations/relationships between events. At one extreme, children in English speaking countries learn to use and produce 60-80,000 words and the complex syntactic rules governing their use. Additionally, most humans spontaneously form memories of long stretches of musical material, easily remembering tempo, ordering and fine acoustic detail [107, 108]. Even simple associative learning requires animals to remember that some stimuli belong together, while others do not. In practice though, the relations even non-human animals must learn are quite complex: bats link memories of food experiences to spatial maps to enable navigation back to distant food sources [109], primates remember dominance hierarchies to ensure safe and appropriate social conduct in groups [110], and, songbirds recall not only the precise content of the song to be imitated, but also the speed and ordering of those sounds [2].

In genetically tractable organisms like mice, some progress has been made in understanding where "engrams" for individual memories may reside. For example, the tagging and reactivation of cells active in the hippocampus during the formation of a fear memory causes animals to freeze, as if remembering this experience [111, 112]. Likewise, artificial memories have been produced by pairing neural reactivations of cells coding for particular locations during sleep, with stimulation of the medial forebrain bundle (MFB), inducing a synthetic preference for that location upon waking [113]. Despite this success, the structured relationships between memory contents (among or within distinct experiences) has proven far more difficult to manipulate. However, vocal imitation in songbirds provides a promising model system for studying this process. Juvenile songbirds form a memory of a tutor song including the ordering of song syllables and the often complex, time-varying pattern

of spectral content which make up each individual vocal element. These memories can be formed rapidly – birds can produce accurate imitations after a mere 75 seconds of cumulative song exposure [114]. Critically, imitating these songs requires that birds remember not just the spectral content of the song, but also the highly structured relationships between the contents, such as the ordering of sounds, relative durations, and the temporal relationships between sounds (a.k.a the song speed or tempo). The memory containing this collection of attributes is often referred to as "the tutor song template".

In the search for the neural locus of template, a pre-motor cortical area known as HVC, has emerged as a promising candidate. Tutoring induces rapid reorganization of HVC connectivity. This includes excitatory synaptic pruning, stabilization and strengthening, and an increase in the density of inhibitory synaptic contacts [88, 102]. Accurate imitation requires NMDA and dopaminergic signaling in HVC during tutoring [22, 101], and neuronal responses exhibit selectivity for the tutor song following tutoring [25, 115, 91]. Strikingly, rhythmic optogenetic stimulation of auditory inputs to HVC can act as a synthetic "tutor"; juvenile birds grow up to produce rhythmic songs, biased by the period of this stimulation [116]. Despite this large amount of evidence for HVC's involvement in the formation of a tutor memory, we know little about how this memory is formed and used to guide learning.

There is strong evidence however, favoring a specific mechanistic account of HVC's function during song motor production. During singing, HVC projection neurons fire in sparse sequences reminiscent of synfire chains, leading some to hypothesize that these sequences act as a motor clock [42]. Consistent with this, focal cooling of HVC during song production uniformly slows all aspects of the song, suggesting local neural dynamics underlie this function [51]. Inspired by this framework, we hypothesized that HVC may play a dual role in auditory and motor processes, acting as a clock for tutor memory formation and recall analogous to its role in motor production. We reasoned that, if HVC dynamics also govern the formation and recall of the tutor memory, cooling HVC during tutoring should lead to a sped up memory, and thus a sped-up imitation of the tutor song. This conclusion is eas-

ily seen by analogy to a tape recorder: while slowing the motor of a tape recorder during playback slows the produced sound, slowing the motor of the tape recorder only during the *recording* phase will lead to a recording that sounds sped up. In this chapter we test this hypothesis using a novel thermoelectric cooling device.

2.3 Methods

2.3.1 Animal care and use

We used male zebra finches (Taeniopygia guttata) from the MIT zebra finch breeding facility (Cambridge, MA). Animal care and experiments were carried out in accordance with the NIH guidelines and reviewed and approved by the Massachusetts Institute of Technology Committee on Animal Care. For all tutoring experiments we prevented exposure to a tutor song prior to experiments; birds were separated from their father on or before 15 days post-hatch and raised by their mother and a foster female. For experiments, birds were singly housed in custom made sound isolation chambers.

2.3.2 Surgical procedures

Juvenile male zebra finches (35-65 dph) were anesthetized with 2% isoflurane gas before being placed in a custom stereotaxic apparatus. After applying a topical anesthetic (0.1% bupivacaine) and performing a midline incision in the skin over the skull, we made craniotomies in the skull at a predetermined distance from the bifurcation of a major blood vessel (the 'lambda sinus'). Bilateral craniotomies were made over HVC and a thermoelectric device was lowered until the two silver pads made contact with the dura above HVC. Craniotomies were sealed with silicone elsatomere (kwik-kast) and the device was secured using dental acrylic. The incision site in the skin was closed with a tissue adhesive (VetBond). Analgesic (Buprenorphine or Meloxicam) was administered 30 minutes prior to surgery and for three

additional days after the end of surgery.

2.3.3 Cooling device

To focally cool HVC during tutoring we designed a thermoelectric device based on prior designs [51, 117, 62]. Our new device featured a modular design which allowed its removal after the end of the tutoring process. A permanent 3-D printed baseplate was cemented to the skull (figure 2.5.4) and connected via two small set screws to a removable 3-D printed component housing the cooling element (figure 2.5.4). Song learning can take several weeks after the end of tutoring, so a removable cooling element allows higher throughput of experiments and, critically, allows birds to spend most of the learning process untethered. Two 2x1mm silver pads (0.005in in thickness) were soldered to a miniature Peltier device (Custom Thermoelectric, 00701-9A30-12RU4) and plated with gold to prevent silver toxicity. A K-type thermocouple (Omega, 5SRTC-TT-K-40-36), used for PID feedback, was affixed to the device with the tip contacting the surface of one of the two silver pads and the tip and pad were covered with thermally conductive paste. For effective and stable cooling, heat-waste generated by the cooling device was dissipated using a custom, water-cooled, copper heat-sink previously described [117]. Current was delivered to the device using a series of thin, braided copper wires (Cooner Wire, Inc. CZ-1187), a programmable current source (TDK-Lambda ZUP36-6), and custom software. Following tutoring, the thermo-electric module was removed and the area was covered with an insulating layer of silicone elastomer and dental acrylic. We calibrated our device after surgical implantation by simultaneously measuring the temperature at the surface of the silver pad and the temperature 0.5mm below the pad surface with two thermocouples. Using custom software, we performed PID feedback to maintain the pad surface temperature at a specified command set-point. We then characterized the steady-state temperature 0.5mm below the pad surface relative to the pad surface for a variety of temperatures (figure 2.9). We also verified that cooling with our device leads to changes in song speed during adult song previously reported [51], and that removal of the

cooling element leads to only very modest changes in temperature at the dura above HVC of approximately 1 degree Celsius (figure 2.11).

2.3.4 Cooling during tutoring

During tutoring sessions, HVC was cooled bilaterally using two pads as described above. A previous study found that the temperature of HVC is typically around 41°C during the day, and 37°C at night [118]. Therefore, HVC temperatures were maintained by PID feedback at 37°C during lights off and 41°C throughout the day. During tutoring, the temperature of HVC was controlled by changing the set point of the PID feedback on the pad surface. To allow the brain to reach steady state temperature, 5 minutes were allowed after changing the temperature set-point prior to introduction of the tutor. HVC was then held at the experimentally imposed temperature until the end of the tutoring period (1hr total). For each individual bird, HVC was held at the same temperature each day throughout tutoring; however, individual birds were arbitrarily assigned to a specific temperature change ranging between -7 and +2.5 degrees relative to baseline. Birds were individually housed after the end of the tutoring period and their vocalizations were monitored until they reached 90-120 days-post-hatch (dph), at which point song imitations were evaluated.

2.3.5 Cooling analysis

The fractional change in motifs, syllables and gaps were calculated as has been previously described [51]. Syllable and motif onsets and offset times were determined using a threshold crossing of sound amplitude. Gaps were then measured as the offset to onset time between pairs of syllables using the same method. We refer to the fractional change of a syllable, motif or gap in the imitation, relative to its respective duration in the tutor song as the pupil dilation. This quantity was calculated as:

$$Dilation = \frac{D_{pupil} - D_{tutor}}{D_{tutor}}$$

where D_{pupil} is the median duration of a syllable, gap or motif, and D_{tutor} is the median duration of the corresponding syllable, gap or motif in the tutor song. The linear effect of temperature on imitation speed at the syllable, gap, and motif level was assessed by pooling all imitations across birds and performing a linear regression between HVC temperature change during tutoring, and song element dilation.

2.3.6 Tutoring protocol

We isolated birds from their father at or before 15 days post hatch (dph) based on prior studies that achieved effective imitation in experimental animals isolated between 10 and 30 dph [1, 88, 119]. These isolated birds remained co-housed with their siblings and mother until 35 dph, at which point they were singly housed for the duration of the experiment [102]. Zebra finches can produce accurate imitations after as little as 75s of tutor exposure [114], but protracted daily tutor exposure impairs imitations [120]. We therefore followed an intermediate tutoring protocol in which pupils received 60 minutes of tutor exposure each day [102]. Birds were housed in a sound attenuating chamber and their songs were recorded daily using either Sound Analysis Pro software (SAP)[121], or custom recording software. Every day for 2-14 days, the tutor was placed in the sound recording chamber with the pupil for 1 hour. All tutor song was recorded during the tutoring sessions to allow for possible variations in the speed of the tutor song for different pupils.

2.4 Results

2.4.1 Using temperature to manipulate neural dynamics

Here we set out to use temperature to manipulate the timescale of neural dynamics during memory formation, an approach that has now been widely adopted during active behaviors, to modulate neural dynamics in a number of different organisms and brain areas [51, 62,

118, 117, 126, 50, 127, 58, 128, 122, 123, 124, 125]. The effect of temperature on neural activity follows from the fact that most biological reactions have a Q_{10} between 2 and 3, meaning a temperature decrease of 10°C tends to slow the reaction rates by two to three fold[1] [130]. As expected from this line of reasoning, the speed of many brain processes are highly temperature dependent [131, 135, 136, 137, 138, 139, 140, 141, 142, 132, 133, 134, 126, 127], and cooling of brain circuits involved in the timing of behavior can have the effect of slowing behavior itself [143, 144, 145, 51, 62, 117, 50, 126, 128, 122, 123, 125][2]. Here we used cooling to ask another important question: how do neural circuits form memories of the timing between events? We reasoned that cooling should affect the timing of neural processes, not only during the production of learned behaviors, but also as memories of temporally structured experiences are formed, allowing us to directly interrogate the learning process with temperature[3].

2.4.2 Perturbing HVC dynamics during auditory memory formation

To test the effect of HVC cooling on song imitation we bilaterally manipulated HVC temperature in isolate juvenile birds as they were naturally tutored by an adult male. Birds were implanted with a thermo-electric cooling device and HVC temperature was controlled using real-time PID feedback (figure 2.1). After birds were implanted with the cooling device, HVC was held at its typical temperature of 41°C during the day, and 37°C at night, as previously reported [118]. Birds were tutored for 1 hour each day for a total of between 2 and 14 days. During this time, the temperature of HVC was held at an experimenter defined temperature which varied across subjects but was constant within each subject. In

[1] There are notable exceptions in which a biological process is actively temperature compensated, for example the biochemical oscillators underlying circadian timekeeping [129].

[2] For a great overview and discussion studies using focal cooling in the brain see [146].

[3] Note that cooling could in principal universally impair learning. At least in hippocampus dependent tasks this is not true: learning is unimpaired by mild cooling of the whole brain [147, 148], but is impaired by selective cooling of the medial septum [127]

experimental birds, the temperature of HVC was changed 5 minutes before the onset of tutoring to allow the temperature of the brain to reach a steady state. A group of control birds were implanted with the same device but were tutored as HVC was maintained at 41°C. At the end of tutoring period, the module housing the cooling element was removed, and the area was covered with thermally insulating material. Birds were then placed in acoustically isolated chambers, and their songs were recorded until they were at least 90dph.

2.4.3 Cooling HVC during tutoring compresses imitation of song motifs

First we wondered whether cooling produced any systematic effects on the speed of the imitated song motif. While only a few birds imitated the entire tutor song motif (likely due to a combination of their abnormally late age at tutoring onset, combined with the presence of the implanted device), most produced imitations of motif-fragments, which we define as unique groups of two or more imitated vocal elements produced in the correct order. For each bird with a motif-fragment (n=23 motifs fragments in 23 birds), the fractional change in the speed of each distinct motif-fragment was quantified and analyzed as a function of the corresponding temperature of HVC during tutoring (Figure 2.5.4 a). The slope of the relationship was measured by performing a linear regression between temperature during tutoring, and the fractional change in duration between each imitated motif-fragment and corresponding fragment in the tutor song (here referred to as dilation of the imitation). Overall, we observed a relationship of 1.9%°C^{-1} (α=1.90 ± 0.41, p<10^{-5}), with cooler temperatures yielding faster imitations and warmer temperatures yielding slower imitations on average. The sign of this effect was opposite to that observed when cooling HVC during vocalizations. Overall, the magnitude was slightly less than the effect of cooling HVC during song production in adults (-2.83± 0.22%°C^{-1} on average) [51], but greater than the effect of cooling HVC during song production in juveniles (-1.36± 0.16%°C^{-1} on average)[62]. There were no obvious effects of temperature on the acoustic structure of the imitated motif-fragments (figure 2.3 b), and

the fragments of the tutor motif that were imitated did not appear to be systematic across subjects.

2.4.4 Cooling HVC during tutoring compresses imitation of song syllables

Next, we asked if there was a systematic effect of temperature on the speed of imitation at the level of song syllables. For all imitated syllables (n=79 syllables in 35 birds), the fractional change between imitated syllable duration and tutor syllable duration was quantified and again analyzed as a function of the corresponding temperature of HVC during tutoring (Figure 2.5.4 b). The slope of the relationship was calculated by a linear regression of fractional change in duration for a given syllable and tutor song syllable. Overall we observed a relationship of 2.0%°C^{-1} (α=2.05 $\pm$ 0.32, p<10^{-9}), nearly identical to the previously described relationship between motif-fragment and HVC temperature during tutoring.

2.4.5 Effects of temperature on syllable-types

Although cooling affects the speed of both motif-fragments and syllables overall, it remains possible that these effects are produced by changes only in some syllable types, while others are either unaffected or even show an inverted relationship. To test this possibility, we analyzed the relationship between temperature and syllable dilation for each syllable type in the tutor song. Because all pupils in our study were tutored with the same tutor, we were able to compare different imitations of the same syllable across many different temperature conditions. For three of the syllables in the tutor song, we observed a significant relationship between temperature and imitation dilation (figure 2.7). In one syllable type we observed a similar trend, which was not significant (figure 2.7 c). The similarity in measured relationships between temperature and dilation across syllable types suggests a uniform mechanism for storing timing relationships, regardless of a syllable's spectral content. It is still possible

that other syllable types not present in our tutor song could exhibit more complex effects from cooling, as has been observed when cooling during song production in other species [123].

2.4.6 Cooling during tutoring has no effect on imitation of silent gaps in the tutor song

We next asked if imitation of the silent gaps between song syllables was also sped up by cooling of HVC during tutoring. Gap imitations were defined as the silent period between all pairs of correctly ordered, imitated syllables. We again calculated the fractional change for each imitated gap (n=33 silent gaps in 23 birds) relative to the corresponding gaps in the tutor song and analyzed as a function of HVC temperature during tutoring (Figure 2.5.4 c). In the case of imitated silent gaps, the slope of the regression was similar to the slope for motif-fragments and syllables but was not significant (α=2.00 $\pm$ 1.52, p=0.20). This may be due to the significantly higher variance of gap dilation for all conditions, or reflect a fundamental mechanistic difference in the way that silent gaps are represented and/or remembered.

2.5 Discussion

2.5.1 HVC's function in tutor memory formation

Our results in this chapter suggest that HVC functions as a clock for the formation and recall of the tutor memory analogous to its role during song production. We hypothesize that during tutoring, tutor evoked activity in HVC undergoes plasticity, producing neural dynamics that form the basis of adult motor sequences. Cooling may bias which synapses are strengthened during tutoring by slowing the rate of neural transmission specifically during this learning phase. The way in which cooling during tutoring compresses the imitation may

also suggest that learning in HVC underlies the formation of a tutor template, perhaps via a projection from HVC to the auditory area Avalanche. Alternatively, cooling could bias the timing of tutor memory readout in HVC without affecting some other more central tutor memory. Based on the results in this chapter, and previous studies implicating HVC in tutor memory storage [149, 22, 21, 103, 20], we predict that HVC activity during tutoring is organized into sparse sequences that are strengthened by Hebbian plasticity. We revisit and test these ideas in subsequent chapters.

2.5.2 Possible biophysical mechanisms explaining the effect of cooling on song imitation.

Here we found that cooling during tutor exposure leads to sped up imitations, with an overall relationship between temperature and song imitation of approximately 2% per degree. Notably, cooling exerts its effect despite birds producing no apparent behavior. This manipulation nonetheless has long lasting effects on the structure of vocal imitations weeks and months later. What physical mechanism might underlie the effect of cooling on song imitation? In principal, any biophysical process involved in neural transmission (from vesicle release at the synaptic cleft, to conductances underlying action potential initiation and propagation) may be slowed by cooling and could contribute to the observed effect. Notably, the effect of cooling on the rate of axonal propagation [138] is very close in magnitude to the behavioral effects observed for singing previously [51] and for imitation in this chapter. During adult vocal production, the effect of cooling was suggested to result from an accumulation of delays due to a hypothesized domino-like nature of network dynamics in HVC [51]. Are HVC dynamics during tutoring likewise organized into "domino-like chains of activity"? While there is at least one report of pre-existing "latent" HVC sequences in older juvenile birds [119], these latent sequences are not likely present in the majority of

juveniles at the time of natural tutoring[4]. Rather, we hypothesize that cooling changes which synapses are potentiated during tutor evoked plasticity. For example, consider two weakly connected neurons activated sequentially (at short latency) by auditory input during tutoring. The synapses connecting this pair could undergo spike-timing-dependent-plasticity (STDP). Here, the degree to which a synapse is strengthened depends strongly upon the precise timing of pre-synaptic release and post-synaptic activation, such that small changes in this timing can have strong effects on both the sign and magnitude of induced plasticity [150, 151]. We suggest that cooling systematically delays the arrival of post-synaptic potentials between pairs of HVC neurons, without affecting the timing of feed-forward input from the auditory system (which is determined solely by the temporal structure of the tutor stimulus). Because cooling slows neural transmission, pre-synaptic potentials that would have arrived too early, now arrive in a more favorable part of the STDP kernel and are strengthened. After cooling, when the brain is returned to its normal temperature, the speed of transmission returns to normal, resulting in faster transmission at these strengthened synapses than would otherwise be expected. If plasticity in HVC during tutoring [88, 102, 22] partially, or completely, underlies the structure of adult motor sequences, then we would expect to observe a systematic relationship between imitation speed and temperature as reported in this chapter (for a visual explanation of this hypothesis, see figure 2.14). Intuitively this can be understood as follows: when HVC is cooled, the synaptic chain formed runs faster than one formed at normal temperatures.

2.5.3 Differential effects on song elements

In this chapter we observed a significant effect of temperature on durations of syllables and motif-fragments, but not on gaps. One possibility is that cooling exerts a similar effect on song gaps, but this effect is masked due to poor imitation of these silent gaps. This is supported by the fact that the variance of the gap durations is much higher than for syllables

[4]We will show that neural sequence are present during tutoring in chapter 3 (figure 3.1 and 3.7), but are likely elicited input to HVC, not internal pre-configured dynamics.

or motifs, yet the slope of the (statistically insignificant) relationship between temperature and gap dilation is very similar to the slope for these other song elements. Recent experimental evidence supports the idea that imitation of silent gap durations is less accurate than the imitation of syllable durations; two recent studies found that the duration of gaps in a tutor song only weakly influences the duration of gaps in the final imitation [152, 153]. Other strong influences on the duration of gaps, for example heritable influences on respiratory physiology (no birds in the present study were tutored by their father), or random processes, could explain why imitated gap durations were far more variable in our data set than syllables and motifs, yet still tended to vary with temperature.

2.5.4 Outlook and future work

Localizing the tutor memory in the songbird has long been a "holy grail" for the field [32, 79, 33]. However, this has been a fraught and difficult process. Many studies implicate auditory and song motor areas in this process through the use of lesions, pharmacological inactivation, as well as optogenetic and electrical stimulation [82, 22, 21, 20, 19]. Although the thrust of this literature is that auditory areas and song motor regions are both likely involved in the storage of the tutor memory, these manipulations can have complex effects on auditory processing and may have non-specific or off-target effects [154]. Here we adopted the approach of mild brain cooling to systemically perturb the memory as it is formed, and provide evidence that HVC stores the timing of the tutor memory during listening that are used to guide learning. Not only does this make a strong contribution to localizing the songbird tutor memory but it also provides, to our knowledge, the only example of a manipulation that allows the structure of a memory to be altered without disrupting its contents. Outside of the songbird field, a fundamental goal of neuroscience is to understand how memories get to have their contents i.e. what about neural representations and storage allows the recall of the remembered objects, their distinctive properties, and relation between them. Together with our cooling results, a picture of how song templates might form in HVC

is beginning to emerge. Future work will give us insights into how this template is stored at cellular and network levels and provide insight into memory formation in other systems.

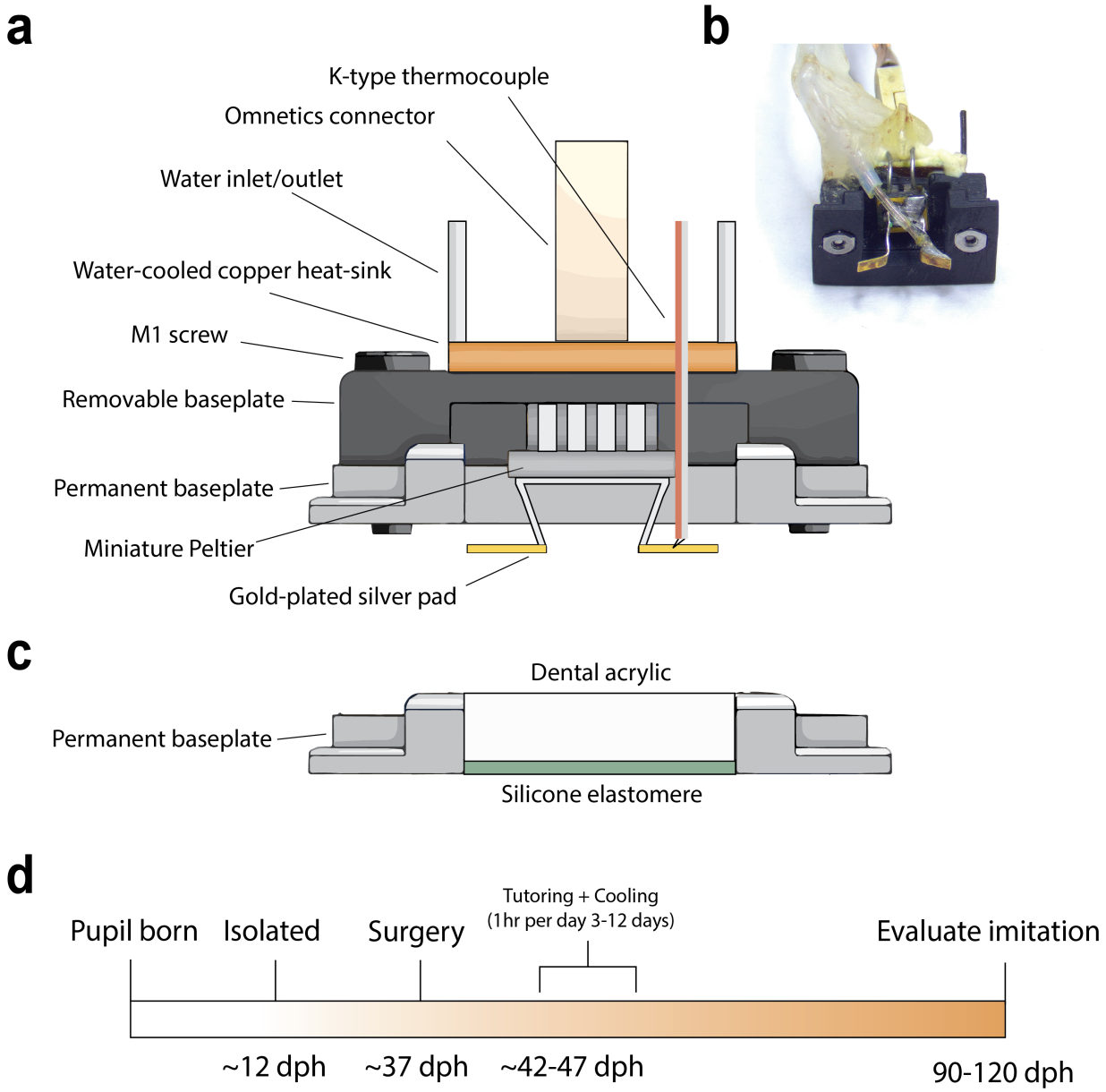


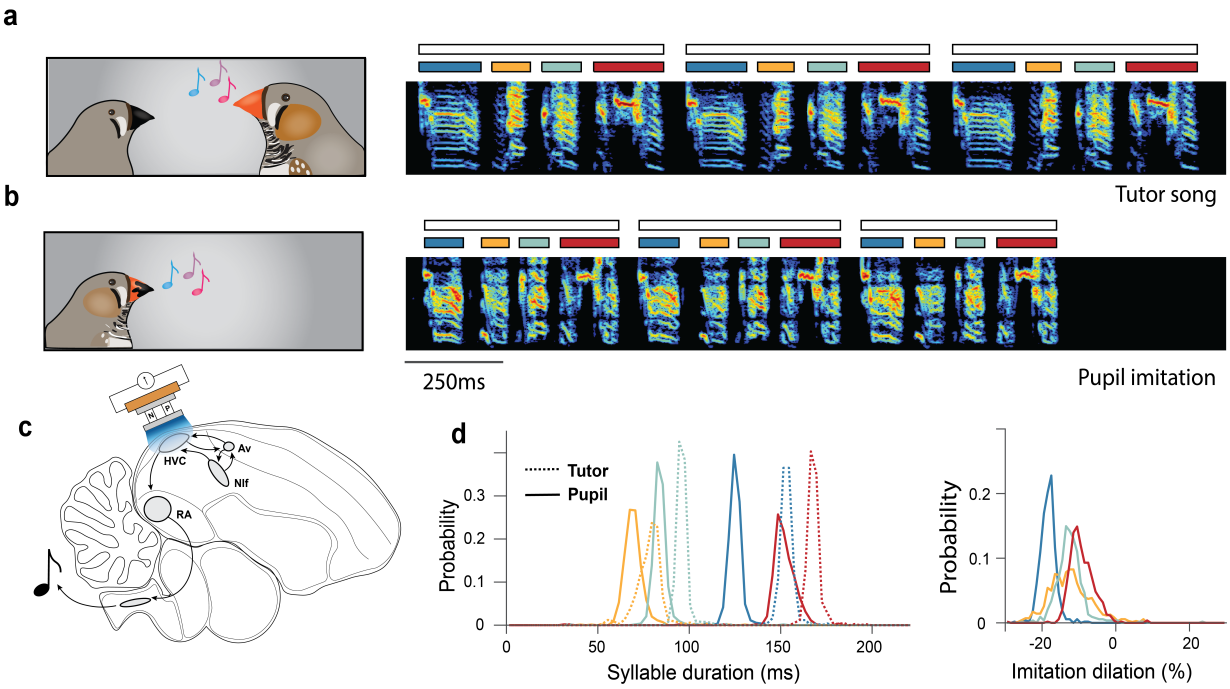
Figure 2.1: Schematic of the thermo-electric cooling device

(a) A schematic showing the thermo-electric cooling device used in cooling experiments. Note the modular permanent and removable baseplates held in place by small set screws, water cooled copper heat-sink (silicone tubing on inlet and outlet not shown) and K-type thermocouple used for feedback control of temperature. Design schematics shown at the end of the chapter

(b) Inset showing a picture of the fully assembled device from the bottom. Note press fit nuts in the permanent baseplate. This allows the removable portion to be freely attached or removed using the set screws.

(c) A schematic showing device after removable baseplate detachment. Area is covered with a thin layer of silicone elstomer (kwikcast) and a thick layer of dental acrylic to thermally insulate area

(d) Timeline of experiments for this chapter. Approximate times of events are shown at the bottom relative to the day of hatching (days-post-hatch, dph)



(a)

Figure 2.3: Example tutor song and song imitation of a bird cooled during tutoring

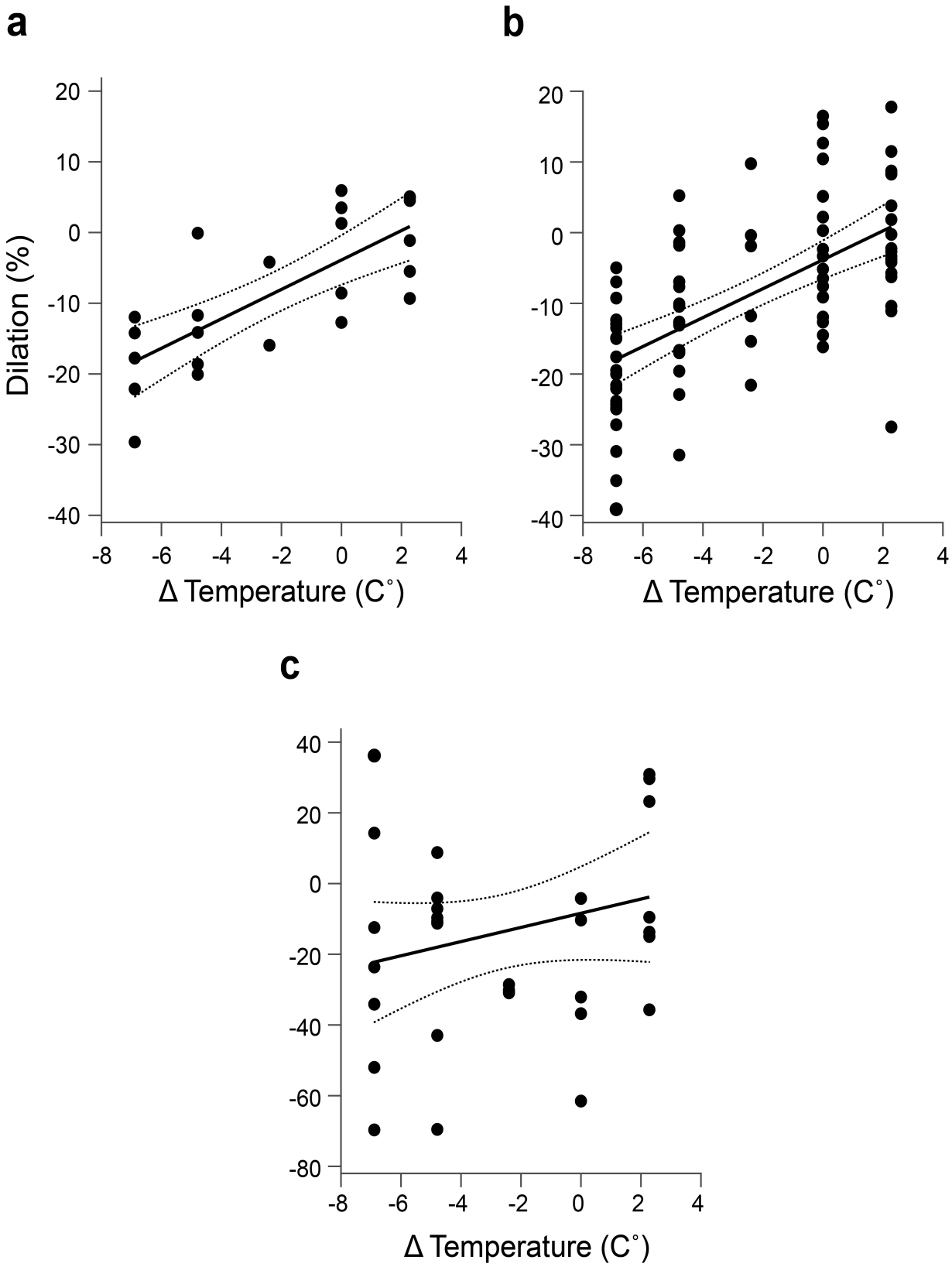
(a) Cartoon showing tutoring process (left) and spectrogram of three motifs of the tutor song (right). Song is composed of four different syllables shown in colored boxes above song spectrogram

(b) Cartoon of the juvenile male producing an imitation (left) and an example specrogram showing the imitation produced by a bird that had HVC bilaterally cooled by 5°C during tutoring. Note that the imitation contains copies of all 4 syllables ordered as in the tutor motif.

(c) A cartoon showing the cooling strategy used in this chapter. Cooling is achieved via the Peltier effect and is focally applied to HVC. Relative locations of important song motor and song auditory brain regions are also shown

(d) Syllable duration distributions for all tutor song syllables and for all imitated song syllables in one example bird. Solid line shows pupil duration distributions and dotted lines show tutor distributions. Colors indicated syllable types as in a and b. Note that pupil distributions for each syllable type are shifted to the left relative to the tutor distributions

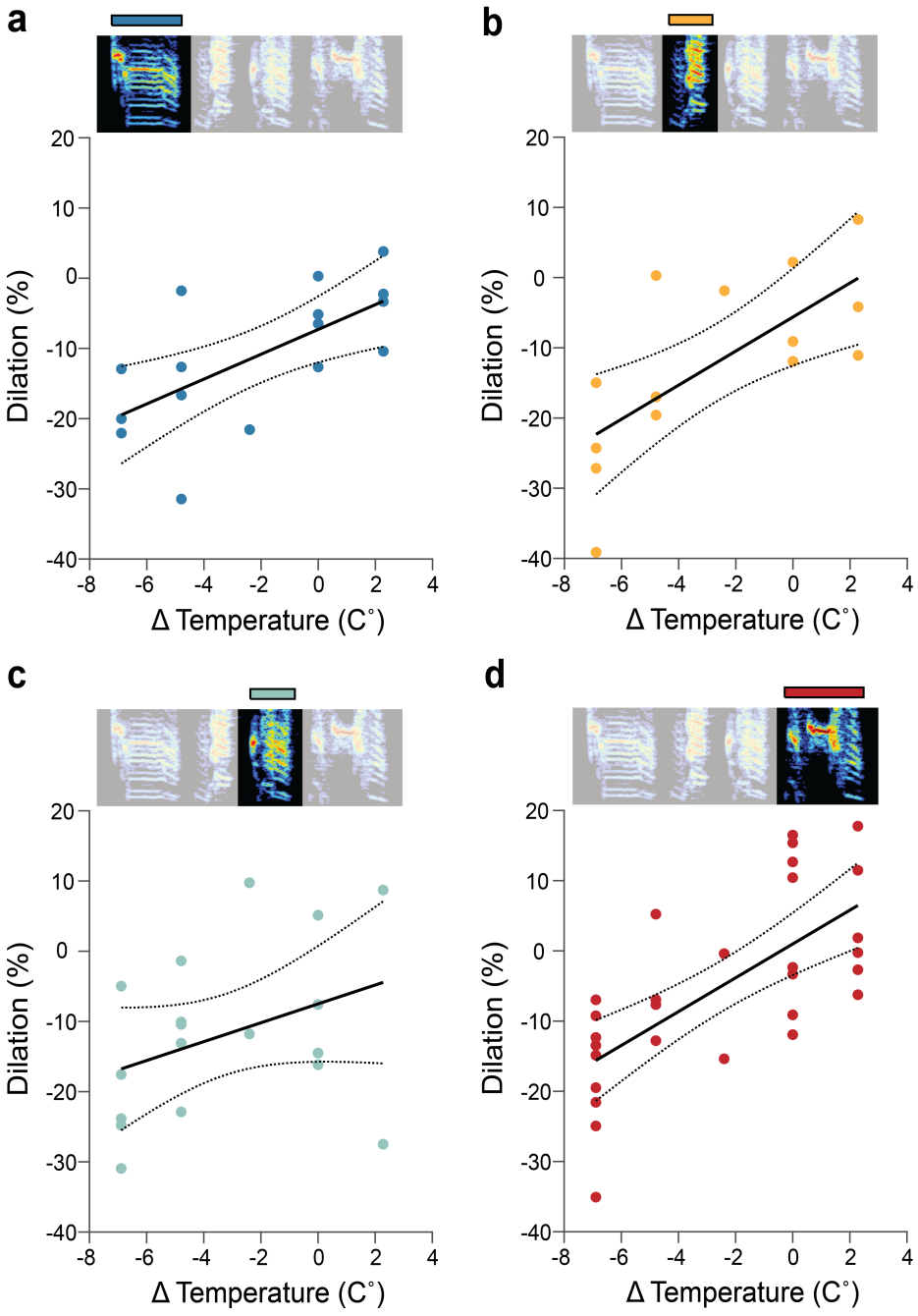
(e) Imitation dilation for all imitated syllables in this example. Imitation dilation indicates fractional change in the durations of imitated syllables relative to tutor syllable durations. Note that syllables were 10 to 20% faster, as indicated by negative dilation values.


a
b
c
Dilation (%)
Δ Temperature (C˚)
Δ Temperature (C˚)
Δ Temperature (C˚)

(a) Fractional dilation of imitated motif-fragments as a function of the deviation of HVC temperature during tutoring from typical HVC temperature ($P<10^{-5}$ (linear regression), $\alpha = 1.90$,$R^2=0.50$,$F=21.2$ and 21 df; n=23 motif fragments in 23 birds). The thick line and dotted lines represent the fit and 95% confidence interval respectively.

(b) Fractional dilation of imitated syllables as a function of the deviation of HVC temperature during tutoring from typical HVC temperature ($P<10^{-9}$ (linear regression), $\alpha = 2.05$,$R^2=0.36$, $F=42.4$ and 76 df; n=79 motif fragments 35 birds). The thick line and dotted lines represent the fit and 95% confidence interval respectively.

(c) Fractional dilation of imitated silent gaps between syllables as a function of the deviation of HVC temperature during tutoring from typical HVC temperature ($P=0.20$ (linear regression), $\alpha = 2.01$,$R^2=0.05$,$F=1.74$ and 21 df; n=31 motif fragments in 23 birds). The thick line and dotted lines represent the fit and 95% confidence interval respectively.



(a)

Figure 2.7: Syllable dilation by syllable type

(a) Fractional dilation of imitated syllables as a function of the deviation of HVC temperature during tutoring from typical HVC temperature for syllable a ($P<10^{-3}$ (linear regression), $\alpha = 1.78$,$R^2=0.43$,$F=10.6$ and 14 df; n=16 birds).The thick line and dotted lines represent the fit and 95% confidence interval respectively.

(b) Fractional dilation of imitated syllables as a function of the deviation of HVC temperature during tutoring from typical HVC temperature for syllable b ($P<10^{-3}$ (linear regression), $\alpha = 2.43$,$R^2=0.50$,$F=11.7$ and 12 df; n=14 birds).The thick line and dotted lines represent the fit and 95% confidence interval respectively.

(c) Fractional dilation of imitated syllables as a function of the deviation of HVC temperature during tutoring from typical HVC temperature for syllable c ($P=0.14$ (linear regression), $\alpha = 1.34$,$R^2=0.13$,$F=2.48$ and 16 df; n=18 birds).The thick line and dotted lines represent the fit and 95% confidence interval respectively.

(d) Fractional dilation of imitated syllables as a function of the deviation of HVC temperature during tutoring from typical HVC temperature for syllable d ($P<10^{-5}$ (linear regression), $\alpha = 2.41$,$R^2=0.47$,$F=24$ and 27 df; n=29 birds).The thick line and dotted lines represent the fit and 95% confidence interval respectively.

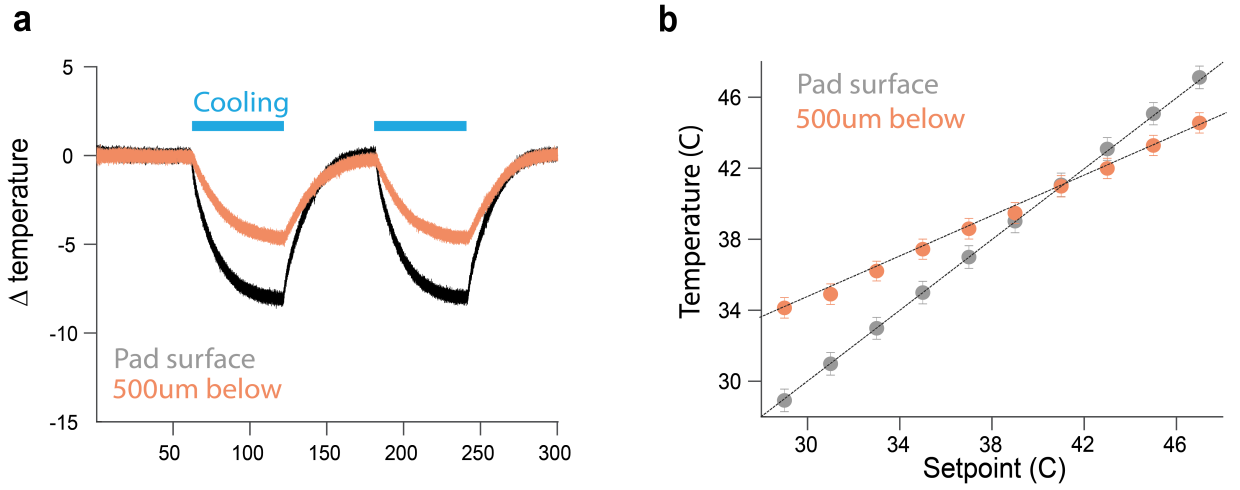


Figure 2.9: Calibration of the cooling device

(a) The temperature of the cooling pad surface and brain temperature 500 microns below the pad during cooling during two short bouts of cooling.

(b) Steady state temperatures of the pad surface, and brain 500 microns below the pad for a variety of pad command temperatures.

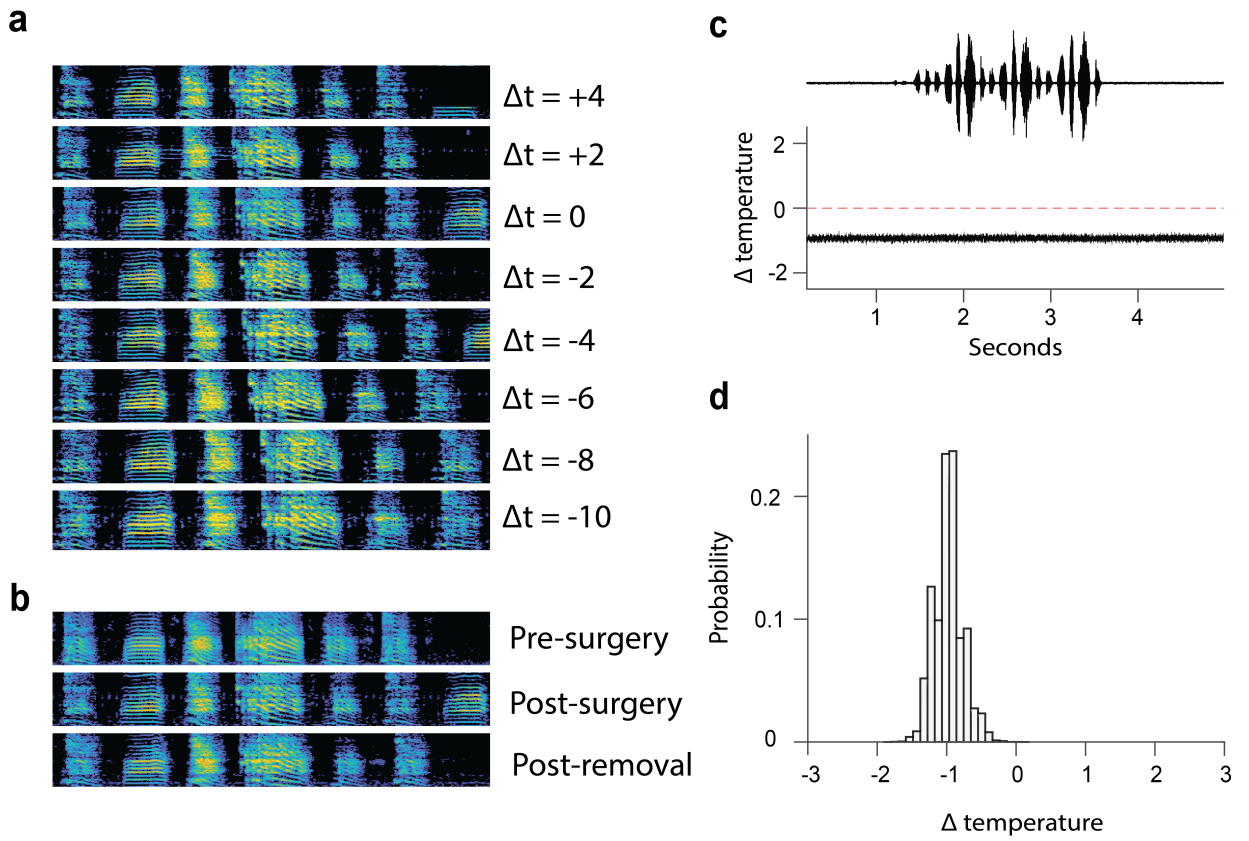


Figure 2.11: Validation of device cooling and removal in an adult bird

(a) Example spectrograms showing the effect of cooling in an adult bird during song production

(b) Example song spectrograms produced before device implantation (pre-surgery), during device implantation (post-surgery), and after removal of cooling element (post-removal).

(c) Temperature of dura surface after device removal (bottom) relative to normal HVC temperature, and sound pressure waveform of song (top). The temperature of the dura above HVC is expected to be slightly less than internal body temperature. Note that as shown in figure 2.9, a 1°C difference in temperature would lead to change in HVC temperature that is only a fraction of this amount.

(d) Histogram showing the distribution of HVC dura temperature post device removal, relative to normal HVC temperature.

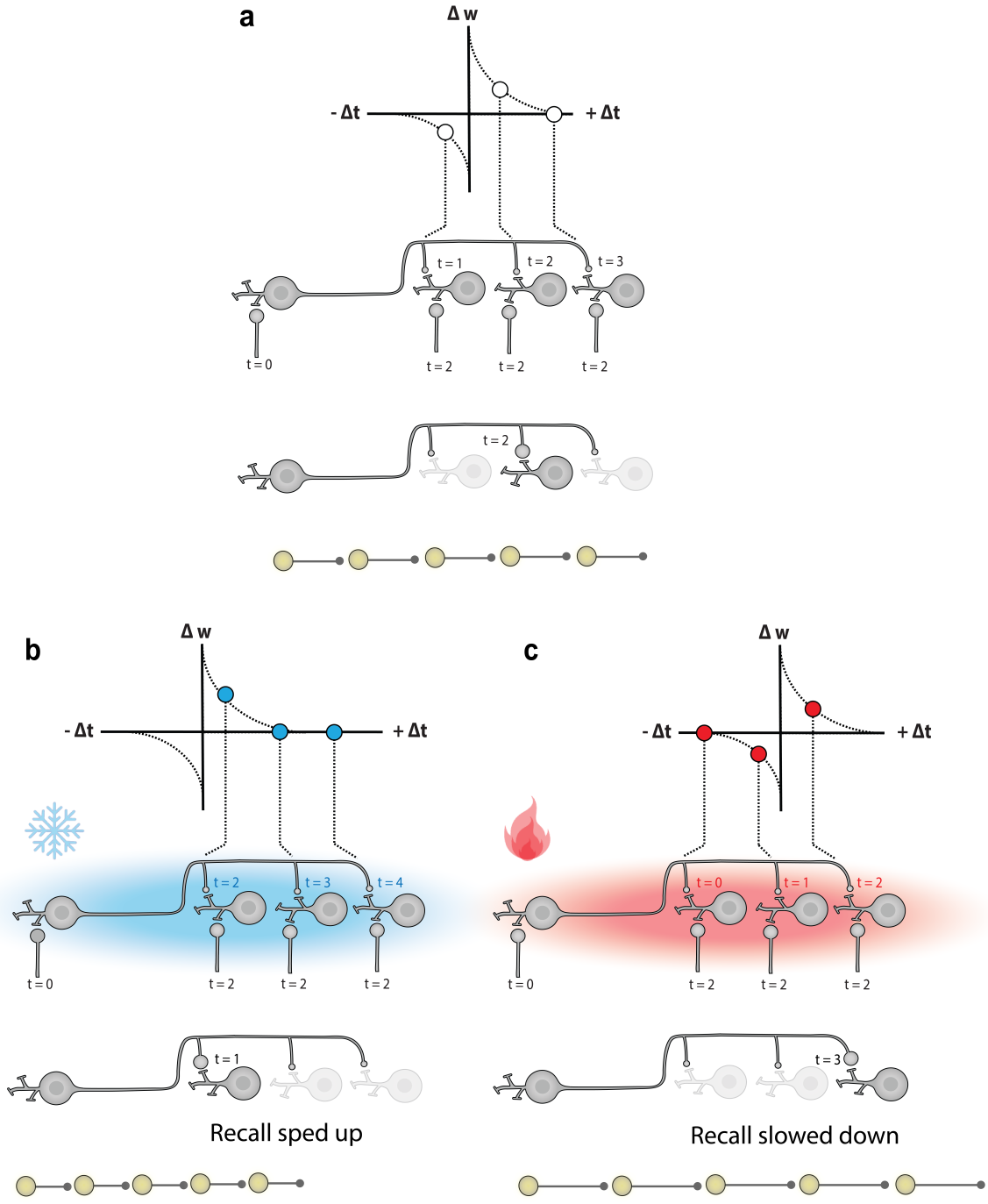
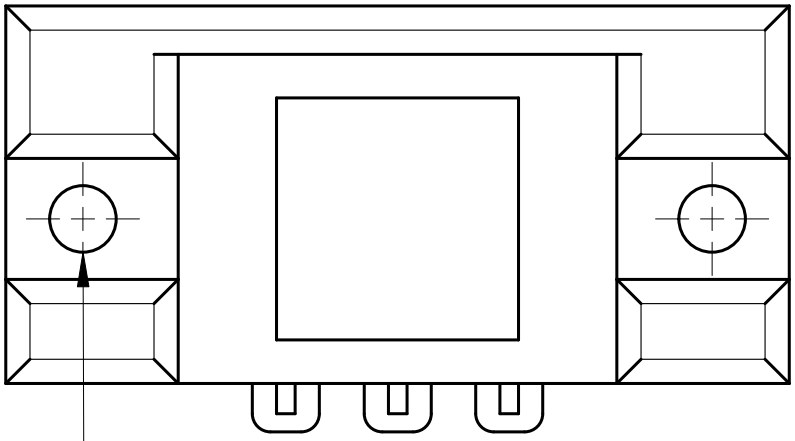

a
Δ w
- Δt
+ Δt
t = 1
t = 2
t = 3
t = 0
t = 2
t = 2
t = 2
t = 2
b
Δ w
- Δt
+ Δt
t = 2
t = 3
t = 4
t = 0
t = 2
t = 2
t = 2
t = 1
Recall sped up
c
Δ w
- Δt
+ Δt
t = 0
t = 1
t = 2
t = 0
t = 2
t = 2
t = 2
t = 3
Recall slowed down

Figure 2.14: Proposed mechanism for observed effect of temperature on temporal structure of imitation

(a) Top: cartoon of hypothesized normal learning mechanism between excitatory HVC neurons during tutoring. A pre-synaptic neuron (driven by strong feedforward auditory input at time t=0) and three post-synaptic partners (driven by strong feedforward auditory inputs at time t=2)). Three synaptic events between these neurons occur at time t=1, t=2 and t=3. Units of time are arbitrary here but assumed to be on the order of milliseconds (Intra-HVC transmission delays were previously estimated to range between 0 and 12ms [54])). Above the schematized neural circuit is a plot showing hypothetical spike-time-dependent-plasticity at each of these synapses. At a normal temperature the input to the middle neuron is strengthened during learning because synaptic input occurs at nearly the same time as post-synaptic spiking. Bottom: a cartoon showing the effective connectivity in this circuit after learning. Note that after learning, the presynaptic neuron activates its post-synaptic partner at time t=2, mirroring the timing during learning. We hypothesize this happens between multiple pairs generating the neural chain shown below.

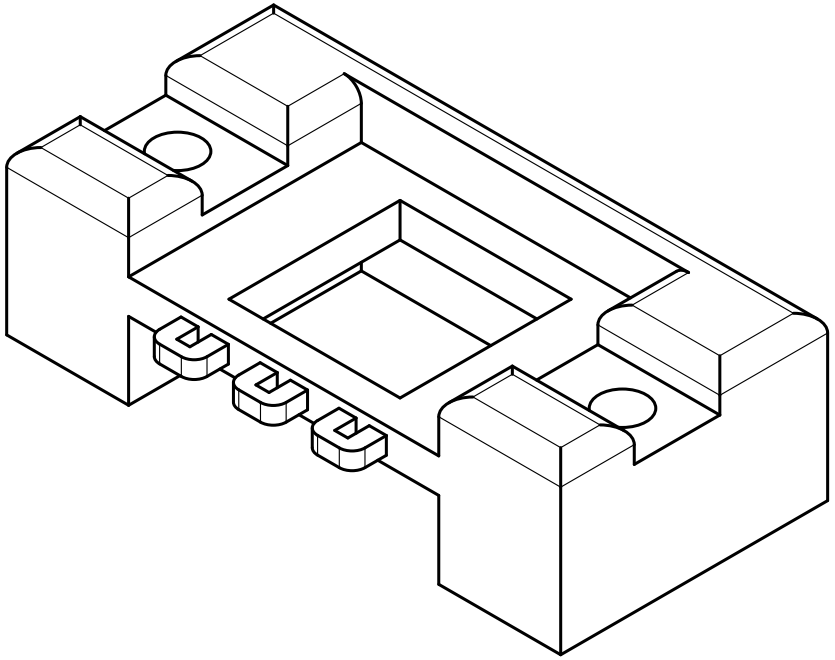
(b) Cartoon showing the same learning on the circuit shown in a, but now during mild cooling. In this case, the timing of synaptic events between synaptically connected HVC neurons is delayed due to presumed slowing of neuronal activation, action potential propagation, synaptic release etc. Note that the relative timing of auditory inputs do not change. This leads to strengthening of a different synaptic connection that normally arrives too fast. After learning, the presynaptic neuron activates its post-synaptic partner at time t=1, faster than the timing during learning. Note this leads to the formation of a neural chain that propagates faster.

(c) Same as in b, now in the case of mild heating. Note that in this case biophysical processes are expected to speed up. After learning, the presynaptic neuron activates its post-synaptic partner at time t=3, slower than the timing during learning. Note this leads to the formation of a neural chain that propagates slower.

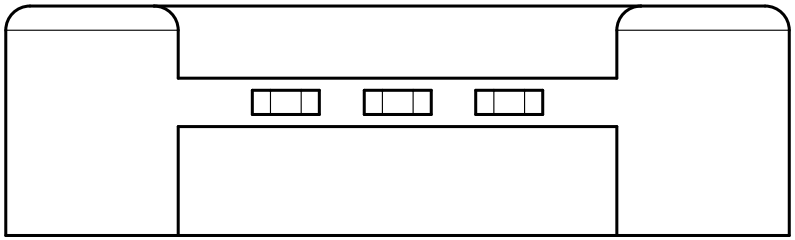


SCALE 8:1

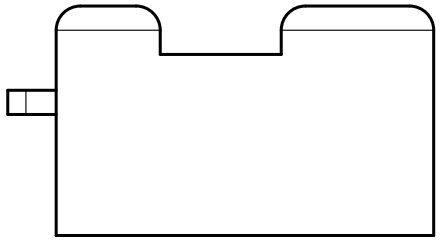
Ø.043 THRU



SCALE 8:1



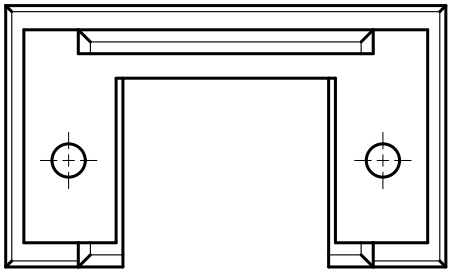
SCALE 8:1

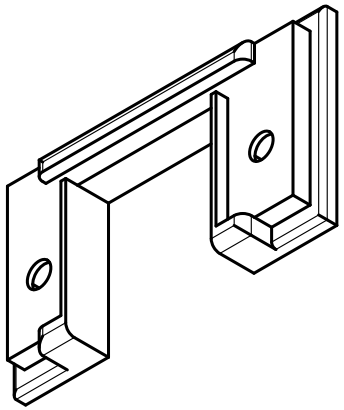


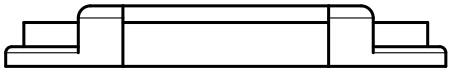
SCALE 8:1

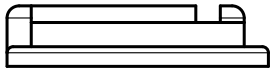
UNLESS OTHERWISE SPECIFIED, DIMENSIONS ARE IN INCHES .XX = ±.0- .XXX = ±.00- ANGULAR = ± ° .XXXX = ±.000- FRACTIONAL = ± SURFACE FINISH		NAME	DATE
	DRAWN	ANDREW BAHLE	03/12/2024
	CHECKED		
	APPROVED		
DO NOT SCALE DRAWING			
BREAK ALL SHARP EDGES AND REMOVE BURRS			
THIRD ANGLE PROJECTION	MATERIAL	FINISH	

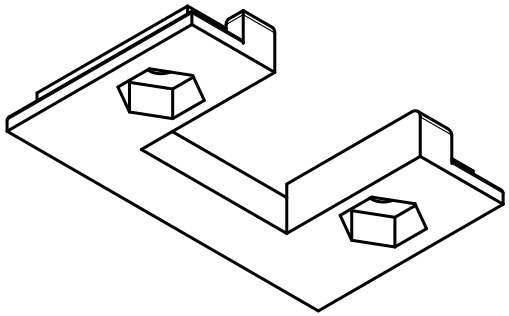
TITLE		
SIZE A	DWG NO.	REV.
SCALE 4:1	WEIGHT	SHEET 1 of 1

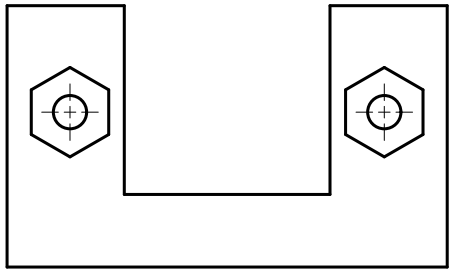












UNLESS OTHERWISE SPECIFIED, DIMENSIONS ARE IN INCHES .XX = ±.0- .XXX = ±.00- ANGULAR = ± ° .XXXX = ±.000- FRACTIONAL = ± SURFACE FINISH		NAME	DATE
	DRAWN	ANDREW BAHLE	03/12/2024
	CHECKED		
	APPROVED		
DO NOT SCALE DRAWING			
BREAK ALL SHARP EDGES AND REMOVE BURRS			
THIRD ANGLE PROJECTION	MATERIAL	FINISH	

TITLE

SIZE A	DWG NO.	REV.
SCALE 4:1	WEIGHT	SHEET 1 of 1

Chapter 3

Neural dynamics in a pre-motor region during auditory memory formation

3.1 Abstract

In songbirds, a pre-motor cortical region known as HVC, has been shown to play a dual role in song motor production and song audition: neurons in HVC display distinctive sparse sequences during motor production of stereotyped song [42, 56] and HVC neurons display auditory responses in sleeping or anesthetized adults as well as awake juvenile zebra finches [72, 23, 25, 89]. Tutor playback evokes activity in juvenile HVC that is gradually suppressed in an imitation dependent matter [89], resulting in a near complete lack of auditory evoked activity in the awake adult bird. This suggests a functional relationship between HVC activity during tutoring and the eventual imitation. Despite the extensive literature on HVC's auditory function, a few key questions remain outstanding: how does HVC respond during tutoring as a population, when do tutor responses arise and how stable are they across time. Here we set out to answer these questions using Neuropixels probes and functional calcium imaging. We first show that HVC neural dynamics during tutor song presentation are organized into sparse sequences that cover the entire song, reminiscent of the neural sequences during adult song production. Such sparse sequential activation is not present in most of HVC's auditory inputs. Using chronic imaging, we find that these sequences are present as early as the very first day of tutoring, but strengthen over time. These data are consistent with a model in which Hebbian plasticity builds dynamical sequences during tutoring that can then sub-serve the long hypothesized tutor song template to guide imitation.

3.2 Introduction

In order for imitation to occur, animals must store a stable sensory memory of the action to be imitated, and use that memory to guide their subsequent actions until they reach a satisfactory level of performance. Songbirds, including the zebra finch (Taeniopygia guttata), naturally imitate the song of an adult male "tutor" when juvenile male "pupils" are exposed to this tutor song during an early sensitive period [2]. During this tutoring process, pupils form a memory of the tutor song that is then used to guide motor learning; an entirely memory guided process which requires no additional additional external feedback from the tutor. A key question is where and how this tutor memory is stored during tutoring, and how this memory is recalled during vocal practice. During adult production of the vocal imitation, excitatory neurons in a pre-motor area known as HVC produce highly stereotyped, temporally sparse, sequences of activity that tile the entire song [42, 41, 40]. Furthermore, neural dynamics in HVC, but not downstream motor areas, control the timing of song production [53, 51]. In awake adult birds, HVC neurons are largely inactive or suppressed during listening [72, 89]. In juveniles however, during the period when birds are tutored, HVC neurons respond to many stimuli, and responses are biased to the tutor song [91, 25]. These responses are gradually suppressed over the course of song learning in a manner proportional to the progression of the imitation [115]. Based on these observations it has been suggested that HVC may be involved in storing and recalling the auditory tutor memory that guides motor imitation [87]. Indeed perturbation of HVC [88] or its auditory inputs [20] during tutoring, severely degrade subsequent imitation, as do lesions of neurons in HVC that project back to the auditory system [21].

Although HVC is likely involved in the formation and recall of the tutor memory, few studies have explored the response properties of HVC neurons during the very first days of tutor memory formation. An additional mystery is how activity in HVC during tutoring may relate, if at all, to HVC's function as a motor clock during adult song production. Thus

we set out to record from large populations of HVC neurons as well as the auditory inputs to HVC during the tutoring process. We first used head-fixed Neuropixel recordings to measure how these regions respond at the population level during exposure to the tutor song. We then examined the relationship between tutor evoked responses in HVC and HVC inputs and explored how this activity may relate to motor sequences produced by HVC during production of the adult imitation. Finally, we used freely-moving calcium imaging to study the responses of HVC during natural tutoring from the very first day of tutor exposure to the final day of tutoring.

3.3 Methods

3.3.1 Animal care and use

We used male zebra finches (Taeniopygia guttata) from the MIT zebra finch breeding facility (Cambridge, MA). Animal care and experiments were carried out in accordance with the NIH guidelines and reviewed and approved by the Massachusetts Institute of Technology Committee on Animal Care. For all tutoring experiments we prevented exposure to a tutor song prior to experiments; birds were separated from their father on or before 15 days post-hatch and raised by their mother and a foster female. For experiments birds were singly housed in custom made sound isolation chambers.

3.3.2 Surgical procedures

Juvenile zebra finches (35-55 dph) were anaesthetized with 2% isoflurane gas before being placed in a custom stereotaxic apparatus. After applying a topical anaesthetic (0.1% bupivacaine) and performing a midline incision in the skin over the skull, we made craniotomies in the skull at a predetermined distance from the bifurcation of a major blood vessel (the lambda sinus). For calcium imaging experiments we used a glass pipette attached to a

pressure injection system (Drummond Nanoject II), to make unilateral injections of a virus containing GCaMP6f (AAV9.CAG.GCaMP6f. WPRE.SV40 [155]) and implanted a gradient refractive index lens (GRIN, 1mm diameter/4mm length, Inscopix) above the injection site. A small durotomy was made at the injection site, and the lens was gently pressed against the brain tissue and secured to the skull using dental acrylic. For head-fixed extra-cellular recordings, unilateral or bilateral craniotomies were made in the same manner as described above and sealed with silicone elsatomere (kwik-kast). Craniotomies were also made over areas used for antidromic verification of recording site identity (Area X was used for identification of HVC recording sites and HVC was used for identification of NIf recording sites). A small durotomy was performed and a custom bipolar stainless steel stimulating electrode [156] was implanted at stereotaxically defined coordinates and secured to the skull using dental acrylic. Finally a small (3.9 x 4.55 x 1.3 mm) stainless steel head-plate with two threaded holes (3.5.3), was affixed to the skull using dental acrylic. In both imaging and electrophysiology experiments, the incision site in the skin was closed with a tissue adhesive (VetBond). Analgesic (Buprenorphine or Meloxicam) was administered 30 minutes prior to surgery and for three additional days after the end of surgery (Meloxicam).

3.3.3 Extracellular recordings

Electrophysiological recordings were carried out in head-fixed birds using Neuropixels 1.0 probes (Phase 3A) [157] or 2.0 probes [158] and acquired using SpikeGLX software (https://billkarsh.github.io/SpikeGLX/). Spikes were sorted offline using Kilosort 2 (www.github.com/MouseLand/Kilosort2) after pre-processing and filtering with CatGT (https://billkarsh.github.io/SpikeGLX/). Clusters were manually curated with phy2 [159] to remove multiunit clusters, noise clusters and duplicates. Waveform characteristics were obtained using Cwaves (https://billkarsh.github.io/SpikeGLX/). Birds were head-fixed for the duration of recordings using a metal post (3.5.3) and swaddled in a foam restraint to prevent excessive movement. Antidromic stimulation was performed using a programmable pulse generator

(Master-8, MicroProbes for Life Sciences) and a stimulus isolator (Iso-flex, MicroProbes for Life Sciences). Stimulus pulses had durations of 100 microseconds and used currents between 100 and 200 microamperes. Target regions were identified as those with electrodes exhibiting antidromic "hash". A unit was defined as within HVC or NIf, if the electrode with the largest amplitude spike waveform fell within the antidromically identified region. Units were defined as in L1/CLM if they were at least 50 microns (to exclude units in the somewhat ambiguous boundary between these areas) above the upper antidromically defined boundary of NIf. We defined HVC_{shelf} as spanning the region beginning 200 microns below the bottom of HVC and ending 50 micron below the bottom of HVC as defined antidromically. This dorsal-ventral extent was conservatively chosen based on a previous study which measured the spread of local stimulation in HVC_{shelf} in slice [73]. In reality these regions are anatomically contiguous, however the 50 micron buffer zone was meant to exclude units near the somewhat ambiguous border.

3.3.4 Analysis of eletrophysiological data

Putative excitatory and inhibitory unit classifications were obtained using k-means clustering of the average firing rate, spike waveform asymmetry and spike width. Spike width was defined as the duration (in milliseconds) between the peak and trough in the average waveform. The peak was defined as the point at which the extracellular waveform reached its maximum and trough was defined as the point at which the average waveform reached its minimum. Spike asymmetry was calculated as the relative height of maximal amplitudes of the regions of the average waveform on either side of the trough [160]. Neurons were classified as significantly modulated by the tutor song using the non-parametric Zeta test[161], $P<0.05$, with Bonferroni correction for multiple comparisons. Average firing rates were computed using 1ms bins smoothed with a 20ms moving average. Peaks in the average motif-aligned firing rate were detected using the MATLAB findpeaks function with a minimum peak distance of 10ms and a minimum peak prominence of 5. Population sparseness was computed as the

fraction of neurons that were not significantly modulated by the tutor stimulus [162].

3.3.5 In-vivo calcium imaging

As described above, the calcium indicator GCaMP6f was expressed in HVC by intracranial injection of the viral vector AAV9.CAG.GCaMP6f.WPRE.SV40 (Addgene) [155]. Although we did not verify tropism, similar viral vectors exhibit expression restricted to excitatory projection neurons in HVC [163, 164, 165]. Even when inhibitory neurons in HVC do express calcium indicators, changes in spike rate rarely produce detectable fluorescence changes [30]. Thus in practice, neurons included in our datasets are likely excitatory projection neurons. After surgery we allowed at least a week before recording, to allow for sufficient viral expression. At that time, a baseplate (Inscopix) was attached to the implant and birds were imaged using either an Inscopix nVista miniature fluorescent microscope, or a lighter custom microscope [166]. Spatial and temporal footprints corresponding to putative single neurons were extracted from calcium imaging data using the EXTRACT algorithm [167]. Results were manually curated to exclude noise and low SNR footprints. Spatial footprints were tracked across days using the cellReg algorithm [168].

3.3.6 Analysis of calcium imaging data

To determine the significance of firing rate peaks during tutoring we used a shuffling procedure similar to one previously described [169]. For each neuron, we first calculated the average tutor aligned fluorescence. We then generated surrogate average fluorescence traces by uniformly circularly permuting the calcium fluorescence on a trial by trial basis. To ensure the number of surrogate shuffled rasters used was appropriate for determining significance at a Bonferroni corrected level of 0.05 we used $2 \times (n/0.05)$ surrogates, where n is the total number of neurons. The p-value for a given neuron was obtained using the frequency with which surrogate datasets generated maxima larger than the original average fluorescence. Peaks in average activity with $P < 0.05$ (Bonferroni corrected) were consid-

ered significant. Using significantly tutor modulated neurons tracked across all seven days of tutoring, we measured the stability of neural responses using two measures: population vector correlation and tuning curve correlation [170]. For population vector correlation, we first computed the population vector for each day of tutoring as the average motif-aligned activity of all extracted footprints. We then calculated the Pearson's correlation between the population vectors on each day of tutoring. We obtained the PV correlation as a function of days of tutoring by averaging the correlation values for PVs differing between one and five days. Tuning curve correlation was measured for each neuron by computing the Pearson correlation between the average firing rate of that neuron and its average firing rate on all other days of tutoring. We again obtained a neural tuning-curve correlation by averaging this measure for pairs of days separated by different numbers of days. Overall tuning curve correlation was the obtained by averaging this measure across all neurons.

3.3.7 Tutoring protocol

We isolated birds from their father at or before 15 days post hatch (dph), based on prior studies that achieved effective imitation in experimental animals isolated between 10 and 30 dph [1, 88, 119]. These isolated birds remained co-housed with their siblings and mother until at least 35 dph, at which point they were singly housed for the duration of the experiment [102]. Zebra finches can produce accurate imitations after as little as 75s of tutor exposure [114], but protracted daily tutor exposure impairs imitations [120]. We therefore followed an intermediate tutoring protocol in which pupils received 60 minutes of tutor exposure each day [102]. Birds were housed in a sound attenuating chamber and their songs were recorded daily using either Sound Analysis Pro software (SAP)[121], or custom recording software. For electrophysiological experiments birds were tutored by a live tutor for 1 hour per day for 5 consecutive days. Neural recordings were made on the seventh day after the onset of tutoring, to allow a night of sleep following the final tutoring session, and acclimatization to the head-fixation apparatus. For imaging experiments, recordings were made as juveniles

were naturally tutored for 1 hour each day. Recording began with the very first exposure of the juvenile to the tutor.

3.3.8 Auditory stimuli

During electrophysiological recordings, the tutor song was presented in bouts of five motifs. A total of 150-200 motifs were presented in 30-40 overall bouts. Bouts of tutor song were interspersed with other songs (data not shown, see chapter 4), and each song type was randomly presented. Presentations were flanked by silent periods of random durations, drawn from a uniform distribution between 3 and 13 seconds. Song recordings were played-back over a speaker (8010A, Genelec or PDMR5, Pyle) placed ∼25cm from the birds head. Playback was acquired via a microphone placed next to the head of the bird, and the sound amplitude of playback was calibrated to match a live tutor singing directed song at the same location as the speaker (SPL of ∼75 dB at 25cm).

3.4 Results

3.4.1 Tutor evoked HVC activity is organized into sparse neural sequences

If learned HVC dynamics control the subsequent timing of song imitation (see chapter 2), then what patterns of neural activity support this function? To investigate auditory population activity in HVC neurons, we carried out high-throughput recordings (1655 neurons in 10 birds), in juvenile birds during presentation of their tutor song. Even at the level of single trials, repeated sequential patterns of activation were visible across the full population of HVC units (figure 3.1b). We further identified a subset of neurons that were significantly modulated by the tutor song using the non-parametric zeta test [161]. We found that individual tutor responsive units identified in this way often fired reliably at a single time in

the tutor song (figure 3.1c). As a population these neurons formed a reliable sequence that spanned the entire tutor motif (figure 3.1d), similar to the sequential firing of HVC neurons observed during adult song production [42, 41, 40].

3.4.2 Putative excitatory and inhibitory HVC neurons both participate in neural sequences

While the firing rate profiles during tutor playback of many HVC neurons were temporally sparse, some neurons fired at high rates throughout the tutor motif as may be expected for inhibitory interneurons. We thus classified units into putative excitatory and inhibitory sub-types based on average firing rate and the shape each individual spike waveform (figure 3.3b). Applying k-means clustering to spike width, spike asymmetry and average firing rate, yielded two clear cell types which are likely to correspond to inhibitory and excitatory cell types. While 33% of recorded putative inhibitory units were significantly modulated by playback of the tutor song, 13% of putative excitatory units were significantly modulated by playback (figure 3.3c). We next asked if the response profiles of putative excitatory and inhibitory cell types tiled the tutor motif uniformly as they do during adult motor production [40]. For each unit, we identified peaks in the average motif-aligned firing rate and found that as a population these peaks were distributed approximately uniformly throughout the tutor motif, including silent gaps (figure 3.3,a). Putative excitatory units tended to have a single peak in their average motif-aligned firing rate, while putative inhibitory neurons tended to have several (figure 3.3 d). This is again reminiscent of the firing rate profiles of HVC excitatory and inhibitory units during adult song production [42, 40, 41].

3.4.3 Tutor evoked responses in HVC are sparser than its auditory inputs.

While it is notable that HVC population activity during tutoring is organized into sparse, stereotyped sequences as during adult song production, one possibility is that this population activity is inherited from the auditory areas that provide input to HVC (unlike the case of motor sequences which emerge within HVC proper [171, 172]). To test this possibility we also recorded from several auditory responsive areas that provide input to HVC: the nucleus interface (NIf), the L1 portion of field L, the caudolateral mesopallium (CLM) and HVC_{shelf} [26, 173, 174, 73, 175]. We first measured the population sparseness of these regions during tutoring and found that HVC tended to be much sparser than NIf, or CLM and L1 (figure 3.5 b). In these areas, roughly 50% of the neurons were significantly active to the tutor song while in HVC, 10-20% of neurons were active during tutor playback. The populations sparseness in HVC and was higher than has been observed an in any other auditory region in the songbird [162]. In addition to the overall population sparsity we also wondered if neurons in areas providing input to HVC were temporally sparse, that is, they tended to have many peaks in their average firing rate as is common in thalamo-recipient auditory areas [176, 162]. Putative excitatory units in HVC tended to have a single peak in their average firing rates, while putative excitatory units in L1/CLM and NIf tended to have more than one peak (figure 3.5). Putative inhibitory units however, were similar in all regions and tended to have several peaks in their average firing rate distributed throughout the tutor motif. Together this indicates that responses in HVC are sparser both in terms of the population recruited during tutoring and in terms of the temporal profile of motif-aligned firing rates, relative to areas that provide input to these regions. Interestingly, responses in HVC_{shelf}, from which HVC is thought to have evolved [177], where highly similar to HVC in terms of population sparseness and frequency of firing rate peaks in excitatory neurons. This may suggest that these related regions are specialized for sparse representation of song. It is also

possible that this arises due to the fact that HVC provides far more input to HVC_{shelf}, than HVC_{shelf} provides to HVC [73].

3.4.4 Chronic calcium imaging of HVC population activity during natural tutoring

We wondered whether HVC activity during unrestrained tutoring was also organized into sparse sequences. We therefore used virally expressed GCaMP6f and head mounted miniature microscopes to record neural activity in freely moving juveniles during natural tutoring. The spatial footprints and temporal activity of putative neurons were identified with the EXTRACT algorithm [167] and the results were curated to remove low SNR and noise units with unphysiological firing patterns. As in our extracellular recordings, we observed individual HVC neurons that fired reliably at a single time in the tutor motif (figure 3.7c). As a population this activity was again organized into a sparse sequences that spanned the entire tutor motif, (figure 3.7d). HVC neurons recorded in this manner were likely excitatory projection neurons based on both their pattern of firing activity [30] as well as the observation in prior studies that expression of a related GCaMP6f virus in HVC tends to be restricted to excitatory neurons [163, 164, 165]. Indeed prior work has shown that even when HVC interneurons do express calcium indicators, transient calcium events are essentially never seen in these cells, perhaps due to their calcium buffering properties [30]. Thus the activity of HVC excitatory neurons is organized into sparse sequences during natural tutoring as well as head-fixed tutor playback.

3.4.5 Tutor evoked responses are stable in time but grow in reliability

Another advantage of calcium imaging is that it allows neurons to be easily tracked across multiple days of tutoring allowing us to ask how tutor evoked sequence in HVC emerge and

change over the course of tutoring. Are responses present from the first day of tutoring or do they develop slowly over time? Additionally, are response profiles of neurons stable from day to day or do they exhibit so-called representational drift, changing their tuning to the tutor song over days and weeks? To address these questions we recorded calcium responses in extracted HVC neuron footprints as we tutored every day for a week, beginning with a juveniles very first day of exposure to the tutor song. Although many neurons in the field of view were lost between subsequent days, in two birds we were able to reliably track populations of HVC neurons over the entire week of tutoring (figure 3.7 b). Restricting our analysis to the first day of tutoring, we indeed found that neural sequences were present from the very first day of tutoring suggesting that sparse neural sequences are generated by hardwired connectivity in HVC (figure 3.9). By tracking neurons across multiple days we found that some neurons were extremely stable across the entire week (figure 3.11 b), however in many neurons responses became stronger and/or more reliable over the course of the week resulting in a stabilization of the overall sequence (figure 3.13).

3.5 Discussion

3.5.1 The role of sparseness in tutor learning

Our findings reveal that HVC excitatory neurons represent the tutor song using an extremely sparse code that likely emerges in HVC itself. Because most HVC excitatory neurons had a single peak in their tutor evoked firing rate, each moment in the tutor song was represented by a unique neural state which was near orthogonal to the representation at all other times in the tutor song. This is notable because even moments in the tutor song with correlated spectral content, or silent gaps, were mapped to unique states in HVC. This mapping process is comparable to the use of hash function in computer science for efficient memory storage and retrieval. Interestingly, our observation that 13% of excitatory neurons in HVC respond to the tutor song is reminiscent of the sparse expansion coding in cerebellar-like circuits [178,

179] such as the olfactory Kenyon cells in insects where individual odors elicit responses from a random 10% of neurons [180, 181, 182, 183]. These sparse representations of odors by KC's are optimal for a class of learning problems [184] and emerge naturally when training machine learning systems with matched neural architectures [185]. Indeed manipulations which alter the spareness of KC cell representations either by altering excitability or the density of synaptic input degrade flies ability to discriminate and form memories [186, 187]. Given that HVC is necessary for the formation and recall of the tutor memory, the sparse sequential responses we observed during tutoring may allow for efficient and rapid storage of the tutor memory, while ensuring that the memory of each moment in the song is stored separately. While it is unclear what mechanisms generate this sparse representation it could arise in principal through a combination of strong inhibition, weakly correlated coincident inputs and/or an expansion in the number of cells representing the song (for example NIf to HVC). Future work should address the mechanisms generating spareness and test whether sparseness of the HVC representation dictates learnability, as is the case of the insect olfactory system.

3.5.2 A possible relationship between tutor sequence and adult motor sequences

The pattern of activity we observed in HVC during tutoring was strikingly similar to HVC motor sequences observed in adults, and as such it is tempting to draw a connection between the activity of HVC during these two very different processes of juvenile tutoring and adult song imitation. Tutor evoked sequences became more stable over the course of tutoring, a property that could be explained by Hebbian strengthening between neurons in the tutor sequence. Consistent with this idea, tutoring leads to rapid changes in excitatory and inhibitory synaptic connections in HVC [102] and blocking NMDA signaling in HVC during tutoring severely impairs song imitation [149]. We hypothesize that tutor evoked sequences form the basis for the motor sequences that emerge during song learning, forming a long hypothesized tutor template memory [188, 105, 189, 87]. This would provide a mechanism

for birds to not only store a sequential memory of tutor song timing, but also to recall it during early vocalizations to guide learning. These tutor memory sequences could then act as a scaffold-like substrate for the growth of motor population dynamics in HVC [56, 55]. This would ensure a one-to-one mapping between the number and length of song elements in the tutor song and in the emerging imitation. A key test of this model would be to record the same sets of HVC neurons during both tutoring and adult song imitation to test whether tutor sequences are reactivated both during early practice and during production of the final imitation. Circumstantial evidence is consistent with this idea; auditory vocal mirroring between song production and playback of birds' own adult song (which ought to be very similar to the tutor song in most cases) has been observed in HVC neurons in Swamp Sparrows [27, 31], Bengalese finches [31, 99] and small number of Zebra Finch neurons during sleep [72]. Future work should focus on tracking tutor sequences from motor learning to adult imitation to directly test this idea.

3.5.3 The instantaneous nature of tutor evoked sequences

The tutor evoked sequences we observed in HVC were present as early as the very first exposure to the tutor song, suggesting that sparse sequential auditory representations in HVC can arise via pre-existing hard-wired connectivity. It may seem surprising that tutor evoked sparse sequences are not learned gradually (although we did observe gradual increase in the reliability of neural responses), but are present from the very first exposures to the tutor song (see also figure 3.9). It is possible however, that the instantaneous nature of these sequences may be the key property that underlies birds' ability to perform near-one-shot learning of the tutor song. Juvenile song birds are faced with the problem of rapidly forming a memory of the timing and content of *any*[1] song they may hear during their sensitive period. To solve this problem songbirds may have evolved an auditory system which decorrelates and

[1]In the case of finches, *any* refers to any imitable or *species-appropriate* song. A similar principal may be applicable in the case of crows, parrots, starlings and other more promiscuous vocal imitators where imitable and species-specific are of course not synonymous and the sensitive period is extended throughout adulthood)

sparsifies auditory representations of songs as you move up the auditory hierarchy, resulting in extremely sparse representations in HVC and ensuring a unique, sequential representation of any song with which a bird is tutored (figure 3.15). These sparse sequences could be intrinsically specified by pre-existing input into HVC and then rapidly mapped into the internal dynamics of HVC itself. This rapid mapping of sequences from an input driven mode, to a fully internally driven mode would allow HVC to autonomously reproduce the precise temporal pattern induced by the tutor song, offline and/or during vocal practice. Sparse tutor sequences could then be repurposed during motor production and mapped to denser and denser motor representations as motor commands move down the motor hierarchy [190], finally resulting in the highly temporally correlated commands sent to the syringeal muscles controlling vocalization (figure 3.15). We hypothesize that this is the function of innately produced sparse tutor sequences in HVC and that the presence of these representations is required for learning, reflecting an inductive bias over song-learnability. Further studies of the songbird system may give important insights into one-shot learning generally, an important aspect of human intelligence that has proved very difficult to capture in artificial learning systems [191, 192, 193].

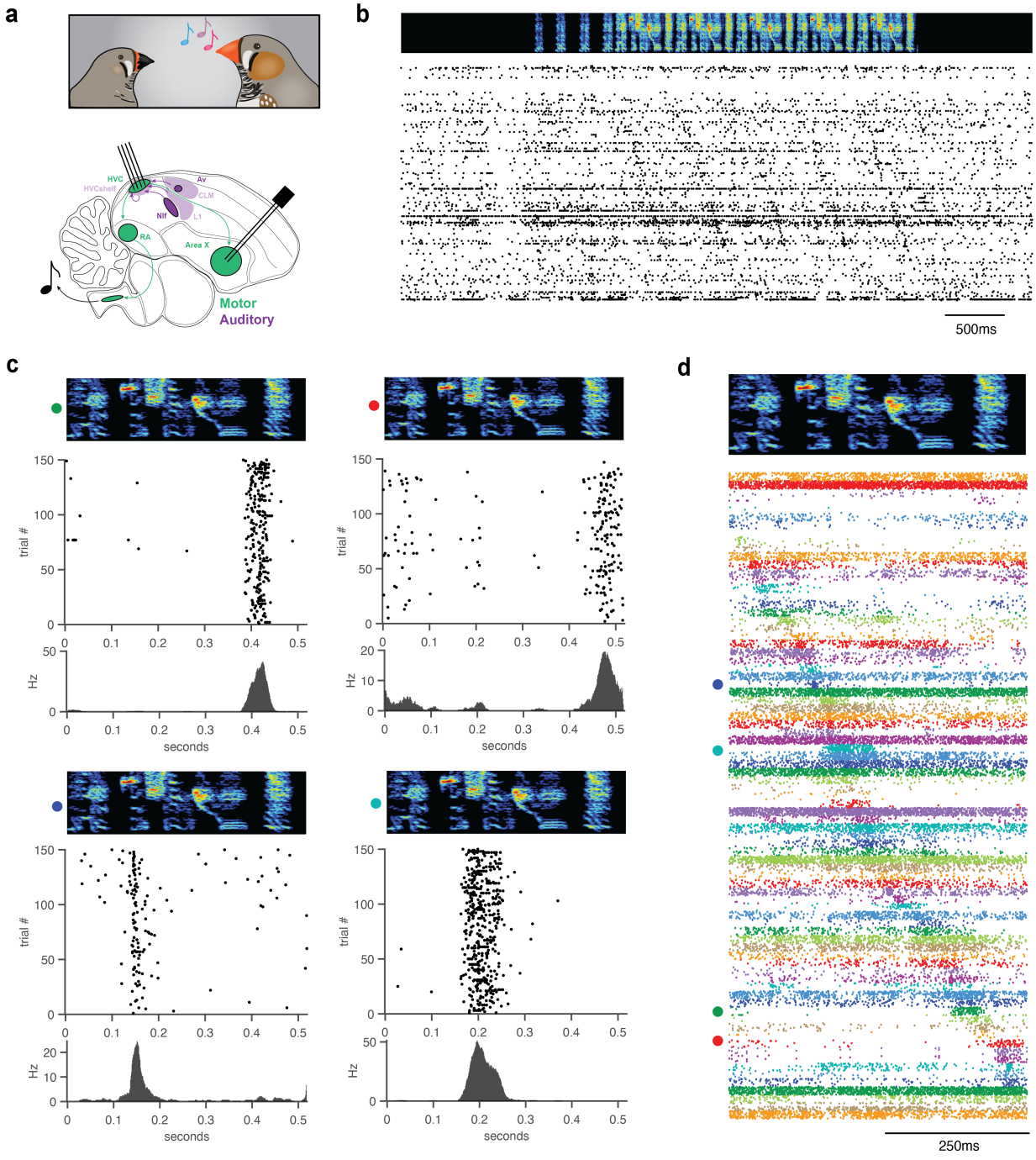


Figure 3.1: Auditory sequences in HVC during tutor revealed with high-density electrophysiology

(a) Cartoon showing tutoring process and a schematic showing anatomy and targeting of electrophysiology. Silicon probes (Neuropixels) were inserted into HVC and boundaries of HVC were determined by antidromic stimulation of HVC axons in Area X.

(b) Spiking activity of 256 simultaneously recorded HVC neurons as a bird listened to playback of a bout of tutor song containing five repetitions of the tutor motif (top). Note that even on a single trial, repeated sequential activations of neurons are visible in the spiking data. Neurons are sorted based on the latency of their peak average firing rate, aligned to the tutor motif across all trials

(c) Four example, putative single units, recorded during tutor playback in HVC. Top: spectrogram of the tutor song. Middle: spike raster plot aligned to the onset of the tutor motif. Bottom: motif aligned average firing rate.

(d) Tutor motif aligned raster plot of all 81 significantly tutor modulated HVC neurons from an example bird. Each color shows a different neuron and position is sorted based on the latency of the maximum of each neurons average firing rate.

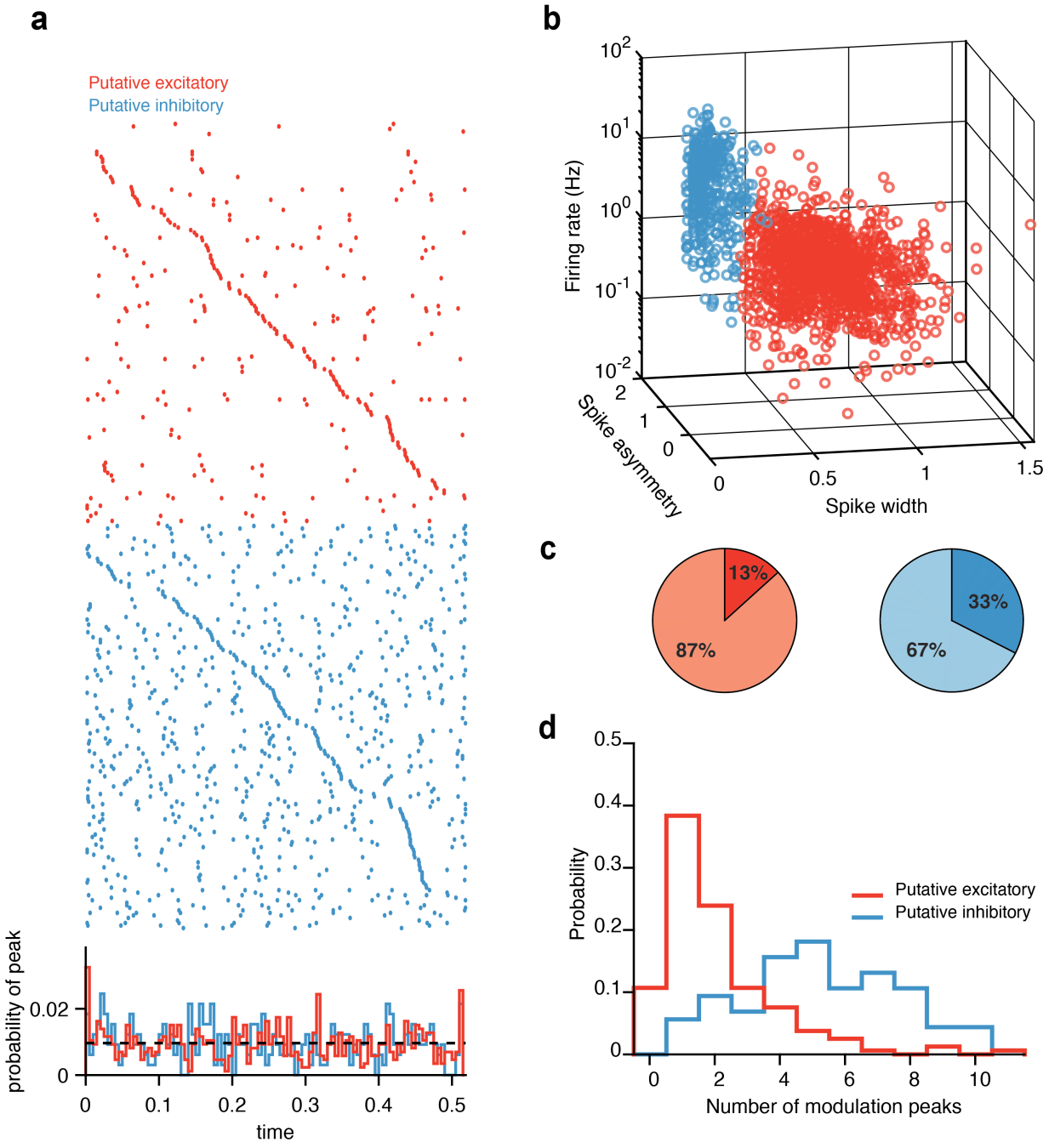


Figure 3.3: Uniform tiling of tutor motifs in putative excitatory and inhibitory HVC units

(a) Top: location of peaks in the PSTH of excitatory neurons (red, n=159 neurons in 10 birds) and inhibitory neurons (blue, n=160 neurons in 10 birds), aligned to the tutor motif. Bottom: histograms showing the density of peaks throughout the tutor motif. Dotted line shows the expectation under a uniform distribution

(b) Plot showing spike waveform characteristics for all single units in HVC (n=1655 in 10 birds). Neurons are colored based on a K-means clustering with k=2. Blue units were assigned to a putative inhibitory class and red units were assigned to a putative excitatory class

(c) Pie plots showing the fraction of units significantly modulated by the tutor song by class. Left: 14% of putative excitatory units were significantly modulated by the tutor song (zeta test, $p<0.0001$, n=1196). Right: 33% of putative inhibitory units were significantly modulated by the tutor song (zeta test, $p<0.0001$, n=492).

(d) Distribution of number of peaks in excitatory (red, n=159 neurons in 10 birds) and inhibitory (blue, n=160 neurons in 10 birds) neurons. Most excitatory neurons had a single peak in their firing rate histogram, while most inhibitory neurons had several.

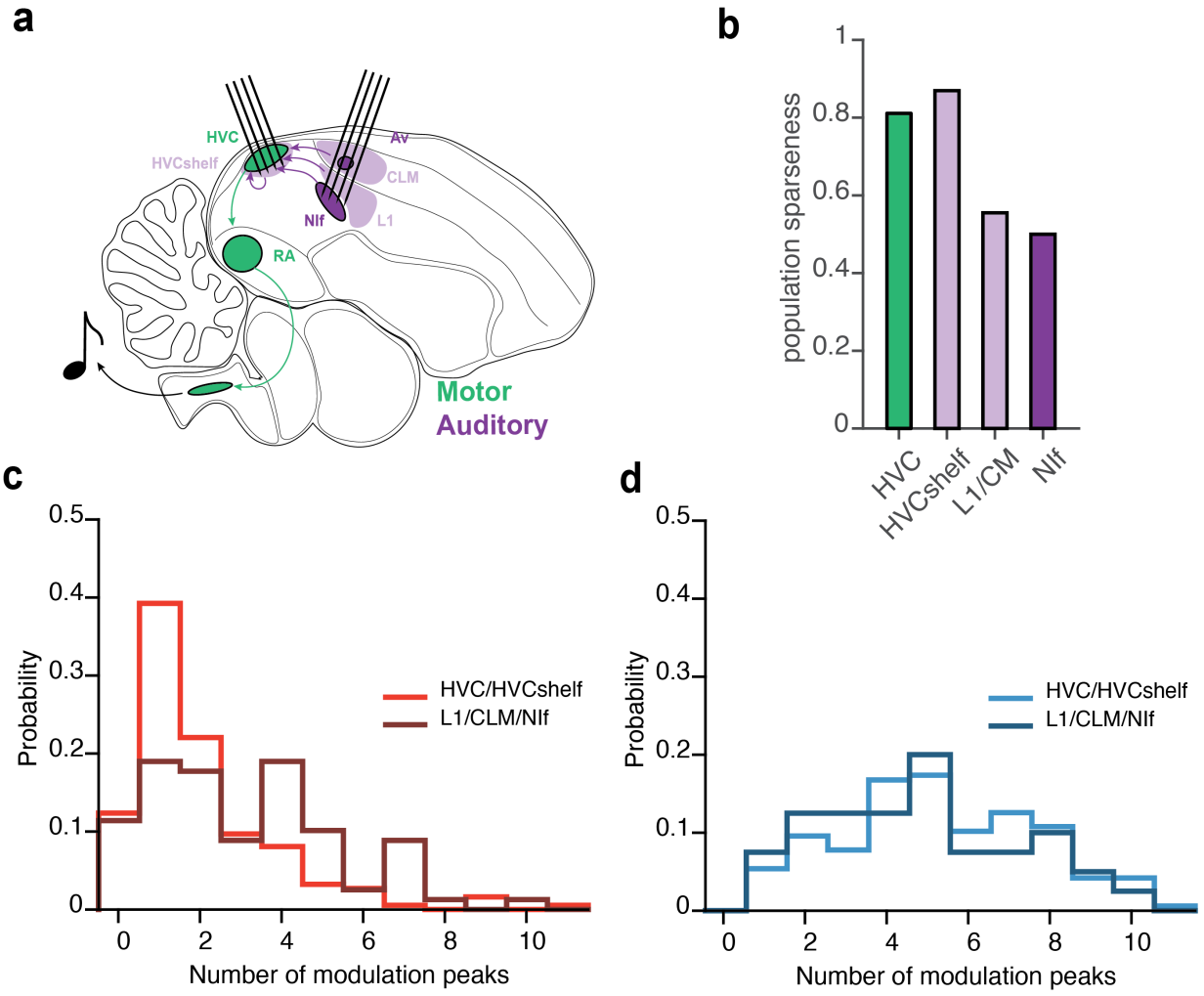


Figure 3.5: Increasing sparseness of tutor evoked activity in HVC and HVC_{shelf}

(a) Schematic showing recording strategy. Electrodes spanned several known auditory inputs to HVC.

(b) Population sparseness of responses in five regions. HVC (n=1688 neurons in 10 birds), HVC_{shelf} (n=260 neurons in 10 birds), L1,CLM (n=173 neurons in 2 birds) and NIf (n=60 neurons in 2 birds). NIf units were identified by antidromic stimulation of NIf axons in HVC. L1 and CLM units were combined as the border between these regions was easily identified. HVC and HVC_{shelf}, were much sparser than L1/CLM and NIf.

(c) Distribution of number of peaks in excitatory neurons recorded in either HVC/HVC_{shelf} or L1/CLM/NIf (auditory areas, n=79 neurons in 2 birds, HVC n=159 in 10 birds). HVC and HVC_{shelf} tended to have a single peak during the tutor motif, while neurons in L1/CLM/NIf tended to have multiple peaks during a single motif

(d) Same as in c but for putative inhibitory neurons (auditory areas, n=40 neurons in 2 birds, hvc n=160 in 10 birds). There was no difference in the distribution of peaks between different regions.

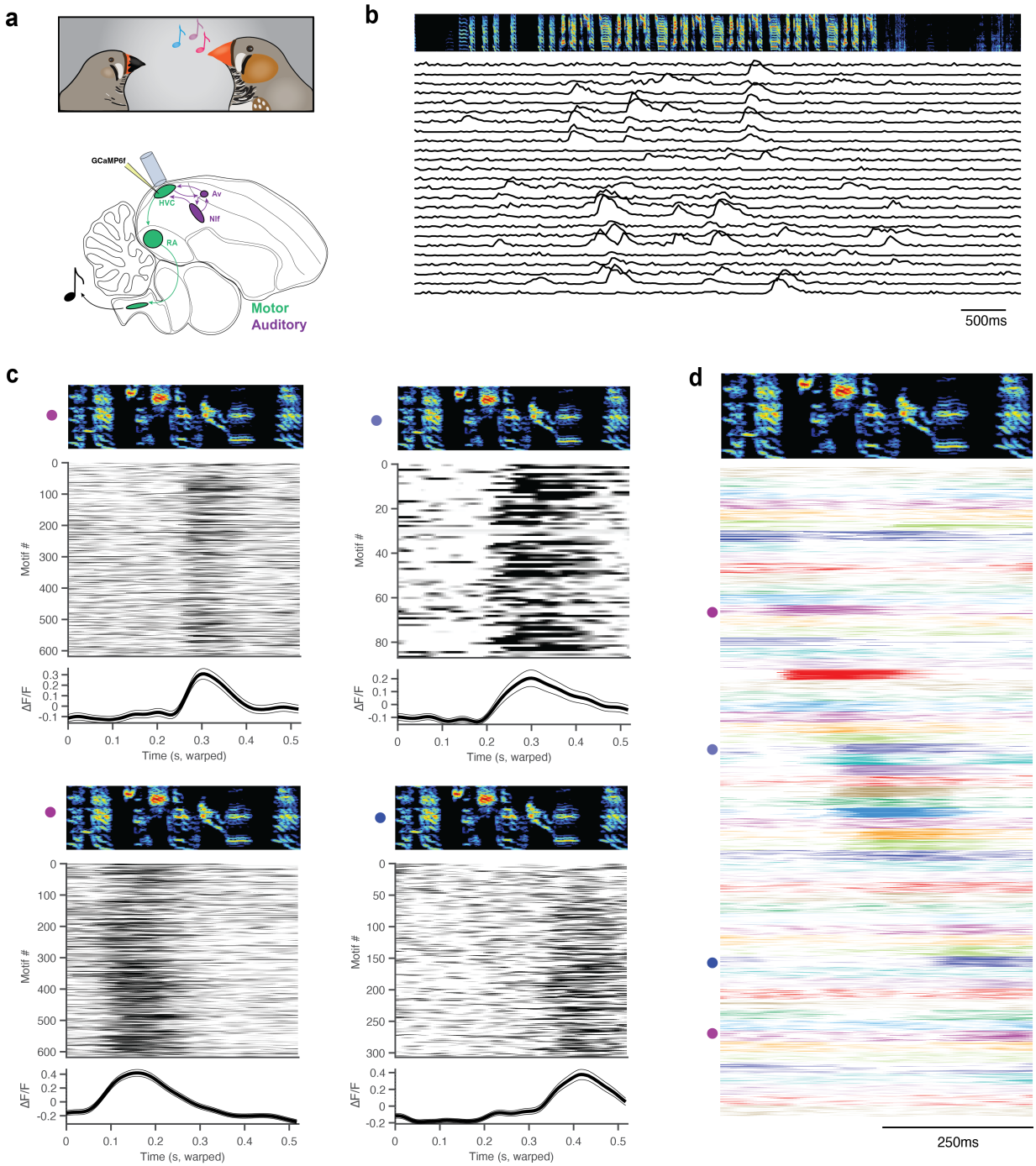


Figure 3.7: Auditory sequences in HVC during tutor revealed with chronic calcium imaging

(a) Cartoon showing tutoring process and a schematic showing anatomy and targeting of viral GCaMP6 and the GRIN lens.

(b) Extracted calcium fluorescence activity of 25 simultaneously recorded HVC neurons as a bird listened to the song of a live performance of a tutor bird (top). Neurons are sorted based on the latency of their peak average calcium intensity, aligned to the tutor motif across all trials

(c) Four examples single units recorded during natural tutoring in HVC. Top: spectrogram of the tutor song. Middle: normalized calcium raster plot aligned to the onset of the tutor motif. Activity is warped to align activity to a single tutor motif. Bottom: motif aligned average calcium activity.

(d) Tutor motif aligned raster plot of 61 HVC neurons from an example bird. Each color shows a different neuron and position is sorted based on the latency of the maximum of a neurons average activity.

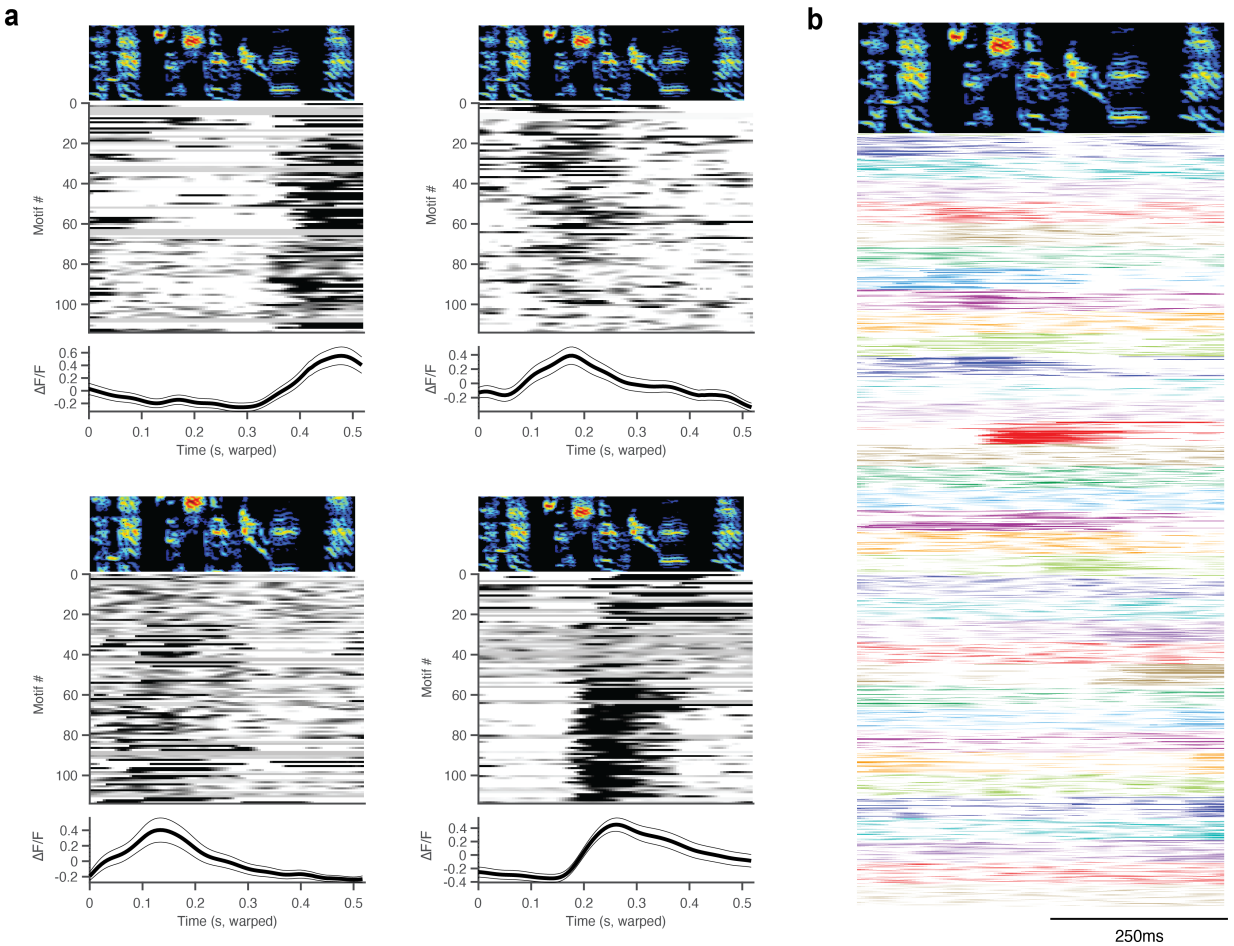


Figure 3.9: Tutor sequences in HVC during are present from the first day of tutoring

(a) Four examples neurons recorded in HVC during the very first hour of natural tutoring. Top: spectrogram of the tutor song. Middle: normalized calcium raster plot aligned to the onset of the tutor motif. Activity is warped to align activity to a single tutor motif. Bottom: motif aligned average calcium activity.

(b) Tutor motif aligned raster plot of 35 HVC neurons from an example bird. Each color shows a different neuron and position is sorted based on the latency of the maximum of a neurons average activity.

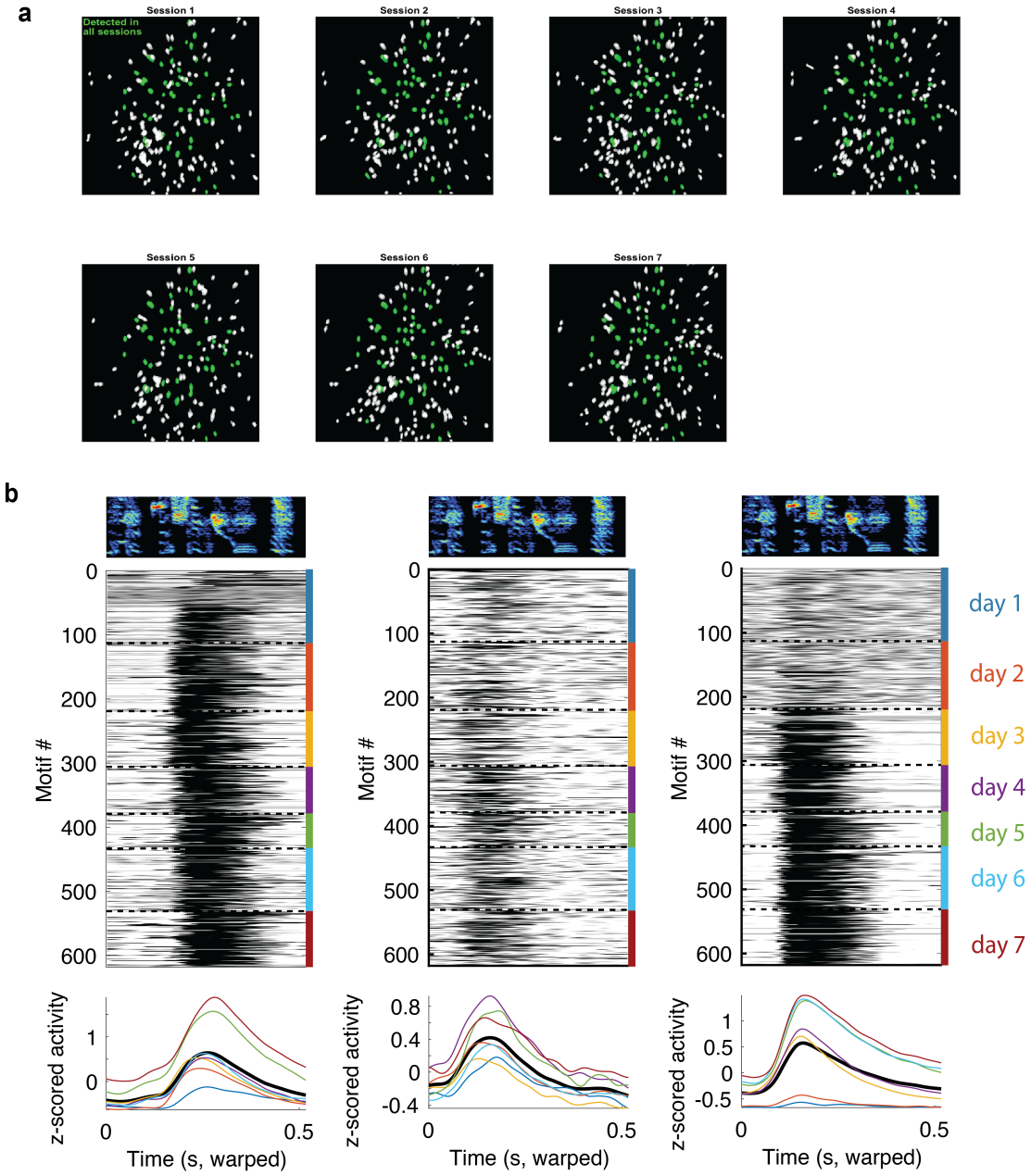


Figure 3.11: Stability of neural responses over a week of tutoring

(a) Tracking of footprints across multiple days of tutoring. Footprints were estimated using EXTRACT [167] and tracked across days using CellReg [168]. Each session shown corresponds to a single day of tutoring and the first session is the very first exposure of the juvenile to its tutor song. Estimated spatial footprints corresponding to single neurons are shown in white and footprints tracked across all 7 days are shown in green.

(b) Three example simultaneously recorded units tracked over a week of tutoring. Top: spectrogram of the tutor song. Middle: normalized calcium raster plot aligned to the onset of the tutor motif. Day boundaries are shown by dotted lines and indicated by color to the right. Activity is warped to align activity to a single tutor motif. Bottom: motif aligned average z-scored calcium activity. Colored lines show average motif aligned activity for individual days and black line shows average motif aligned activity for the entire week.

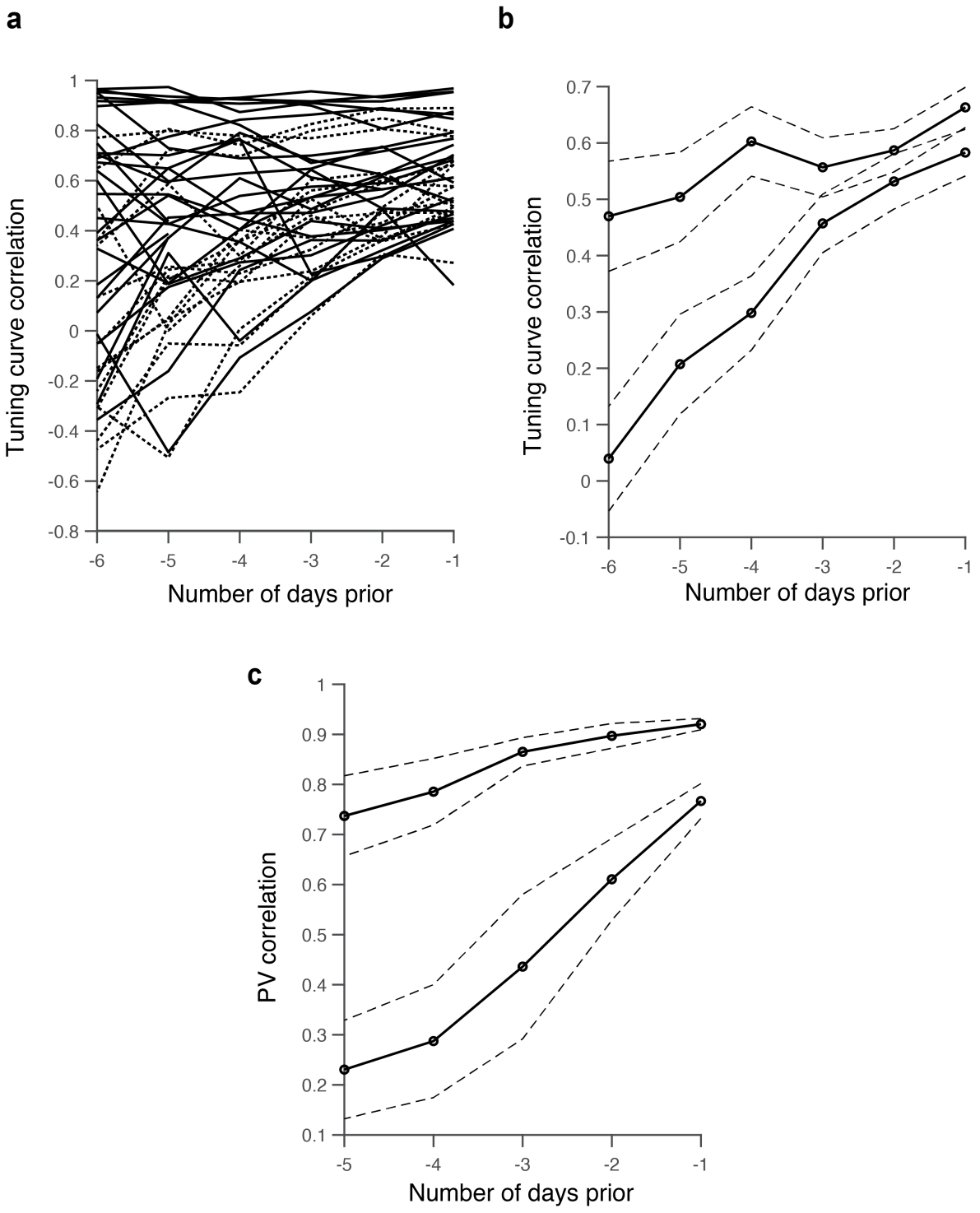


Figure 3.13: Stability of neural responses over a week of tutoring

(a) Tuning curve correlation for all tutor modulated neurons tracked over a week of tutoring in two birds (bird 1 neurons indicated by a solid line n=24, bird 2 neurons shown with dotted lines n=14). Each line is an individual neurons. Data points shows the average correlation between a neurons motif and on a given day and its tuning curve some number of days prior(between 1 and 6).

(b) Average tuning curve correlation for the two birds shown in a. Dotted lines show mean -/+ SEM

(c) Average population vector correlation between all days of tutoring over a week for two birds. Dotted lines show mean -/+ SEM

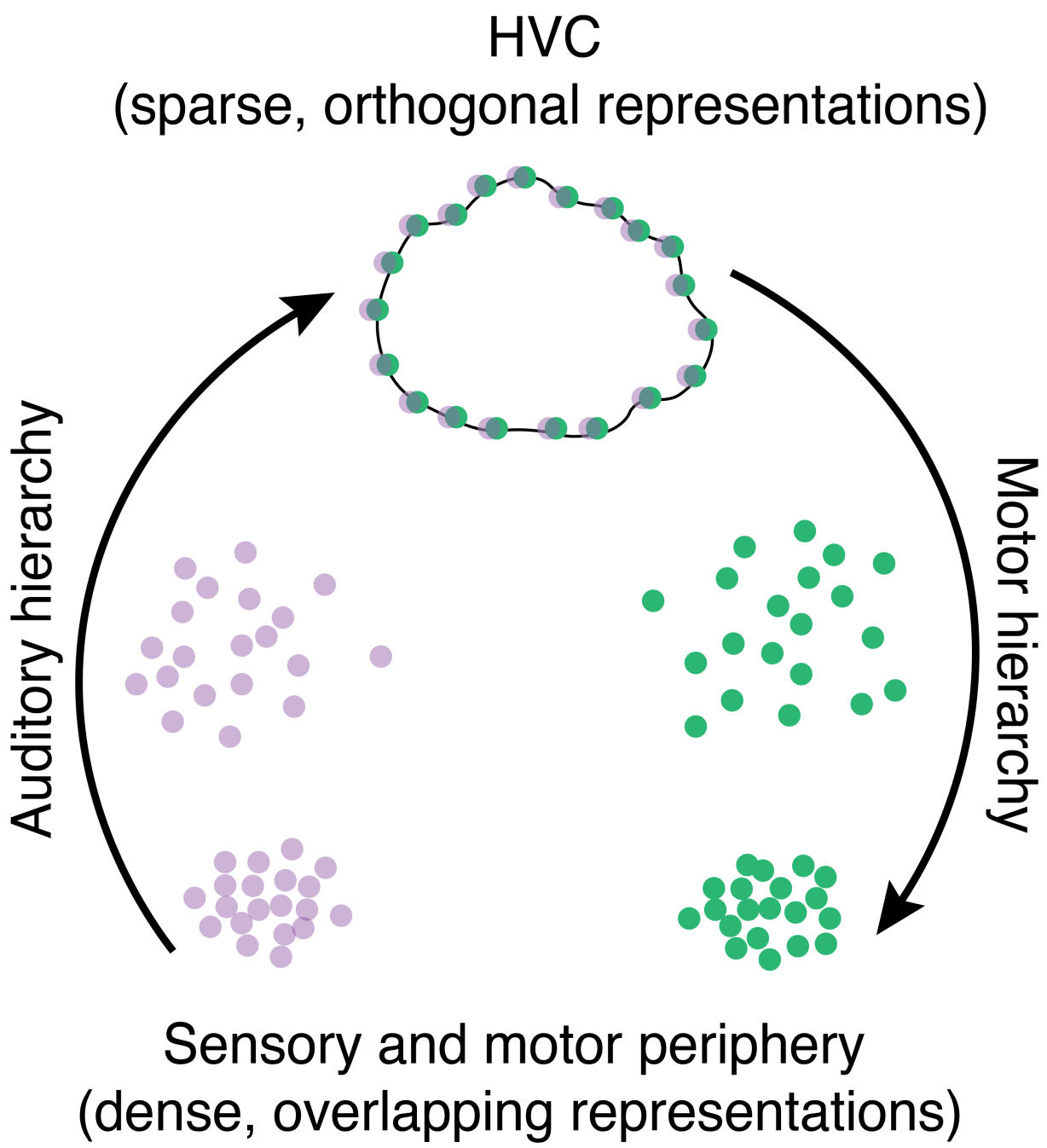
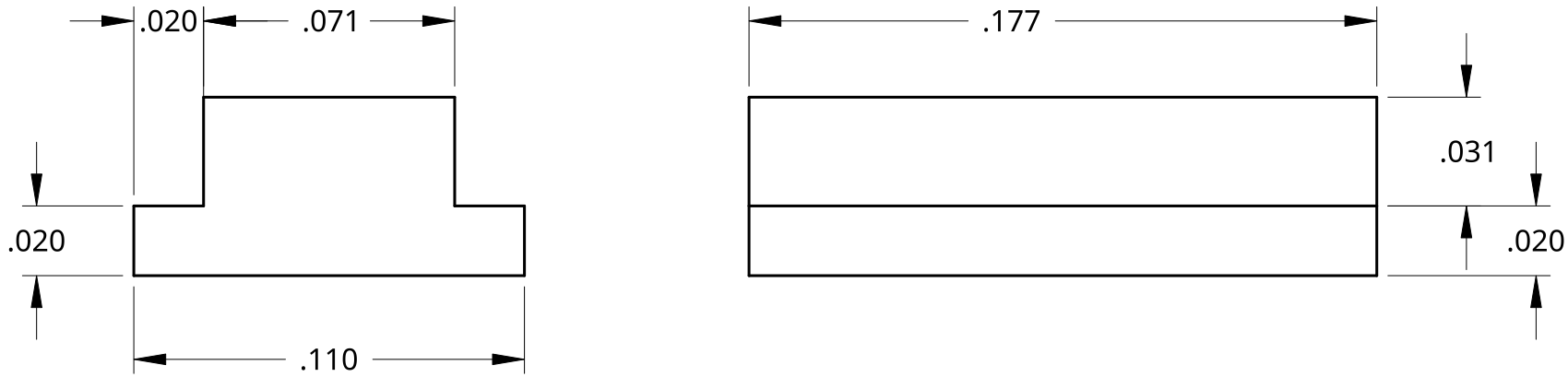
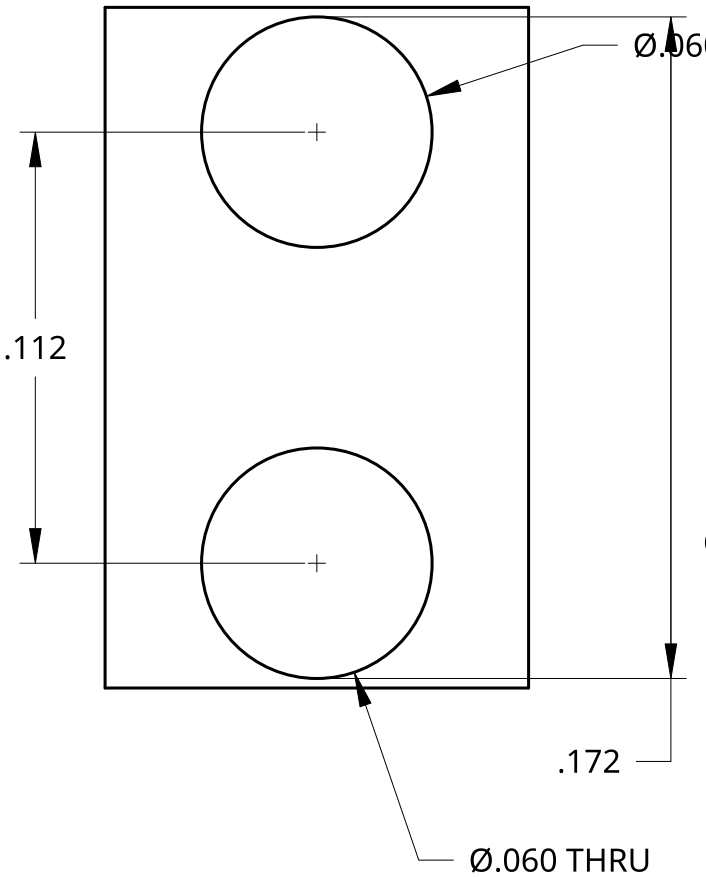
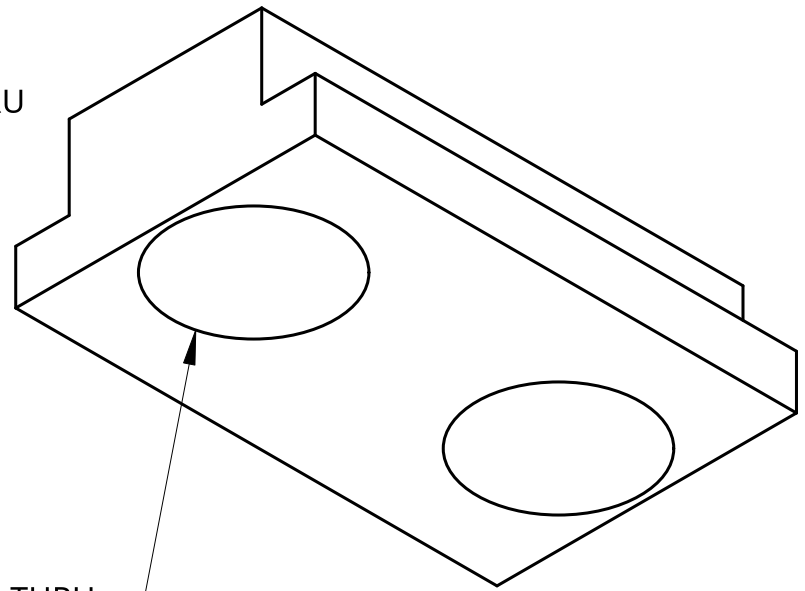


Figure 3.15: Conceptual sketch of HVC's function in auditory and motor hierarchies

A conceptual sketch of HVC's roles in both the auditory and motor hierarchies. In this framework the auditory system sparsens and decorrelates the representation for different sounds until it reaches HVC where each song produces a sparse sequence of activity (i.e. each time has its own unique and sparse neural state). For the motor system HVC plays the opposite role. HVC dynamics provide dynamics that uniquely map to each time in the developing imitation. These states can be arbitrarily mapped to a dense pattern of motor commands needed to control the motor periphery. Thus HVC representation links the language of sensory and motor systems, allowing for efficient memory formation and readout

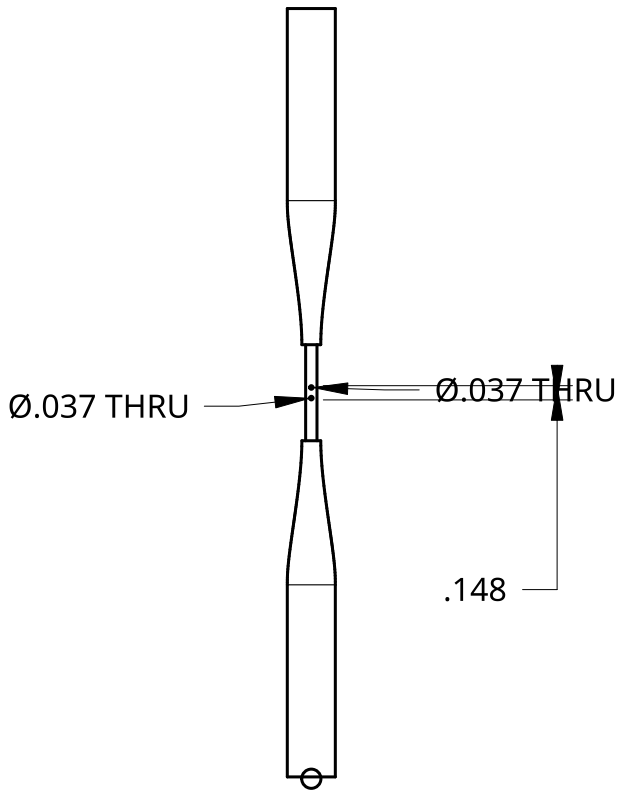




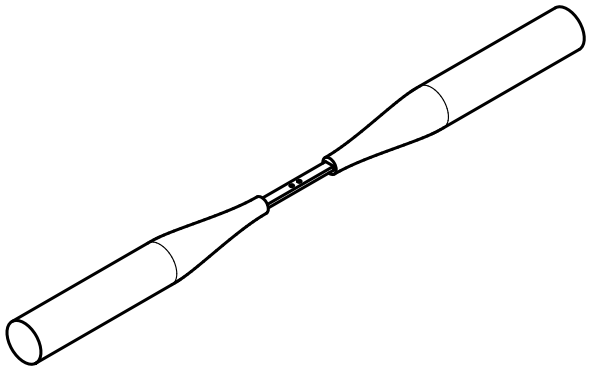


UNLESS OTHERWISE SPECIFIED, DIMENSIONS ARE IN INCHES .XX = ±.0- .XXX = ±.00- ANGULAR = ± ° .XXXX = ±.000- FRACTIONAL = ± SURFACE FINISH		NAME	DATE
	DRAWN	ANDREW BAHLE	11/21/2020
	CHECKED		
	APPROVED		
DO NOT SCALE DRAWING			
BREAK ALL SHARP EDGES AND REMOVE BURRS			
THIRD ANGLE PROJECTION	MATERIAL	FINISH	

TITLE		
SIZE A	DWG NO.	REV.
SCALE 20:1	WEIGHT	SHEET 1 of 1



SCALE 1:2



SCALE 1:2

SCALE 1:2

UNLESS OTHERWISE SPECIFIED, DIMENSIONS ARE IN INCHES .XX = ±.0- .XXX = ±.00- ANGULAR = ± ° .XXXX = ±.000- FRACTIONAL = ± SURFACE FINISH		NAME	DATE
	DRAWN	ANDREW BAHLE	11/28/2018
	CHECKED		
	APPROVED		
DO NOT SCALE DRAWING			
BREAK ALL SHARP EDGES AND REMOVE BURRS			
THIRD ANGLE PROJECTION	MATERIAL	FINISH	

TITLE		
SIZE A	DWG NO.	REV.
SCALE 1:5	WEIGHT	SHEET 1 of 1

Chapter 4

Tutor dependent changes in auditory responses in song motor regions

4.1 Abstract

Auditory experience in songbirds leads to long lasting neurophysiological changes throughout the auditory system [194, 195]. Experience with a particular "tutor" song, leads to dramatic changes in behavior, culminating in a vocal imitation of this song. Which changes, in which auditory brain regions underlie this capacity for vocal imitation? While tutor experience leads to changes in neural responses to the tutor song in many regions [76, 196, 197, 198, 199, 194, 195], genetic and pharmacological lesions indicate that song motor areas (including HVC and NIf) and motor associated auditory regions (Avalanche) are particularly important for the auditory memory underlying vocal imitation [22, 21, 20, 80, 81][1]. We therefore asked how the representation of the tutor song by HVC neurons is affected by tutor experience. Following up on our previous finding that tutoring evokes sparse sequences in HVC, we found that unfamiliar songs also evoke unique sparse sequences, both before and after tutoring. In HVC however, responses became more selective for the tutor song following tutoring, perhaps reflecting an auditory memory trace of this song. Since HVC receives diverse auditory inputs [173, 174, 26, 84], we next asked if these change in selectivity arise in auditory regions that provide synaptic input to HVC. By recording in NIf, L1, CLM and HVC_{shelf}, we found that HVC was more selective for the tutor song than these other regions. This suggests that tutor selectivity may emerge from local circuit reorganization within HVC itself. Interestingly, this difference in selectivity was smallest when comparing HVC and HVC_{shelf}, a regions out of

[1]The extensively studied area NCM does seem to be important for a tutor song memory used in recognition [83] suggesting that there are different tutor memories for different behaviors. For a fascinating discussion of the generalized vs functionally individuated memory systems and more see [200].

which, HVC has been proposed to have evolved [177].

4.2 Introduction

Auditory experience has profound effects on future behavior. In humans, early exposure to language is essential for adult perceptual abilities and language fluency [201, 202, 203], and auditory cortical areas display specializations for the representation of human vocalizations and native-language speech [204, 205, 206]. In rodents, the contents and statistics of auditory experience dictate the spatial pattern of auditory cortical receptive fields [207, 208, 209], and adult experience can dramatically reorganize this tuning [210, 211, 212]. It is clear from these cases, that auditory experience changes behavior, and alters the statistics of neural tuning, but how does auditory plasticity underlie the formation of an individual auditory memory, for example the instantaneously formed, and long-lasting memory of a catchy musical tune [108]?

In juvenile male songbirds, experience with a particular tutor song leads to the rapid formation of an auditory memory used for vocal imitation [2, 106, 105]. This auditory memory is likely stored by plasticity in auditory responsive regions, but which changes in which brain regions make up this memory? The auditory system of the songbird is similar to the layered cortical structures found in mammals [213, 214, 162], characterized by dense connectivity within and between primary and secondary regions [173, 215]. Tutoring leads to changes in neural tuning to auditory stimuli in many of these areas [76, 196, 197, 198, 199, 194, 195, 91, 25, 23], and inactivations or lesions of these different regions can lead to deficits in song learning [19, 82]. Recent work suggests that some auditory regions that show tutor dependent plasticity, are not required for song learning (L1 and NCM) [80, 81, 22], while song motor circuits are instead essential for accurate song copying (HVC, NIf and Av) [22, 21, 20]. Here we test the hypothesis that all zebra finch songs (and thus potential tutor songs) are represented in these song regions in young birds of tutoring age. We further

hypothesize that tutoring biases the representation of the tutor song in some or all of these regions, forming a long lasting memory that underlies song learning.

4.3 Methods

4.3.1 Animal care and use

We used male zebra finches (Taeniopygia guttata) from the MIT zebra finch breeding facility (Cambridge, MA). Animal care and experiments were carried out in accordance with the NIH guidelines and reviewed and approved by the Massachusetts Institute of Technology Committee on Animal Care. For all tutoring experiments we prevented exposure to a tutor song prior to experiments; birds were separated from their father on or before 15 days post-hatch and raised by their mother and a foster female. For experiments birds were singly housed in custom made sound isolation chambers.

4.3.2 Surgical procedures

Juvenile zebra finches (35-55 dph) were anaesthetized with 2% isoflurane gas before being placed in a custom stereotaxic apparatus. After applying a topical anaesthetic (0.1% bupivacaine) and performing a midline incision in the skin over the skull, we made craniotomies in the skull at a predetermined distance from the bifurcation of a major blood vessel (the lambda sinus). For head-fixed extra-cellular recordings, unilateral or bilateral craniotomies were made in the same manner as described above and sealed with silicone elsatomere (kwik-kast). Craniotomies were also made over areas used for antidromic verification of recording site identity (Area X was used for identification of HVC recording sites and HVC was used for identification of NIf recording sites). A small durotomy was performed and a custom bipolar stainless steel stimulating electrode [156] was implanted at stereotaxically defined coordinates and secured to the skull using dental acrylic. Finally a small (3.9 x 4.55 x 1.3

mm) stainless steel head-plate with two threaded holes (3.5.3) was affixed to the skull using dental acrylic. In both imaging and electrophysiology experiments, the incision site in the skin was closed with a tissue adhesive (VetBond). Analgesic (Buprenorphine or Meloxicam) was administered 30 minutes prior to surgery and for three additional days after the end of surgery (Meloxicam).

4.3.3 Extracellular recordings

Electrophysiological recordings were carried out in head-fixed birds using Neuropixels 1.0 probes (Phase 3A) [157] or 2.0 probes [158] and acquired using SpikeGLX software (https://billkarsh.github.io/SpikeGLX/). Spikes were sorted offline using Kilosort 2 (www.github.com/MouseLand/Kilosort2) after pre-processing and filtering with CatGT (https://billkarsh.github.io/SpikeGLX/). Clusters were manually curated with phy2 [159] to remove multiunit clusters, noise clusters and duplicates. Waveform characteristics were obtained using Cwaves (https://billkarsh.github.io/SpikeGLX/). Birds were head-fixed for the duration of recordings using a metal post (3.5.3) and swaddled in a foam restraint to prevent excessive movement. Antidromic stimulation was performed using a programmable pulse generator (Master-8, MicroProbes for Life Sciences) and a stimulus isolator (Iso-flex, MicroProbes for Life Sciences), with currents between 100 and 200 microamperes. Target regions were identified as those with electrodes exhibiting antidromic "hash". A unit was defined as within HVC or NIf if the electrode with the largest amplitude spike waveform fell within the antidromically identified region. Units were defined as in L1/CLM if they were at least 50 microns (to exclude units in the somewhat ambiguous boundary between these areas) above the upper antidromically defined boundary of NIf. We defined HVC_{shelf} as spanning the region beginning 200 microns below the bottom of HVC and ending 50 micron below the bottom of HVC as defined antidromically. This dorsal-ventral extent was conservatively chosen based on a previous study which measured the spread of local stimulation in HVC_{shelf} in slice [73]. In reality these regions are anatomically contiguous, however the 50 micron buffer zone was

meant to exclude units near the somewhat ambiguous border.

4.3.4 Analysis of eletrophysiological data

Putative excitatory and inhibitory unit classifications were obtained using k-means clustering of the average firing rate, spike waveform asymmetry and spike width. Waveform characteristics were calculated using the average waveform on the electrode with the largest amplitude. Spike width was defined as the duration (in milliseconds) between the peak and trough in the average waveform. The peak was defined as the point at which the extracellular waveform reached its maximum and trough was defined as the point at which the average waveform reached its minimum. Spike asymmetry was calculated as the relative heights of the maximal amplitudes on either side of the trough [160]. Neurons were classified as significantly modulated by the tutor song using the non-parametric Zeta test[161] at a level of $P<0.05$, with Bonferroni correction for multiple comparisons.

To measure the temporal sparseness of auditory evoked responses we adopted a previously used measure [216, 217]. Average firing rates were computed using 1ms bins smoothed with a 20ms moving average, and normalized to produces a probability density, p_i for each of N bins such that $\sum_{i=1}^{N} p_i = 1$. The sparseness index of a given neuron was then computed as:

$$S.I. = 1 + \frac{\sum_{i=1}^{N} p_i \log p_i}{N}$$

This measure is equal to one if the activity is restricted to a single bin and zero if it is equally distributed across all bins.

The inter-motif correlation coefficient (IMCC) was used to measure the reliability of neural firing patterns across all trials, as previously described [217, 59, 18, 218]. Motif aligned, mean subtracted, firing rates were computed for each trial (r_i), using 1ms bins smoothed with a 20ms moving average. The IMCC for a given neuron was then computed

as the mean pairwise correlation coefficient (CC) across all pairs of trial (n):

$$IMCC = \left\langle \sum_{j>i}^{n} CC_{ij} \right\rangle$$

where

$$CC_{ij} = \frac{r_i r_j}{\sqrt{r_i^2 r_j^2}}$$

Neural selectivity was quantified using two measures. A discriminability index (D') [26] and a song selectivity index (SI) [95]. The D' measure was calculated for each neuron as:

$$D' = \frac{2(\langle r_a \rangle - \langle r_b \rangle)}{\sqrt{\sigma_a^2 + \sigma_b^2}}$$

where $\langle r_x \rangle$ is the mean firing rate during stimulus x (in this case a or b) and σ_x^2 is the corresponding variance. Selectivity was calculated for each neuron as:

$$SI = \frac{\langle r_a \rangle - \langle r_b \rangle}{\langle r_a \rangle + \langle r_b \rangle}$$

where $\langle r_x \rangle$ is the average activity for a given neuron during stimulus x. This measure ranges between 1, for a neuron perfectly selective for song a and -1, for a neuron perfectly selective for song b.

4.3.5 Tutoring protocol

We isolated birds from their father at or before 15 days post hatch (dph) based on prior studies that achieved effective imitation in experimental animals isolated between 10 and 30 dph [1, 88, 119]. These isolated birds remained co-housed with their siblings and mother until at least 35 dph, at which point they were singly housed for the duration of the experiment [102]. Zebra finches can produce accurate imitations after as little as 75s of tutor exposure [114], but protracted daily tutor exposure impairs imitations [120]. We therefore followed an

intermediate tutoring protocol in which pupils received 60 minutes of tutor exposure each day [102]. Birds were housed in a sound attenuating chamber and their songs were recorded daily using either Sound Analysis Pro software (SAP)[121] or custom recording software. Birds were tutored by a live tutor for 1 hour per day for 5 consecutive days. Neural recordings were made on the seventh day after tutoring to allow a night of sleep following tutoring and acclimatization to the head-fixation apparatus.

4.3.6 Auditory stimuli

Auditory stimuli used for playback were constructed using motifs recorded from birds in our colony, and included the tutor bird used in these experiments. Songs were presented as bouts of five motifs preceded by intro notes. For HVC and HVC_{shelf} recordings, 40 bouts of each song were presented for a total of 200 motifs per song. For auditory cortical recordings (NIf and L1/CLM), 30 bouts of the tutor song were presented, 20 bouts of control song 1 were presented, 10 bouts of control song 2 were presented and 5 bouts of control song 3 were presented for a total of 150,100,50 and 25 motifs respectively. Bouts from each song type were randomly presented and presentations were flanked by silent periods with random durations drawn from a uniform distribution between 3 and 13 seconds. During neural recordings, tutor song recordings were played-back over a speaker (8010A, Genelec) placed ∼25cm from the birds head. Playback was acquired via a microphone placed next to the head of the bird, and the sound amplitude of playback was calibrated to match a live tutor singing directed song at the same location as the speaker (SPL of ∼75 dB at 25cm).

4.4 Results

4.4.1 Song evokes sparse sequences in HVC in naive and tutored birds

We first wondered whether unfamiliar songs also evoke sparse sequential responses in HVC similar to those we observed during tutoring (figure 3.1 and 3.7). Since HVC sequences for the tutor song were present as early as the very first day of tutoring (figures 3.11 and 3.13), we reasoned that *any* zebra finch song should elicit a sparse and unique neural sequence. To test this hypothesis, we used Neuropixels probes to record HVC neurons in head-fixed birds, during presentation of multiple songs. Tutored birds were presented four song stimuli. These stimuli included three songs chosen at random from our colony, as well as their tutor's song. Naive birds were presented with the same four stimuli as tutored birds, but had no prior experience with any song. We observed that HVC neurons were organized into sparse sequences during presentation of each song, in both naive and tutored birds (figure 4.1). Many song responsive neurons were significantly modulated by a single song type (44.4% in naive birds, 47.1% in tutored birds), while a relatively small fraction were significantly modulated by all four songs (12.8% in naive birds, 12.4% in tutored birds). We observed that the ordering of neural responses was unique for each presented song. This was apparent when visualizing the evoked activity for each song, sorting neurons either by the latency of their peak response during the same song, or by the latency of their peak response in other songs. Sequential organization was only observed when responses to a given song were sorted according to activity in that song condition, not other song types (figure 4.1).

4.4.2 Tutoring does not change the fraction of neurons selective to the tutor song

We next wondered if tutor exposure led to an increase in the fraction of neurons modulated by the tutor song. We found that this was not the case. The proportion of cells significantly modulated by the tutor song, as determined by the Zeta test [161], did not differ between naive and tutored birds (mean fraction modulated by tutor song: naive birds = 15.9%, n = 2; tutored birds = 17.2%, n = 8; two sample t-test, t(8)=-.1716, p=0.87). In addition, the mean fraction of neurons modulated by each of the four songs was similar across birds, for both naive (n=2 birds; two way ANOVA F(3,1) = 0.79, p = 0.58) and tutored birds (n=8 birds; two way ANOVA F(3,7) = 1.58, p = 0.22). Together this suggests that HVC neurons are recruited in equal proportions by playback of any zebra finch song, and that tutoring does not significantly alter the fraction of neurons responsive to the tutor song.

4.4.3 Changes in reliability and sparseness of neural responses after tutoring

Although tutoring does not alter the fraction of tutor responsive HVC neurons, it may lead to an increase in the trial to trial reliability of neural responses, or the temporal sparseness of tutor evoked firing rates. To quantify these properties, we computed two measures of the firing response profiles; the inter-motif-correlation-coefficient (IMCC) [217, 59, 18, 218], which measures the average pairwise correlations for a neuron across all trials, and an entropy based measure of temporal spareness [216, 217]. We did not find evidence for changes in these measures that were specific to the tutor song. Comparing the cumulative distributions of IMCC between tutored and untutored birds, we found that the IMCC increased following tutoring for both the tutor song and one of the control songs song (tutor song; $p<0.01$, song 1; $p<0.05$, one tailed two way Kolmogorov–Smirnov test). Temporal sparseness increased non-specifically for all songs following tutoring, suggesting this may be the

product of physiological changes unrelated to the particular tutor song (tutor song; $p<0.05$, song 1-3; $p<0.0001$, one tailed two way Kolmogorov–Smirnov test).

4.4.4 Average firing rates are shifted towards the tutor song following tutoring

As described in the previous sections, tutoring does not lead to a greater fraction of tutor modulated neurons, nor an increase in reliability or temporal sparseness of the response to the tutor song over unfamiliar songs. What if anything, does change in HVC as a result of experience with the tutor song? By comparing the mean firing rate in HVC neurons for the tutor song and control songs, we found that in tutored birds, HVC neurons exhibit higher firing rates to the tutor song relative to other song (mean difference in firing rates between tutored and untutored birds 1.14hz, n=8 birds; two way ANOVA $F(3,401) = 33.03$, $p < 0.0001$). This effect was present in both excitatory and inhibitory neurons, as previously reported [91] and was absent in naive birds (figure 4.5 a). This suggests that preferential firing to the tutor song is acquired via auditory experience, and is not due to an incidental/random firing rate preference for the particular song used for tutoring.

4.4.5 Tutoring increases the tutor selectivity of the HVC population

We next sought to directly quantify the neural tuning preference of populations of HVC neurons, both before and after tutoring. For each neuron, we computed the selectivity for the tutor song relative to each control song. We employed two commonly used measures of firing rate selectivity: a neural discriminability index (D') and a song selectivity index (SI). We calculated D' and SI for the tutor song relative to each unfamiliar song, for all HVC neurons. In tutored birds, the distributions of both D' and SI scores were shifted toward greater selectivity to the tutor song, relative to distributions in naive birds (figure 4.7). This

shift was present at the level of the entire population (figure 4.7 a, D': $p<0.01$, in each case. SI: $p<0.0001$ in each case, one tailed two way Kolmogorov–Smirnov test) and inhibitory cells (figure 4.7 c, D': $p<0.01$ in each case. SI: $p<0.0001$ in each case, one tailed two way Kolmogorov–Smirnov test) as well as at the level of excitatory neurons, although here the shift was not statistically significant (figure 4.7 b). In light of this shift in selectivity following tutoring, we wondered what the qualitative firing patterns of the neurons most selective for the tutor song looked like. Figure 4.9 shows three units in the 95th percentile of tutor song selectivity (here calculated as the average selectivity across all control songs), for naive birds (figure 4.9 a) and tutored birds (figure 4.9 b). In untutored birds, even the cells most selective for the tutor song, tended to fire unreliably, and at similar rates for each song. In tutored birds on the other hand, the most selective cells tended to fire at much higher rates during presentation of the tutor song, relative to control songs. These selective neurons tended to fire briefly during the tutor song and at high rates.

4.4.6 Selectivity to the tutor song in neurons with temporally modulated firing rates

In the previous chapter, we highlighted the sparse sequential activation observed during tutoring. We wondered if the sub-population of neurons with these temporally modulated, tutor-evoked firing rates, themselves exhibited selectivity to the tutor song. Alternatively, tutor selectivity at the population level could arise via a temporally non-specific increase in the overall firing rate during the tutor song for all neurons. To test this we measured the cumulative distributions of the D' and SI metrics in the subset of HVC neurons with firing rates significantly temporally modulated by the tutor song. We found that neurons that were significantly modulated by the tutor song exhibited even more dramatic increases in their selectivity after tutoring than the overall HVC population (figure 4.11). When restricting our analysis to only neurons significantly modulated by the tutor song, we found that, as a population these neurons were significantly more selective after tutoring (figure 4.11 a,

D': $p<0.01$, in each case. SI: $p<0.01$ in each case, one tailed two way Kolmogorov–Smirnov test), as well all as the subset inhibitory neurons (figure 4.11 c, D': $p<0.05$, in each case. SI: $p<0.05$ in each case, one tailed two way Kolmogorov–Smirnov test). In this case, the subset of excitatory neurons was significantly more selective for the tutor song in two of the three control songs (figure 4.11 b, D': Tutor vs song 2 $p<0.01$, Tutor vs song 3 $p<0.05$; SI: Tutor vs song 2 $p<0.05$, Tutor vs song 3 $p<0.05$, one tailed two way Kolmogorov–Smirnov test). This suggests that the increase in tutor selectivity, is primarily due to changes in neurons with temporally modulated firing rates and is not due to a coarse overall increase in the firing rates of neurons during tutor song presentation.

4.4.7 Selectivity to the tutor song first emerges in HVC

Changes in HVC tutor response selectivity may be due to synaptic plasticity within the local HVC network. Alternatively, this selectivity may be inherited from one or more of the auditory areas providing input to HVC. To distinguish between these two possibilities, we made recordings from several structures known to provide auditory input to HVC: NIf, L1, CLM, and HVC-shelf. We measured the selectivity of neural responses in each of these regions, relative to selectivity in HVC. This was done in tutored birds, with matched amounts of tutor song exposure. The selectivity of responses, as measured with both D' and SI, were higher in HVC relative to both NIf, L1/CLM and to a lesser extent, HVC_{shelf} (figure 4.13). This supports the idea that tutor selective representations, perhaps underlying a tutor memory, emerge specifically in HVC. This could occur as a result of the plasticity and synaptic reorganization previously observed in HVC following tutoring [149, 22, 102]. This result stands in contrast to BOS-selectivity in adult birds, which is present not only in HVC, but two of its inputs, NIf and the Avalanche subregion of CLM [98, 219, 29, 84].

4.5 Discussion

4.5.1 Sequences for any song

In chapter 3, we observed that tutor evoked sequences could be present from the very first day of tutoring, and could be stable over an entire week. Here we showed that sparse sequential responses in HVC are evoked by playback of all songs tested. This was true in both naive and in tutored birds, perhaps suggesting that *any* zebra finch songs would also produce similar sparse sequences. This observation is consistent with the tutor template formation mechanism, discussed in chapter 3. We hypothesize that each possible tutor song can produce a unique sparse sequence of activity in HVC. We further hypothesize that during tutoring, this sequence is rapidly mapped, from an input driven regime, to one driven by internal HVC dynamics. In this way, initially input driven auditory sequences become a dynamic neural clock, reflecting an auditory memory of song timing. A possible consequence of this model might be the presence of sequential neural replay. If a tutor evoked sequence is stored in the internal dynamics of HVC, we might expect to see replay of this sequence offline, during quiet wakefulness or sleep. Even if these dynamics are initially too weak to support coherent replay of the entire sequence, evidence of this plasticity may be observed in the form of neural correlations between neurons co-active, during tutoring. This prediction should be tested in future work.

4.5.2 What does the emergence of tutor selectivity tell us about the mechanisms of template formation?

Despite a clear role for HVC in the formation of the auditory tutor memory [149, 22, 21, 101, 20, 89], tutor song selectivity during the earliest days of tutoring has been studied relatively infrequently. Tutor selective neural responses have been reported in multi-unit recordings [25], and in a very small number of single units, recorded in HVC [91, 220, 221]. Here we

built upon this prior work, using contemporary methods of high-density electrophysiology, to record large populations of HVC neurons. We replicated the previous observation of tutor selective responses in HVC [25, 91], and were able to quantify the prevalence of such selectivity before and after tutoring in HVC as well as its auditory inputs. What does the presence of tutor selective responses mean? We hypothesize that tutor selective responses are a signature of learned dynamics in HVC that underlie the song-template, and guide subsequent imitation learning. Selectivity could reflect Hebbian plasticity, between tutor responsive neurons. This is the same property we hypothesized could account for the effect of cooling on imitation in chapter 2, and for the increase in stability of the tutor response over a week of tutoring (chapter 3). Here however, we showed that this change was specific to the tutor song. This was a critical prediction taken for granted in these earlier chapters. Future work could test if selectivity truly reflects a memory of the temporal structure of the tutor song by using playback of artificially stretched or compressed songs. We hypothesize that the specificity of the timing selectivity in HVC sets a bound on the maximum possible temporal accuracy of the imitation.

4.5.3 Inhibition as a gate for the sensitive period

We found that HVC interneurons developed a particularly clear preference for the tutor song over other songs. This fits well with previous findings that tutor song evoked inhibition grows proportional to song imitation [115]. Here however, we observed an increase in inhibitory activity even before birds had a chance to produce a vocal imitation. Previous work suggested that increases in inhibition in HVC, could act as a gate, protecting learned changes in song and finally closing the sensitive period [115]. This hypothesis agrees with theories of critical period plasticity developed in the domain of ocular dominance [222]. Upon monocular deprivation, inhibition in binocular visual areas drops dramatically, before slowly increasing to compensate for increased excitability in excitatory neurons with a contra-lateral preference [223]. Once inhibition reaches normal levels, the window for plasticity is effectively closed,

and the pattern of inhibition in the network maintains the visual tuning of neurons. This prevents compensatory plasticity from restoring binocular tuning once the deprived eye is uncovered [222]. We hypothesize that a parallel situation arises in during tutoring in HVC. At the onset of tutoring, inhibition is effectively low, and excitatory neurons are initially unselective to auditory stimuli. During tutoring, there is a large increase in inhibitory synapaptic contacts throughout the network [102], and a stabilization of excitatory synapses [88]. In addition there is an increase in tutor evoked firing rates in inhibitory neurons, as observed in this chapter and in previous work [89]. Once inhibition in the network matures in this manner, it serves to fix the network connectivity in place, preventing subsequent auditory plasticity. In other systems, similar "protective" inhibition can be manipulated by dissolving peri-neuronal nets (PNNs) surrounding PV interneurons. This manipulation has strong effects on plasticity. It can reopen critical period plasticity [224], and causes adult fear memories to be erased rather than extinguished by behavioral habituation [225]. PNNs are also present on PV interneurons in HVC and other song nuclei [226], but attempts to dissolve PNNs in adults produces almost no effect on song [227]. Future avenues for studying critical period closure and the role of PNNs should focus on studying these processes in juveniles during tutor memory formation.

4.5.4 Where and how does tutor selectivity emerge

We observed that tutor song selectivity was present in HVC to a greater extent than in auditory regions providing input to HVC (figure 4.13). This is in contrast to BOS selectivity in adults which is present in both NIf [29], as well as the Avalanche sub-region of CLM [84]. We observed the largest difference in tutor selectivity between HVC and NIf, the area which provides the largest source of auditory input to HVC [174, 20, 26]. Selectivity differences were more modest for CLM/L1 and HVC_{shelf}, which are far smaller sources of input to HVC [173, 228, 84, 26]. Our observations are also consistent with the requirement for synaptic plasticity during tutoring [22, 101] and the observed changes in synaptic connectivity in HVC

following tutoring [149, 22, 102]. Selectivity could emerge in HVC due to strengthening of excitatory synapses, pruning of inactive synapses, and through the presence of tutor selective inhibition throughout the network (as in previous models computational models of HVC [55, 56]). This interplay could additionally lead to the winner-take-all mechanisms proposed to underlie HVC sequential dynamics [229]. Because tutor selectivity is easily observed in head-fixed juveniles, it may be possible to directly study the synaptic mechanisms governing the real-time induction tutor selectivity. For example, by two-photon imaging of dendritic spines during tutoring, using calcium or glutamate reporters like GCaMP or iGluSnFR [230, 231, 232, 233]. This is a rich area for future study given the proliferation of tools for spine imaging and the emerging picture of tutor plasticity in HVC involving a rich interplay of dopaminergic signaling [101], synaptic pruning [149], and multiple neural subtypes. Together this could allow for experimental tests of several competing theoretical models of HVC sequence dynamics, by exploring where and how plasticity occurs during tutoring [229, 54, 57, 55, 56, 234, 235, 236, 237]

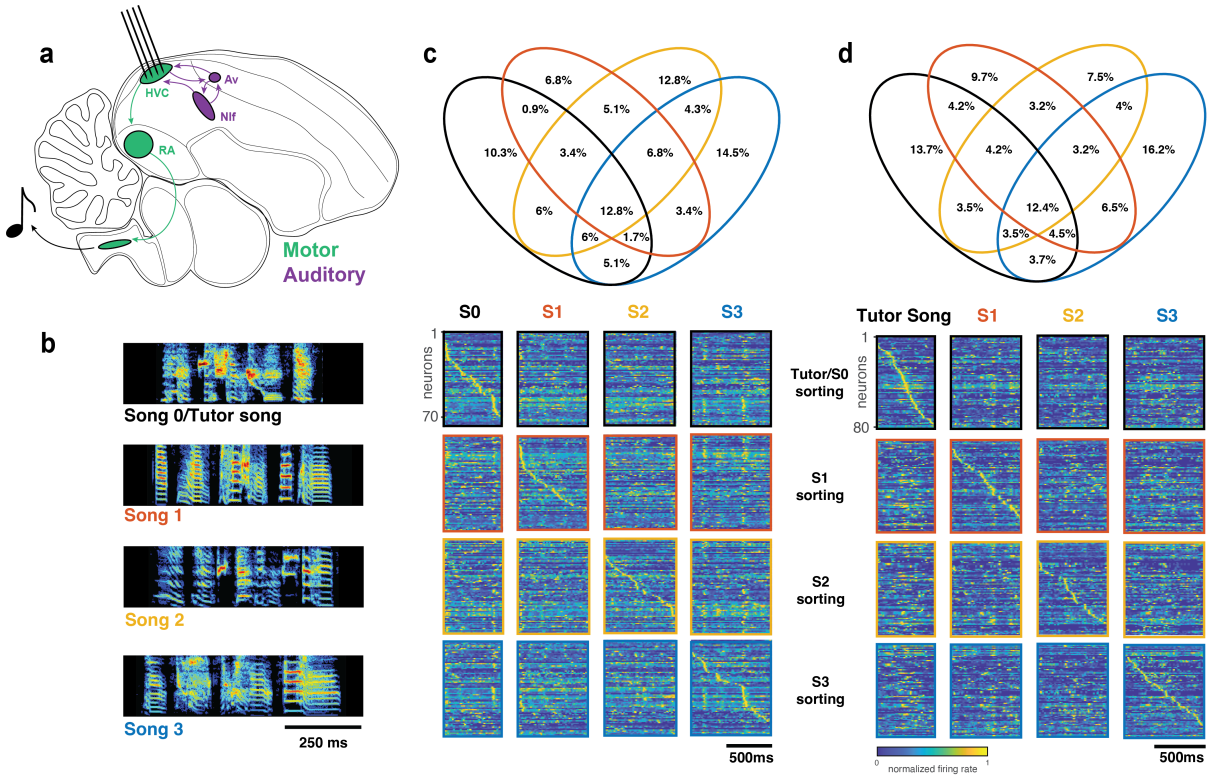


Figure 4.1: Sparse sequences are present for all songs both before and after tutoring

(a) A schematic showing the location of recording electrodes in HVC and its relation to song auditory and motor regions

(b) All four song stimuli used in this chapter. The first song was used as a tutor song for the tutored birds but was unfamiliar song for naive birds and is thus referred to as Song 0.

(c) top: venn diagram showing the proportion of neurons modulated by each song in naive birds (n=2 birds). Colors are the same as in B. Overlapping regions indicates proportions of neurons that were modulated by multiple song types. bottom: motif-aligned average firing rates of all song modulated neurons in an example naive bird. Each column shows the the activity during a different song and each row shows the same neurons sorted by the peak latency during a different song. For example the first column shows the activity during song 0 and the first row shows neurons sorted by the peak latency in this song.

(d) Same as in C but for tutored birds (n=8 birds). top venn diagram showing proportion of neurons modulated by each song and the overlap between these groups. Bottom: same as in C but for an example tutored bird.

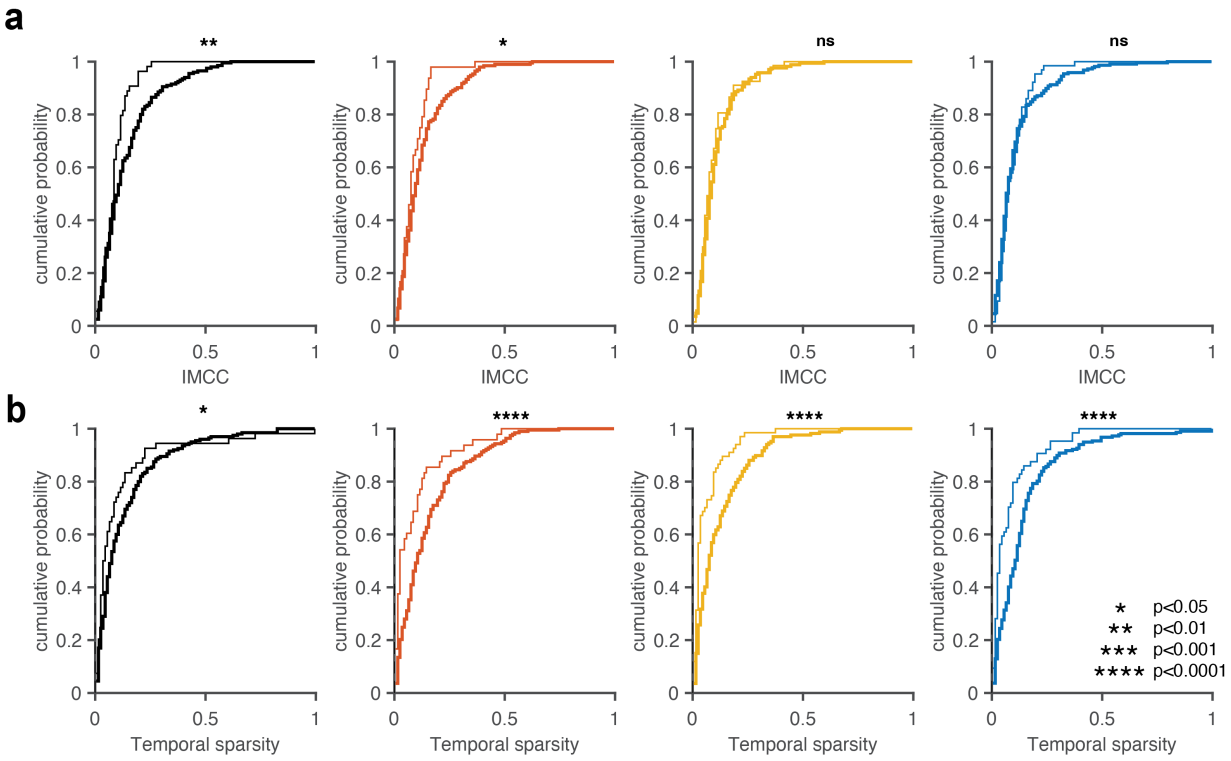


Figure 4.3: Reliability and temporal sparsity of HVC auditory responses

(a) Cumulative distribution of inter-motif correlation coefficients for HVC neurons in naive birds (thin lines) and tutored birds (thick lines) during playback of four different songs. Coloring is the same as in figure 4.1. Black plots show tutor song. Red, yellow and blue plots show data for control songs 1-3

(b) Cumulative distribution of temporal sparsity of the motif aligned average firing rates for HVC neurons in naive and tutored birds as in panel a.

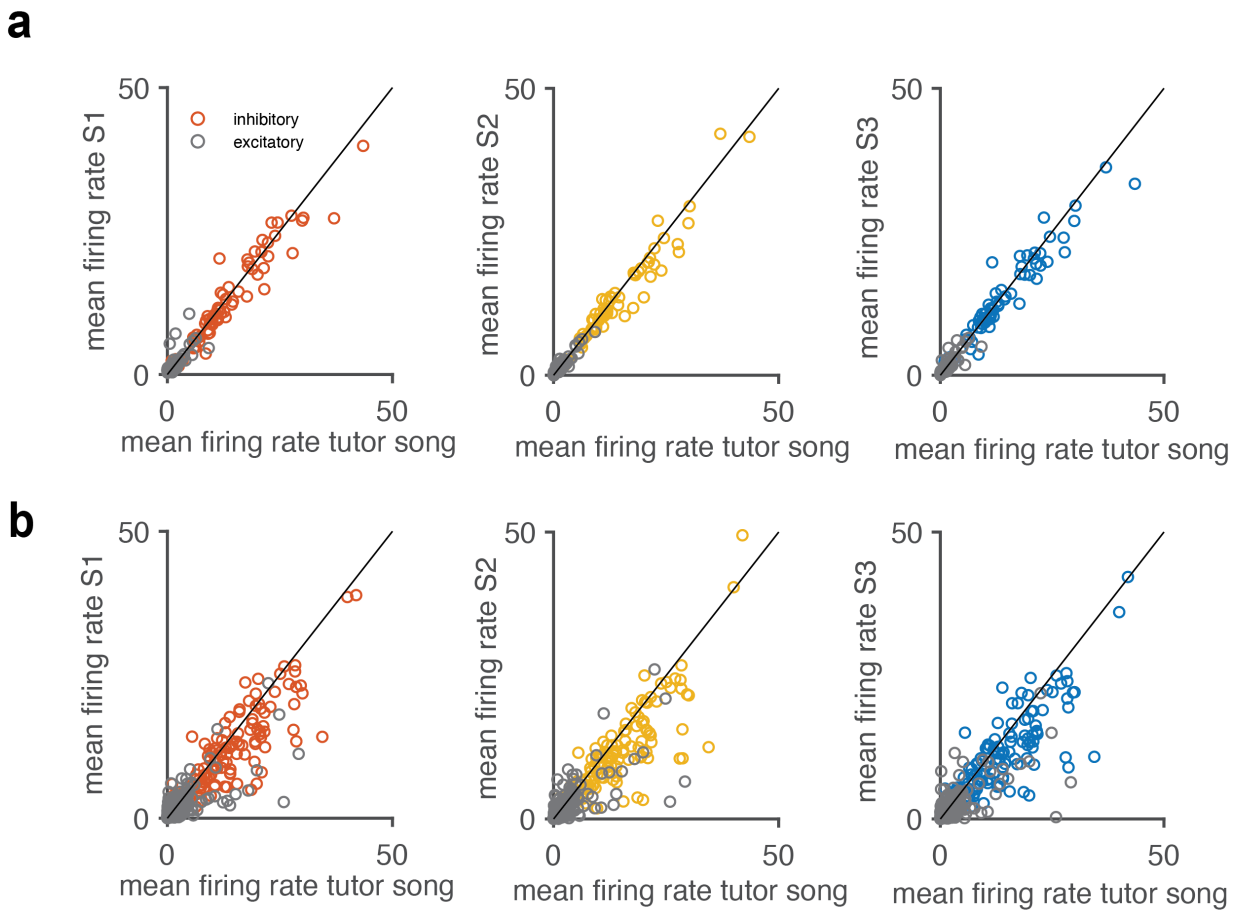


Figure 4.5: A comparison of the average firing rate of HVC neurons during playback of different songs

(a) Mean firing rates during playback of the tutor song or a control song in naive birds. Putative excitatory neurons are shown in grey and putative inhibitory neurons are shown in color. From left to right, plots show comparisons of the tutor song and control song 1-3.

(b) Same as in panel a but for all neurons recorded in tutored birds.

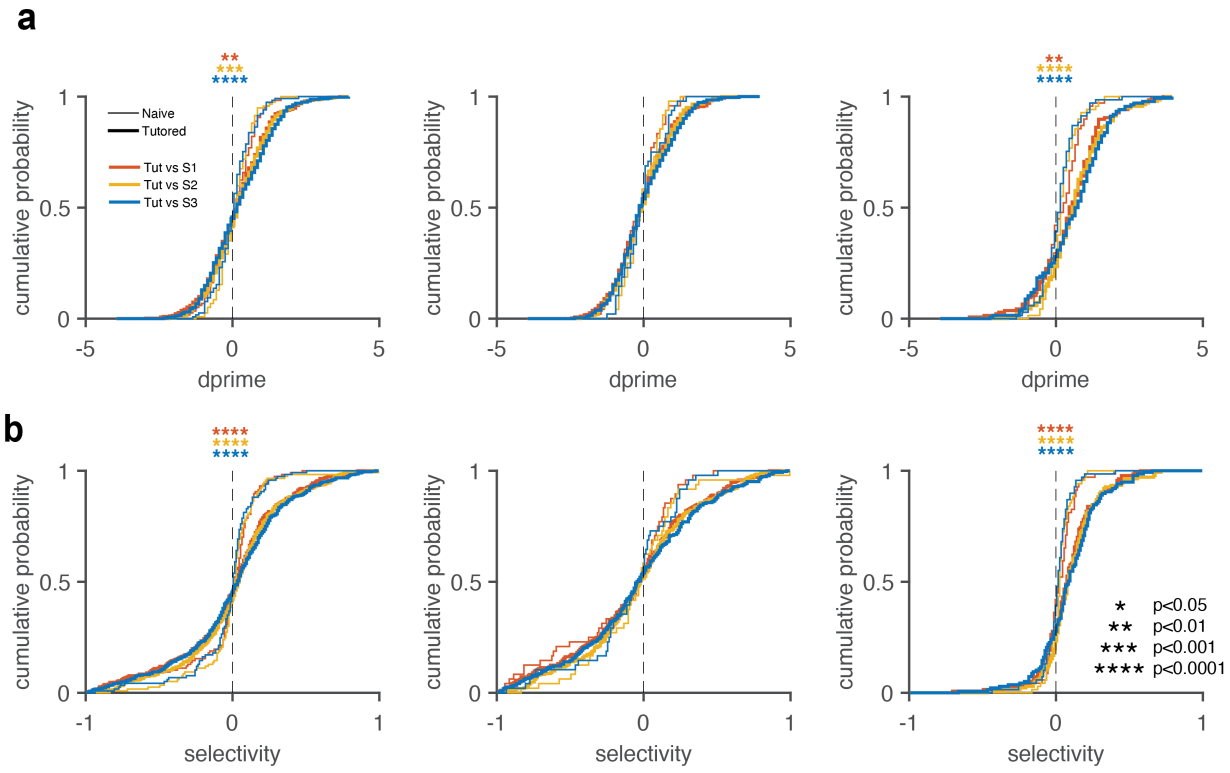


Figure 4.7: Tutoring induces and increase in the selectivity of neural responses to the tutor song in HVC

(a) Top row: Cumulative d-prime distribution for HVC neurons recorded in naive birds (thin lines) or tutored birds (thick lines) that were significantly modulated by any song. Different color indicate which control song the tutor activity is being compared to. Shifts to the right indicate an increase in the selectivity of neurons for the tutor song relative to control songs. Individual panels show all neurons (left), just excitatory neurons (middle), just inhibitory neurons (right).

(b) Bottom row: same as above but for cumulative distribution of each neuron's selectivity index

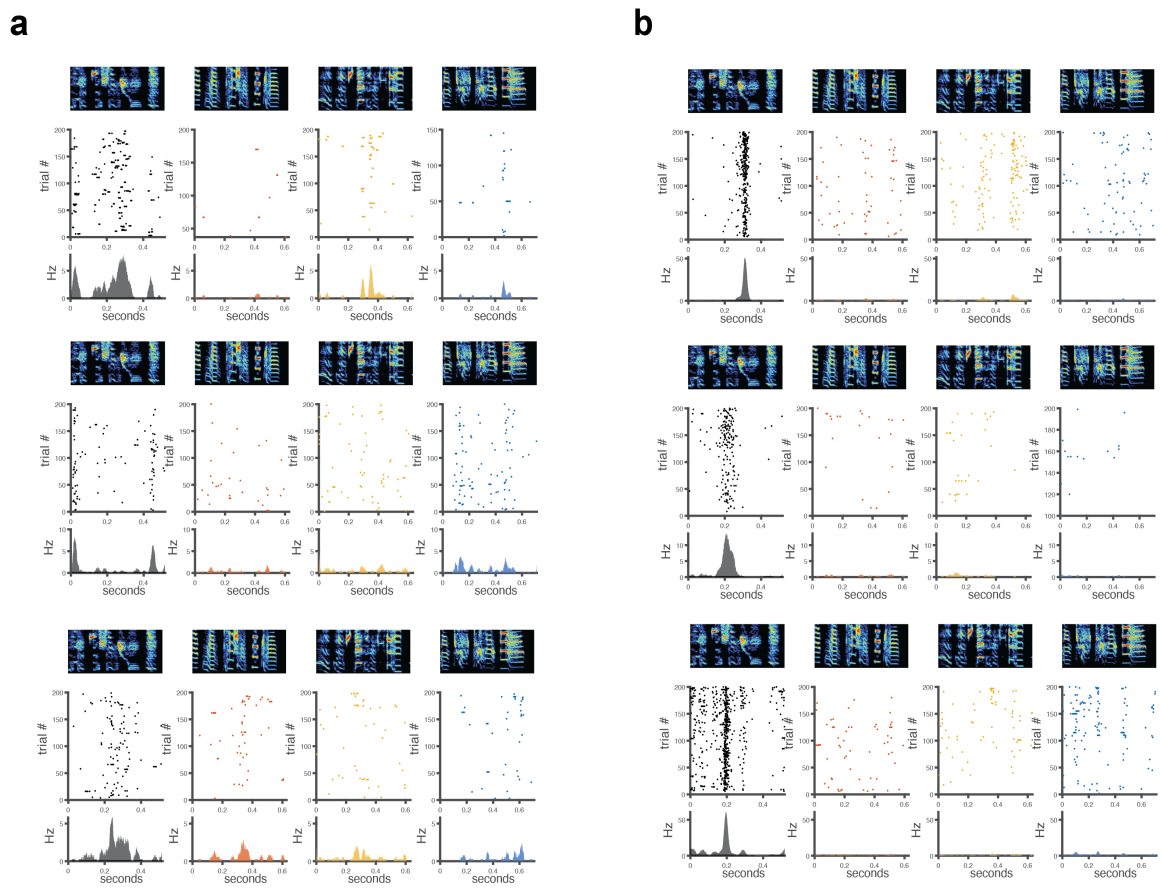


Figure 4.9: example selective units

(a) Three example units in the top 5% of neurons most selective to the tutor song (S0) in a naive bird. For each neurons the left most panel shows responses to the tutor song and each other panel shows responses to control songs. Within each panel a spectrogram of the song is shown (top), a raster of motif aligned neural spiking is shown (middle) and the average motif aligned activity is shown (bottom)

(b) The same as in panel a but from an example tutored bird

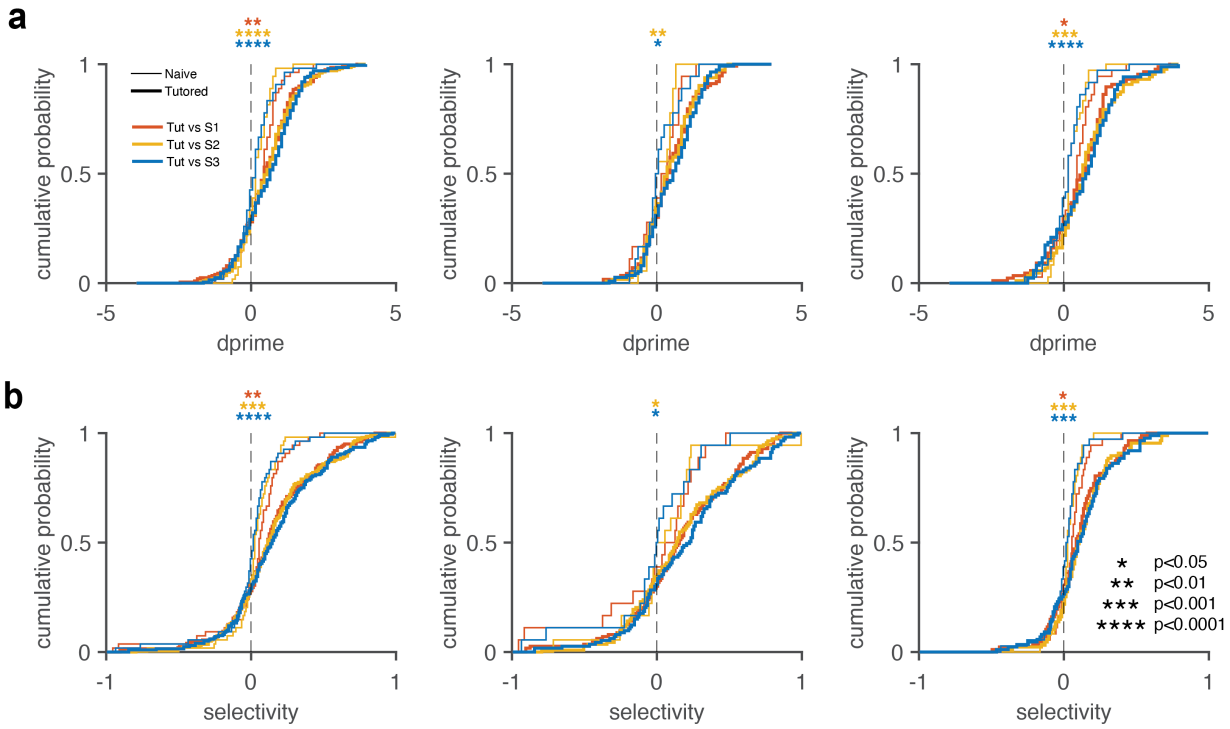


Figure 4.11: selectivity in neurons significantly modulated by the tutor song

(a) Top row: Cumulative d-prime distribution for HVC neurons recorded in naive birds (thin lines) or tutored birds (thick lines) that were significantly modulated by the tutor song. Different color indicate which control song the tutor activity is being compared to. Shifts to the right indicate an increase in the selectivity of neurons for the tutor song relative to control songs. Individual panels show all neurons (left), just excitatory neurons (middle), just inhibitory neurons (right).

(b) Bottom row: same as above but for cumulative distribution of each neuron's selectivity index

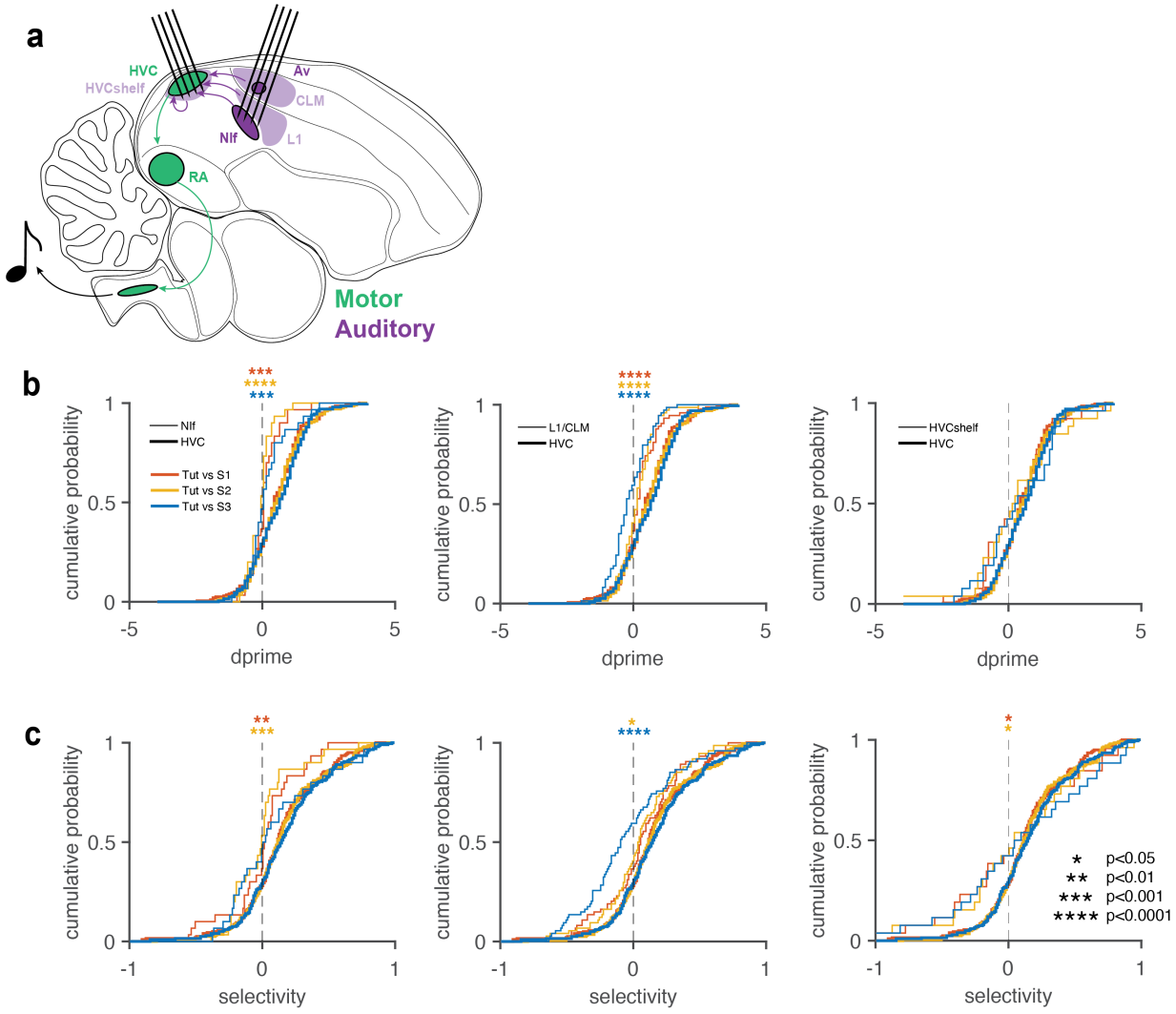


Figure 4.13: stability

(a) Top row: Cumulative d-prime distribution for HVC neurons recorded in tutored birds that were significantly modulated by the tutor song. Different color indicate which control song the tutor activity is being compared to. Shifts to the right indicate an increase in the selectivity of neurons for the tutor song relative to control songs. Individual panels show neurons in HVC (thick lines) compared with NIf (thin lines) neurons (left), L1/CLM (thin lines) neurons (middle), or HVC_{shelf} (thin lines) neurons (right).

(b) Bottom row: same as above but for cumulative distribution of each neuron's selectivity index

Chapter 5

A mechanistic hypothesis for HVC's role in the formation and recall of the tutor memory

We proposed in chapter 2, that the effect of cooling on imitation speed is evidence that HVC acts as a clock for tutor memory storage and recall. What does this entail? In the following section I outlining a complete hypothesis for tutor memory storage at the circuit level. The general idea is that HVC neurons form a temporal basis which acts as an index, calling up the content of the tutor memory at the correct speed and in the correct sequential order. We hypothesize that this recall occurs when HVC sequences that form during tutoring, are replayed during vocal practice. This replay reactivates a tutor memory in auditory cortex via HVC's projection to Avalanche [21], enabling song evaluation, and providing an error signal to reinforcement learning circuitry in the basal ganglia [12, 15]. Our model builds on previous proposals which developed the idea that HVC motor sequences might transmit a representation of the tutor memory to auditory cortex for evaluation [87, 88, 90, 86].

5.0.1 A temporal basis for the tutor song memory

We hypothesize that HVC acts as a temporal basis for the formation of the tutor memory. Mathematically, a basis refers to a set of functions that can be used in linear combination to approximate or represent a more complex function [238], in this case one defined in time. Particularly useful temporal basis functions are those that are orthogonal (or ortho-normal). For example, a complete set of Kronecker delta functions, defined at every time, can exactly reproduce any discrete function of time by linear combination. Our hypothesis that HVC acts as a temporal basis is in effect, a hypothesis that HVC neurons represent the tutor song using a near orthogonal basis set.

5.0.2 Stage I: Tutoring evokes sparse sequential activity in HVC

If HVC neurons do form a temporal basis for the tutor song memory, the simplest basis would be one in which each HVC neuron becomes active at one unique time during tutoring, similar to HVC activity during singing. Our observations in chapter 3 are consistent with this idea; excitatory neurons are active throughout the tutor motif, forming a sparse sequence (5.1 a). However, for a sequential basis set to be useful, for example in recalling the tutor memory, these sequences would need to be stored in a manner that would enable them to be recalled later during practice. Our observation in chapter 2 that cooling of HVC during tutoring leads to a sped up memory strongly suggest that this is the case.

5.0.3 Stage II: Hebbian plasticity during tutoring forms sequence ensembles in HVC

How could the sparse neural sequences observed during tutoring be stored and later recalled? We hypothesize that repeated exposure to the tutor song leads to Hebbian plasticity between neighboring neurons in the tutor evoked sequence, resulting in storage of the entire sequence in the weights of HVC (5.1 b). Multiple lines of evidence support this possibility. As

previously described, NMDA [149, 22] and dopamine [101] mediated plasticity in HVC during tutoring is required for initiation. In addition, both excitatory and inhibitory synapses in HVC exhibit rapid changes in density and stability following a single day of tutoring [149, 102]. Our observation in chapter 4 that tutor song selectivity increases in HVC after tutoring is also consistent with this idea. Finally the observation of sleep replay in the song system, and its potential relation to song learning, provides indirect evidence for Hebbian plasticity during tutoring [104, 239, 240, 241, 242]. HVC sleep replay is driven by NIf, its primary auditory input [243, 244], and replay frequency increases following the onset of tutoring [103]. This strongly suggests an association between tutoring and replay, however future work should directly test the prediction that tutor sequences themselves are replayed during sleep.

5.0.4 Stage III: Anti-Hebbian plasticity stores a negative image of the tutor song.

Once a temporal basis is stored in the weights of HVC, it can then be recalled during vocal practice. But how could this temporal basis serve to recall a memory of the spectral content of the tutor song? While there is little evidence favoring any particular mechanism, we propose that this could occur by the formation of a negative image of the tutor song on the output of the HVC tutor sequence. As in other systems, negative image formation in Avalanche could be mediated by anti-Hebbian plasticity (5.1 c). More specifically, we hypothesize that the HVC temporal basis sends axons to auditory cortex that target local inhibitory interneurons. During tutoring, we hypothesize that synaptic output from these inhibitory cells onto excitatory neurons, is strengthened in an anti-Hebbian manner, canceling or suppressing tutor evoked activity in Avalanche (5.1 c). At the end of this stage of learning, tutor sequences in HVC drive inhibitory neurons to produce a temporally specific negative image of the tutor song in Avalanche. Mechanisms like this have been proposed to underlie self-image cancellation in the electrosensory system of the electric fish [245], de-

tection of novelty in the mushroom body of the fly [246], predictive coding in the retina [247], and cancellation of self-generated sounds in cortex [248, 249, 250], the dorsal cochlear nucleus [251] and many other systems [252, 253, 254]. In songbirds, the observation of song specific adaptation (SSA) in various auditory regions has been suggested to result from a similar mechanism [74, 75, 76, 77], and was previously hypothesized to underlie tutor memory formation [78]. We propose that the observed song "adaptation", could result from active generation of a negative image.

5.0.5 Stage IV: The tutor memory is recalled during vocal practice.

Once the tutor memory is stored in the synaptic weights of HVC neurons how is it used during singing? We hypothesize that HVC tutor sequences are reactivated during vocal practice (perhaps by NIf [255]) and drive motor activity in downstream RA and LMAN to produce juvenile song (5.1 d). At the same time, these sequences recall a negative image of the tutor song in Avalanche that is synchronized with vocalization. Although, evidence for such a negative image is lacking in songbirds, motor circuits produce widespread suppression in auditory cortical areas during vocalizations in both marmosets and humans [256, 257], consistent with our model.

5.0.6 Stage V: Comparison of song with the tutor memory produces an error signal

At this time, the juvenile bird produces immature vocalizations which are a poor match to the tutor song. Thus the pattern of auditory evoked activity in Avalanche differs from that of the tutor song, and is therefor not matched/suppressed by the negative image (5.1 d). However, as the juvenile bird produces more and more vocalizations, some songs are a closer match to the tutor song than others. Because of the negative image, the overall amount of neural activation in Avalanche reflects the degree to which the produced vocalization matches

the tutor song. This constitutes an error signal which is minimized when an imitation is produced. Crucially, this moment-by-moment error signal is precisely that which is required by reinforcement learning models of song motor learning [12], and could account for the error signals observed in the auditory system [71, 19] and in VTA [15, 16].

5.0.7 Predictions and future work

This hypothesis makes myriad predictions that could be tested in future work. Key predictions include:

1. Sequences are reactivated in the same order during tutoring and vocal learning.

2. Auditory responses in Avalanche to the tutor song become suppressed by repeated tutor exposure, reflecting learning of the predictive cancellation.

3. After responses are in Avalanche are suppressed, neurons exhibit error-like responses to perturbations of the tutor song.

4. Both processes in 2 and 3 are dependent on HVC activity.

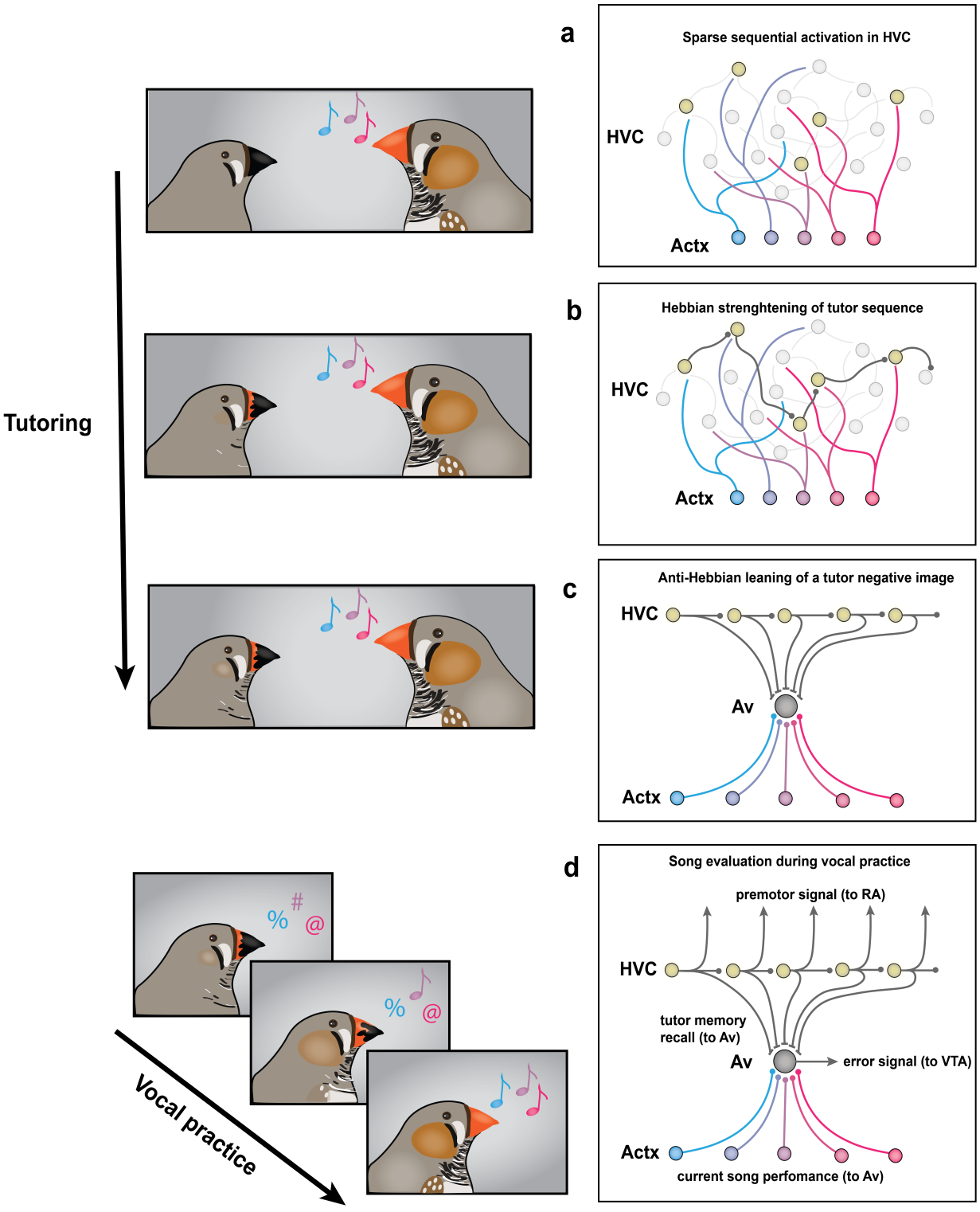


Figure 5.1: A cartoon showing a mechanistic model for HVC's role in tutor memory formation and recall. For details see text above.

Chapter 6

Discussion

In this thesis I have attempted to formulate and test a mechanistic model of HVC's role in the formation and recall of the tutor memory. In the following chapter I will connect the work contained in this thesis to various other aspects of the literature which did not fit neatly into the chapters above.

6.0.1 What is the biophysical basis of the relationship between imitation and HVC temperature?

What causes changes in learning due to temperature? In the case of motor production, it has been suggested that changes in the rate of axonal propagation are the primary cause of song slowing during cooling [146]. This is based on the observation that axon conduction velocity in unmeylinated cortical neurons slows during cooling by 2.5% C^{-1} [138]. The magnitude of this relationship is very close to behavioral effect of cooling during singing in adults [51] and to the effect of cooling on song imitation observed in chapter 2 (figure 2.5.4) . An obvious question is how this squares with the fact that most other neural processes, such as synaptic release [131, 137, 132] or spontaneous firing rate are also slowed by cooling, often to a greater degree [142]. One possibility is relative to these other processes, axon propagation makes up a much larger proportion of the overall transmission time. Consistent

with this, a recent study showed that HVC has particularly slow conduction through axons collaterals. Authors estimated that conduction latency typically adds between 1 and 7.5 milliseconds of delay between a somatic spike and synaptic transmission [54]. Variation in the contributions of different biophysical processes to this overall time[1] could explain the variability of behavioral effects elicited by cooling across different regions and/or different organisms [126, 51, 133, 144]. For example, in some brain areas, the overall time of neural transmission might be taken up by a biophysical process with a Q_{10} near 4. When cooling this hypothetical region, we might expect to observe a steeper relationship between behavior and temperature, relative to a region where transmission time is dominated by a process with a Q_{10} near 1.

6.0.2 Why are the auditory representations in HVC so sparse?

We found that a modest population of HVC excitatory neurons (13%) is active during the tutor song, and that each of these neurons tends to fire at a single time in the tutor song. This sparseness and approximately even tiling of time are reminiscent of the HVC activity during adult production. However, during adult song a very large proportion of excitatory neurons in HVC are recruited; it is estimated that approximately half of all RA projectors are active during singing [13], and upwards of 80% of X projecting excitatory neurons [43]. Even if our recordings were somehow biased towards RA projecting neurons, the auditory evoked sequences recruit a far smaller fraction of HVC neurons than motor sequences. Why might this be? Sparse representations have myriad advantages and have been a topic of great interest in neuroscience. A notable comparison is the olfactory system of insects in which about 10% of kenyon cells respond during presentation of a given odor [180, 181, 182, 183]. As discussed in chapter 3, there is a large literature of theoretical and experimental work pursuing the hypothesis that sparse representations, such those in the olfactory system,

[1] For example there is huge variation between brain regions in the propagation velocity of action potentials depending on axon diameter and myelination [258]. This property can in some cases be very tightly controlled, as in the case of inter-aural timing circuits used for sound localization of the barn owl [259, 260]

cerebellum or dentate gyrus, are useful for learning [178, 186, 187, 184, 185]. We suggest that a similar principle may be at play in HVC, transforming the denser auditory inputs we observed in NIf, L1 and CLM into spatio-temporally sparse representations. As birds are faced with the problem of learning one of an infinite set of possible songs, we suggest this arrangement means that for any given song (of similar duration) a random 13% of HVC neurons will be sparsely and sequentially active. What advantages does this confer? First it ensures that each time in the song is unambiguously represented with near orthogonal neural states and is therefor highly "untangled" (i.e. no times have similar states that evolve in similar ways) [261, 262] . Low tangling was previously proposed as a requirement for a system's activity to be deterministically generated by internal dynamics, not external commands. Here we suggest that the low tangling of HVC activity is evidence, not that these sequences are in fact generated by internal dynamics, but that such dynamics can be learned in principle. By this logic, a song's ability to produce reliable sparse sequences in HVC may be a requirement for learnability and act as a "prior" or inductive bias over song learnability. Such biases are essential for learning algorithms that generalize properly [263, 264]. An easy way of seeing this is by considering song "catchiness". Catchy presumably refers a property of a song that make it easy for the brain to form a memories of. This memory formation occurs to the point of promiscuous encoding; people tend to form strong memories of these experiences automatically (often against their will). We propose that a similar mechanism explains why songbirds find some songs (for example from their own species) highly imitable. Indeed human musical memory and songbird vocal imitation are both examples of near "one-shot" learning. We hypothesize that the combination of highly sparse and untangled representations simplify learning problems immensely, to the point that one-shot learning becomes not only possible but highly probable. This comes at the expense of generality i.e. the auditory systems of humans and birds only allow one-shot learning over a small class of stimuli, not any possible stimuli. The neural properties that govern catchiness in humans and birds is an example of an inductive bias, and illustrates the

trade-off all learning systems must make [265].

6.0.3 Could the small number of auditory projecting HVC neurons support tutor memory recall?

Though sparseness may confer the benefits discussed above, wouldn't one expect that storing a high-fidelity long-term memory, vital for mating success, would require the recruitment of a large percentage or the brain region underlying this capacity? Perhaps in support of this notion, birds with larger brains tend to be better at general cognitive tasks and can also learn more complex vocal imitations [266]. Another possibility is that only a particular class of HVC excitatory neurons respond to the tutor song, but that a high fraction of these neurons are recruited. Indeed several previous studies in other species found that only basal ganglia projecting HVC excitatory neurons are auditory responsive in awake adult birds [27, 9, 99]. This class of HVC neurons comprise a far smaller fraction of the overall population (roughly 10,000 neurons per hemisphere) than motor projecting HVC projection neurons (roughly 40,000 per hemisphere) [267]. If only basal ganglia projecting HVC neurons are active during tutoring then it may be that a large fraction of these neurons is recruited by a given tutor song. If all tutor responsive neurons are basal ganglia projecting (and we assume a uniform probability of recording each cell-type) then this would mean more than half of the relevant neurons are recruited by a given song during tutoring. Interestingly although only a fraction of RA projecting neurons fire during adult song production, most basal ganglia projecting neurons fire during singing. If this class of neurons is involved recalling the tutor memory during singing, then reactivation of a large subset these neurons is consistent with this role.

6.0.4 Auditory responses to song in HVC

One of the most robust findings in the songbird literature is the observation of neurons selective for the birds own song (BOS), in song motor regions [24, 240, 268, 95, 269, 98, 219, 84]. Interestingly, although this property is observed in awake adult birds of other species [27, 31], in zebra finches, this phenomenon is only observed in sleeping or anesthetized birds [240, 268, 270, 72]. This suppression is specific to song motor regions, since responses in auditory cortical areas that provide inputs to HVC are not suppressed during wakefulness [240]. BOS selectivity in HVC emerges over development, and a preference can already be observed in sleeping juvenile birds [23]. Responses to BOS in sleeping zebra finches are observed in all cell types in HVC [271], although awake responses in other species are restricted to excitatory neurons that project to the basal ganglia [27, 31]. A subset of inhibitory interneurons also seem to display waking responses to BOS playback and perhaps act to mediate the suppressed responses during wakefullness, through inhibition of the HVC network [100]. BOS responsive activity is not isolated to HVC but propagates to its down stream targets producing BOS responses in song motor regions in the cortex [104] and basal ganglia [95] and even motor periphery [272]. BOS selectivity also exists in most areas projecting to HVC, including in the nuclues interface (NIf) [26, 29], Avalanche [84] and the thalamic area Uvaformis (Uva) [269]. BOS responses in HVC are abolished when both NIf and Avalanche are inactivated [29, 26]. As we observed during tutor song presentation, NIf neurons tend to fire at multiple points during BOS playback, while HVC projections neurons fire much more sparsely, often with a single burst of activity [271, 29].

6.0.5 Zero latency auditory vocal mirroring

A striking feature of BOS selectivity is that the auditory evoked activity has zero latency relative to the motor activity [104, 27, 31, 99, 72]. More specifically, this means that if a hypothetical HVC neuron fires a bust of spikes at time t during singing, that same neuron also

fires a burst at time t during song playback. Initially this finding is somewhat puzzling, since pre-motor activity leads sound output by approximately 30ms [40], and auditory activity in HVC lags the sound source, again by approximately 30ms due to synaptic delays [29]. How can HVC pre-motor activity and auditory evoked activity exhibit a zero-lag correlation despite an overall 60ms delay between a pre-motor command and its auditory consequence? Some have suggested that the presence of this zero-latency delay is evidence that the motor system constructs a predictive basis to account for the delay between motor commands and their consequences [104]. Interestingly, our model of memory formation in HVC predicts a zero-latency delay as straight-forward consequence of learning. When sequences are formed during tutoring, the pattern of HVC activity is entirely determined by auditory input, i.e. all activity in HVC lags the sound signal by approximately 30ms. However in our model, tutor sequences are replayed during juvenile song, recalling a memory the tutor song content and generating an error signal. At this point in learning, neurons produce no stereotyped motor output but do produce a stereotyped memory recall. This recalled memory completely determines what motor output a neuron ends up producing at the end of learning. Thus, what looks like a predictive mechanism actually emerges as a consequence of memory sequences being recalled during vocal learning in a manner aligned to vocal output. One possible issue with this proposition, is that the first 30ms of the song motif would be missing from such a representation. However, the repeating nature of zebra finch song means that neurons which fire at the end of the tutor motif could potentially fill in this gap, forming a closed cyclical representation of the song. This may explain why repetition is such strong feature of bird song in general and is a favored strategy for learning even in human musicians.

A secondary issue that may be raised with this explanation comes from the technicality that previous findings of zero-latency mirroring used BOS, not the tutor song. If imitation is successful, of course BOS and tutor songs are highly similar and thus neural responses to these stimuli are somewhat confounded [273]. But what about a zero-lag auditory motor correlation in poor imitators? Birds with damaged syringeal nerves still exhibit BOS

selectivity despite highly degraded imitations [274]. How can we explain the zero-latency auditory vocal mirroring observed in birds with poor or degraded imitations? An alternative explanation could be that during practice, auditory inputs to HVC neurons are potentiated onto pre-motor neurons. Although auditory inputs are mostly gated off in awake singing birds [72], birds sing thousands of renditions per day. Despite such gating, it is possible that the extremely high correlation between HVC excitatory neurons and their sensory consequences, could strengthen auditory inputs onto HVC neurons. This could also account for the observed BOS responses in sleeping or anesthetized adults but would not depend on imitation success.

6.0.6 Hemispheric dominance

It is well established that auditory processing of speech in humans exhibits a left hemisphere dominance [275, 276]. Interestingly, while we have not studied this phenomenon here, a similar left hemisphere dominance appears to be present in auditory processing in songbirds. Lesions of left hemisphere auditory thalamus, degrade adult canaries ability to discriminate between songs much more than lesions of the right hemisphere [277]. Additional findings related to lateralization of tutor song processing however are interesting if not somewhat confusing. The secondary auditory area NCM is left-lateralized in processing the tutor song, but not novel songs [278], and the degree of left-lateralization in NCM is correlated with imitation accuracy [278, 279]. HVC responses to playback of the tutor song (as well as spontaneous activity) are left-lateralized in awake juveniles [278], but are right-lateralized in sleeping adults [279]. On the motor side of things, transections of the right tracheosyringeal nerve (which controls contractions of the vocal musculature), tends to degrade song more than transections of the left side [280]. Unilateral cooling of left or right HVC tends to slow somewhat different portions of the song [51], and brief unilateral stimulation of either left or right hemisphere in song motor regions disrupts different portions of the song [281]. Despite this somewhat confusing set of result, the songbird model system is especially well suited

for studying lateralization due to its well understood neurophysiology. Future work should compare tutor responses across hemispheres to test whether adult motor lateralization can be related to this early auditory exposure.

6.0.7 HVC avalanche connectivity as a for tutor song recall

In this thesis we proposed that the mechanism by which HVC tutor sequences recall the content of the tutor memory is by its projection to the auditory area Avalanche [84]. We hypothesize that during tutoring plasticity at HVC synapses in Avalanche undergo plasticity onto local inhibitory neurons to produce a negative image of the tutor song, canceling tutor evoked activity in excitatory neurons in this region. This negative image is then recalled by HVC sequences during singing and produces an evaluation signal which is sent to VTA; Avalanche excitatory activity is minimized when the pupil produces a close match to the tutor song. This hypothesis leads to several strong yet untested predictions. First, HVC neurons active during tutoring must be reactivated in the same manner during song production. Second HVC tutor sequences must contain a significant proportion of neurons that project to the auditory system (Avalanche, HVCshelf and/or other regions). Finally, those auditory projecting neurons should synapse onto local interneurons (although there are other versions of the model in which this assumption could be relaxed [90]).

There are a number of lines of evidence that indirectly support these points. Optogenetic stimulation paired with patch clamp recordings suggests that Avalanche projecting HVC neurons (HVC_{Av}) receive direct mono-synaptic input from NIf neurons [282] suggesting these neurons could be recruited during tutoring. HVC_{Av} neurons are also known to fire sparsely during adult song production, similar to other excitatory projection neurons [21]. A concern though, is whether HVC_{Av} neurons constitute a large enough population to serve a functional role. A previous study suggested that these neurons make up less than 1% of HVC neurons (less than 100 cells in an adult!) suggesting to some that this cell type constitutes more or less a "rounding error" in terms prevalence. This claim though appears to come from directly

counting the number of cell bodies labeled by retrograde tracer injected in to Avalanche which is indeed close to 100 [21]. Such a method of estimating cell counts assumes that 100% of neurons will be labeled by retrograde tracing and provides and artificially low estimate since retrograde tracers in fact only label a small fraction of neurons that project. Another experiment in the same study compared the number of cell bodies labeled by injections into the basal ganglia to the number labeled by injections in the Avalanche. They found that injections in Avalanche labeled 30-50% as many neurons as injections into the Area X in the same birds. Since these basal ganglia projecting neurons constitute about 10,000 neurons [267], this gives an estimate of between 3,000 and 5,000 avalanche projectors. Finally, a recent study estimated that HVC_{Av} cells make up about 2% of excitatory neurons in adult HVC based on single-cell transcriptomics data [283]. All estimates were obtained in adult birds, which likely also underestimates the fraction of HVC_{Av} neurons present in juveniles, since HVC_{RA} projectors grow dramatically in number during song learning. Although estimating absolute numbers of neurons of different cell-types is difficult and relies on many assumptions, taken together it seems likely that Avalanche projecting neurons likely comprise more than 1% of HVC neurons and this number may reach as high as 5% [2].

Finally we have the question of whether HVC_{Av} neurons could transmit a negative image of the tutor song during vocal practice via inhibitory internurons in Avalanche. Avalanche neurons are indeed active during vocal practice as revealed by Arc, Zenk, Erg1 and c-Fos activity dependent gene expression [284, 285], however the cell types of these active neuron are unknown. Interestingly, although genetic lesions of HVC_{Av} neurons impair imitation in juvenile birds, the same manipulation in adult birds does not impair birds ability to modify a syllables pitch in conditional auditory feedback experiments (CAF) (perhaps indicating this process does not require template comparison) [21]. When conditional auditory feedback is turned off, adult control birds rapidly undo these learned pitch changes while birds with le-

[2]Experience in our lab suggests highly variable labeling of HVC neurons after injections into Avalanche. It is unclear whether this variability is a result of inter-animal variablity in cell number or perhaps in spatial location of Avalanche itself

sions of HVC_{Av} neurons, do not. Similarly, deafened adult birds songs rapidly degrade which has been interpreted as a consequence of the fact that they constantly perceive themselves as making errors when they sing, however lesioning HVC_{Av}neurons prevents degradation. Together these results indicate that HVC_{Av} neurons may be involved in the error computation. In mice, auditory cortical projections from M2 selectively engage local inhibitory interneurons in auditory cortex [250] and a recent presentation at the Society for Neuroscience Conference [286] demonstrated that auditory areas are suppressed during singing by motor circuits. However, whether the HVC produces inhibition (and perhaps a negative image of the tutor song) in Avalanche during vocal practice remains to be established by future work.

References

[1] O Tchernichovski, P P Mitra, T Lints, and F Nottebohm. "Dynamics of the vocal imitation process: how a zebra finch learns its song". en. In: *Science* 291.5513 (Mar. 2001), pp. 2564–2569.

[2] K Immelmann. "Song development in the zebra finch and other estrildid finches". In: *Bird Vocalizations* (1969), pp. 61–77.

[3] Philip H Price. "Developmental determinants of structure in zebra finch song". In: *J. Comp. Physiol. Psychol.* 93.2 (Apr. 1979), pp. 260–277.

[4] Christopher M Glaze and Todd W Troyer. "Temporal structure in zebra finch song: implications for motor coding". en. In: *J. Neurosci.* 26.3 (Jan. 2006), pp. 991–1005.

[5] A J Doupe and P K Kuhl. "Birdsong and human speech: common themes and mechanisms". en. In: *Annu. Rev. Neurosci.* 22 (1999), pp. 567–631.

[6] S A White. "Learning to communicate". en. In: *Curr. Opin. Neurobiol.* 11.4 (Aug. 2001), pp. 510–520.

[7] Ikuko Teramitsu, Lili C Kudo, Sarah E London, Daniel H Geschwind, and Stephanie A White. "Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction". en. In: *J. Neurosci.* 24.13 (Mar. 2004), pp. 3152–3163.

[8] Andreas R Pfenning et al. “Convergent transcriptional specializations in the brains of humans and song-learning birds”. en. In: *Science* 346.6215 (Dec. 2014), p. 1256846.

[9] Jonathan F Prather, Kazuo Okanoya, and Johan J Bolhuis. “Brains for birds and babies: Neural parallels between birdsong and speech acquisition”. en. In: *Neurosci. Biobehav. Rev.* 81.Pt B (Oct. 2017), pp. 225–237.

[10] Kenji Doya and Terrence J Sejnowski. “A novel reinforcement model of birdsong vocalization learning”. In: *Adv. Neural Inf. Process. Syst.* 7 (1994).

[11] Lei Xiao and Todd F Roberts. “What Is the Role of Thalamostriatal Circuits in Learning Vocal Sequences?” en. In: *Front. Neural Circuits* 15 (Sept. 2021), p. 724858.

[12] M S Fee and J H Goldberg. “A hypothesis for basal ganglia-dependent reinforcement learning in the songbird”. en. In: *Neuroscience* 198 (Dec. 2011), pp. 152–170.

[13] Michale S Fee, Alexay A Kozhevnikov, and Richard H R Hahnloser. “Neural mechanisms of vocal sequence generation in the songbird”. en. In: *Ann. N. Y. Acad. Sci.* 1016 (June 2004), pp. 153–170.

[14] Richard Mooney. “Neural mechanisms for learned birdsong”. en. In: *Learn. Mem.* 16.11 (Nov. 2009), pp. 655–669.

[15] Vikram Gadagkar, Pavel A Puzerey, Ruidong Chen, Eliza Baird-Daniel, Alexander R Farhang, and Jesse H Goldberg. “Dopamine neurons encode performance error in singing birds”. en. In: *Science* 354.6317 (Dec. 2016), pp. 1278–1282.

[16] Alison Duffy, Kenneth W Latimer, Jesse H Goldberg, Adrienne L Fairhall, and Vikram Gadagkar. “Dopamine neurons evaluate natural fluctuations in performance quality”. en. In: *Cell Rep.* 38.13 (Mar. 2022), p. 110574.

[17] Andrea Roeser, Vikram Gadagkar, Anindita Das, Pavel A Puzerey, Brian Kardon, and Jesse H Goldberg. "Dopaminergic signals for reward, performance and social outcomes are dynamically gated during courtship". In: *BioRxiv* (2019), p. 822817.

[18] Sarah C Woolley, Raghav Rajan, Mati Joshua, and Allison J Doupe. "Emergence of context-dependent variability across a basal ganglia network". en. In: *Neuron* 82.1 (Apr. 2014), pp. 208–223.

[19] Yael Mandelblat-Cerf, Liora Las, Natalia Denisenko, and Michale S Fee. "A role for descending auditory cortical projections in songbird vocal learning". en. In: *Elife* 3 (June 2014).

[20] Wenchan Zhao, Francisco Garcia-Oscos, Daniel Dinh, and Todd F Roberts. "Inception of memories that guide vocal learning in the songbird". en. In: *Science* 366.6461 (Oct. 2019), pp. 83–89.

[21] Todd F Roberts, Erin Hisey, Masashi Tanaka, Matthew G Kearney, Gaurav Chattree, Cindy F Yang, Nirao M Shah, and Richard Mooney. "Identification of a motor-to-auditory pathway important for vocal learning". en. In: *Nat. Neurosci.* 20.7 (July 2017), pp. 978–986.

[22] Todd F Roberts, Sharon M H Gobes, Malavika Murugan, Bence P Ölveczky, and Richard Mooney. "Motor circuits are required to encode a sensory model for imitative learning". en. In: *Nat. Neurosci.* 15.10 (Oct. 2012), pp. 1454–1459.

[23] Teresa A Nick and Masakazu Konishi. "Neural auditory selectivity develops in parallel with song". en. In: *J. Neurobiol.* 62.4 (Mar. 2005), pp. 469–481.

[24] D Margoliash and M Konishi. "Auditory representation of autogenous song in the song system of white-crowned sparrows". en. In: *Proc. Natl. Acad. Sci. U. S. A.* 82.17 (Sept. 1985), pp. 5997–6000.

[25] Teresa A Nick and Masakazu Konishi. “Neural song preference during vocal learning in the zebra finch depends on age and state”. en. In: *J. Neurobiol.* 62.2 (Feb. 2005), pp. 231–242.

[26] Eric E Bauer, Melissa J Coleman, Todd F Roberts, Arani Roy, Jonathan F Prather, and Richard Mooney. “A synaptic basis for auditory–vocal integration in the songbird”. In: *Journal of Neuroscience* 28.6 (2008), pp. 1509–1522.

[27] J F Prather, S Peters, S Nowicki, and R Mooney. “Precise auditory–vocal mirroring in neurons for learned vocal communication”. en. In: *Nature* 451.7176 (Jan. 2008), pp. 305–310.

[28] S Peters, S Nowicki, and R Mooney. “Persistent representation of juvenile experience in the adult songbird brain”. In: *Journal of* (2010).

[29] Melissa J Coleman and R Mooney. “Synaptic transformations underlying highly selective auditory representations of learned birdsong”. en. In: *J. Neurosci.* 24.33 (Aug. 2004), pp. 7251–7265.

[30] Wendy Y X Peh, Todd F Roberts, and Richard Mooney. “Imaging auditory representations of song and syllables in populations of sensorimotor neurons essential to vocal communication”. en. In: *J. Neurosci.* 35.14 (Apr. 2015), pp. 5589–5605.

[31] Jonathan F Prather, Susan Peters, Stephen Nowicki, and Richard Mooney. “Persistent representation of juvenile experience in the adult songbird brain”. en. In: *J. Neurosci.* 30.31 (Aug. 2010), pp. 10586–10598.

[32] Patrice Adret. “In search of the song template”. en. In: *Ann. N. Y. Acad. Sci.* 1016 (June 2004), pp. 303–324.

[33] Heather Williams. “Birdsong and singing behavior”. en. In: *Ann. N. Y. Acad. Sci.* 1016 (June 2004), pp. 1–30.

[34] F Nottebohm, D B Kelley, and J A Paton. “Connections of vocal control nuclei in the canary telencephalon”. en. In: *J. Comp. Neurol.* 207.4 (June 1982), pp. 344–357.

[35] J M Wild. “Neural pathways for the control of birdsong production”. en. In: *J. Neurobiol.* 33.5 (Nov. 1997), pp. 653–670.

[36] Niangui Wang, Patrick Hurley, Carolyn Pytte, and John R Kirn. “Vocal control neuron incorporation decreases with age in the adult zebra finch”. en. In: *J. Neurosci.* 22.24 (Dec. 2002), pp. 10864–10870.

[37] M E Gurney. “Hormonal control of cell form and number in the zebra finch song system”. en. In: *J. Neurosci.* 1.6 (June 1981), pp. 658–673.

[38] D S Vicario and F Nottebohm. “Organization of the zebra finch song control system: I. Representation of syringeal muscles in the hypoglossal nucleus”. en. In: *J. Comp. Neurol.* 271.3 (May 1988), pp. 346–354.

[39] J M Wild. “Descending projections of the songbird nucleus robustus archistriatalis”. en. In: *J. Comp. Neurol.* 338.2 (Dec. 1993), pp. 225–241.

[40] Galen F Lynch, Tatsuo S Okubo, Alexander Hanuschkin, Richard H R Hahnloser, and Michale S Fee. “Rhythmic Continuous-Time Coding in the Songbird Analog of Vocal Motor Cortex”. en. In: *Neuron* 90.4 (May 2016), pp. 877–892.

[41] Michel A Picardo, Josh Merel, Kalman A Katlowitz, Daniela Vallentin, Daniel E Okobi, Sam E Benezra, Rachel C Clary, Eftychios A Pnevmatikakis, Liam Paninski, and Michael A Long. “Population-Level Representation of a Temporal Sequence Underlying Song Production in the Zebra Finch”. en. In: *Neuron* 90.4 (May 2016), pp. 866–876.

[42] Richard H R Hahnloser, Alexay A Kozhevnikov, and Michale S Fee. “An ultra-sparse code underliesthe generation of neural sequences in a songbird”. en. In: *Nature* 419.6902 (Sept. 2002), pp. 65–70.

[43] Alexay A Kozhevnikov and Michale S Fee. “Singing-related activity of identified HVC neurons in the zebra finch”. en. In: *J. Neurophysiol.* 97.6 (June 2007), pp. 4271–4283.

[44] Michael A Long, Dezhe Z Jin, and Michale S Fee. “Support for a synaptic chain model of neuronal sequence generation”. en. In: *Nature* 468.7322 (Nov. 2010), pp. 394–399.

[45] Vamsi K Daliparthi, Ryosuke O Tachibana, Brenton G Cooper, Richard Hr Hahnloser, Satoshi Kojima, Samuel J Sober, and Todd F Roberts. “Transitioning between preparatory and precisely sequenced neuronal activity in production of a skilled behavior”. en. In: *Elife* 8 (June 2019).

[46] Anthony Leonardo and Michale S Fee. “Ensemble coding of vocal control in birdsong”. en. In: *J. Neurosci.* 25.3 (Jan. 2005), pp. 652–661.

[47] A C Yu and D Margoliash. “Temporal hierarchical control of singing in birds”. en. In: *Science* 273.5283 (Sept. 1996), pp. 1871–1875.

[48] Bence P Ölveczky, Timothy M Otchy, Jesse H Goldberg, Dmitriy Aronov, and Michale S Fee. “Changes in the neural control of a complex motor sequence during learning”. en. In: *J. Neurophysiol.* 106.1 (July 2011), pp. 386–397.

[49] Iris Adam and Coen P H Elemans. “Increasing Muscle Speed Drives Changes in the Neuromuscular Transform of Motor Commands during Postnatal Development in Songbirds”. en. In: *J. Neurosci.* 40.35 (Aug. 2020), pp. 6722–6731.

[50] Kosuke Hamaguchi, Masashi Tanaka, and Richard Mooney. “A Distributed Recurrent Network Contributes to Temporally Precise Vocalizations”. en. In: *Neuron* 91.3 (Aug. 2016), pp. 680–693.

[51] Michael A Long and Michale S Fee. “Using temperature to analyse temporal dynamics in the songbird motor pathway”. en. In: *Nature* 456.7219 (Nov. 2008), pp. 189–194.

[52] F Nottebohm, T M Stokes, and C M Leonard. “Central control of song in the canary, Serinus canarius”. en. In: *J. Comp. Neurol.* 165.4 (Feb. 1976), pp. 457–486.

[53] E T Vu, M E Mazurek, and Y C Kuo. “Identification of a forebrain motor programming network for the learned song of zebra finches”. en. In: *J. Neurosci.* 14.11 Pt 2 (Nov. 1994), pp. 6924–6934.

[54] Robert Egger, Yevhen Tupikov, Margot Elmaleh, Kalman A Katlowitz, Sam E Benezra, Michel A Picardo, Felix Moll, Jörgen Kornfeld, Dezhe Z Jin, and Michael A Long. “Local Axonal Conduction Shapes the Spatiotemporal Properties of Neural Sequences”. en. In: *Cell* 183.2 (Oct. 2020), 537–548.e12.

[55] Ila R Fiete, Walter Senn, Claude Z H Wang, and Richard H R Hahnloser. “Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity”. en. In: *Neuron* 65.4 (Feb. 2010), pp. 563–576.

[56] Tatsuo S Okubo, Emily L Mackevicius, Hannah L Payne, Galen F Lynch, and Michale S Fee. “Growth and splitting of neural sequences in songbird vocal development”. en. In: *Nature* 528.7582 (Dec. 2015), pp. 352–357.

[57] Yevhen Tupikov and Dezhe Z Jin. “Addition of new neurons and the emergence of a local neural circuit for precise timing”. en. In: *PLoS Comput. Biol.* 17.3 (Mar. 2021), e1008824.

[58] Margot Elmaleh, Devorah Kranz, Ariadna Corredera Asensio, Felix W Moll, and Michael A Long. “Sleep replay reveals premotor circuit structure for a skilled behavior”. en. In: *Neuron* 109.23 (Dec. 2021), 3851–3861.e4.

[59] Bence P Olveczky, Aaron S Andalman, and Michale S Fee. “Vocal experimentation in the juvenile songbird requires a basal ganglia circuit”. en. In: *PLoS Biol.* 3.5 (May 2005), e153.

[60] Mimi H Kao and Michael S Brainard. “Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability”. en. In: *J. Neurophysiol.* 96.3 (Sept. 2006), pp. 1441–1455.

[61] Laurie Stepanek and Allison J Doupe. “Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability”. en. In: *J. Neurophysiol.* 104.5 (Nov. 2010), pp. 2474–2486.

[62] Dmitriy Aronov and Michale S Fee. “Analyzing the dynamics of brain circuits with temperature: design and implementation of a miniature thermoelectric device”. en. In: *J. Neurosci. Methods* 197.1 (Apr. 2011), pp. 32–47.

[63] Kosuke Hamaguchi and Richard Mooney. “Recurrent interactions between the input and output of a songbird cortico-basal ganglia pathway are implicated in vocal sequence variability”. en. In: *J. Neurosci.* 32.34 (Aug. 2012), pp. 11671–11687.

[64] Jonnathan Singh Alvarado, Jack Goffinet, Valerie Michael, William Liberti 3rd, Jordan Hatfield, Timothy Gardner, John Pearson, and Richard Mooney. “Neural dynamics underlying birdsong practice and performance”. en. In: *Nature* 599.7886 (Nov. 2021), pp. 635–639.

[65] Lukas A Hoffmann, Varun Saravanan, Alynda N Wood, Li He, and Samuel J Sober. “Dopaminergic Contributions to Vocal Learning”. en. In: *J. Neurosci.* 36.7 (Feb. 2016), pp. 2176–2189.

[66] Samuel D Gale and David J Perkel. “Anatomy of a songbird basal ganglia circuit essential for vocal learning and plasticity”. en. In: *J. Chem. Neuroanat.* 39.2 (Mar. 2010), pp. 124–131.

[67] Malavika Ramarao, Andrea Roeser, Caleb Jones, and Jesse H Goldberg. “Songbird mesostriatal dopamine pathways are spatially segregated before the onset of vocal learning”. en. May 2023.

[68] Erin Hisey, Matthew Gene Kearney, and Richard Mooney. “A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning”. en. In: *Nat. Neurosci.* 21.4 (Apr. 2018), pp. 589–597.

[69] Ruidong Chen, Pavel A Puzerey, Andrea C Roeser, Tori E Riccelli, Archana Podury, Kamal Maher, Alexander R Farhang, and Jesse H Goldberg. "Songbird Ventral Pallidum Sends Diverse Performance Error Signals to Dopaminergic Midbrain". en. In: *Neuron* 103.2 (July 2019), 266–276.e4.

[70] Matthew Gene Kearney, Timothy L Warren, Erin Hisey, Jiaxuan Qi, and Richard Mooney. "Discrete Evaluative and Premotor Circuits Enable Vocal Learning in Songbirds". en. In: *Neuron* 104.3 (Nov. 2019), 559–575.e6.

[71] Georg B Keller and Richard H R Hahnloser. "Neural processing of auditory feedback during vocal practice in a songbird". en. In: *Nature* 457.7226 (Jan. 2009), pp. 187–190.

[72] Kosuke Hamaguchi, Katherine A Tschida, Inho Yoon, Bruce R Donald, and Richard Mooney. "Auditory synapses to song premotor neurons are gated off during vocalization in zebra finches". en. In: *Elife* 3 (Feb. 2014), e01833.

[73] J Wang, M Sokabe, and H Sakaguchi. "Functional connections between the HVC and the shelf of the zebra finch revealed by real-time optical imaging technique". en. In: *Neuroreport* 12.2 (Feb. 2001), pp. 215–221.

[74] Kai Lu and David S Vicario. "Familiar But Unexpected: Effects of Sound Context Statistics on Auditory Responses in the Songbird Forebrain". en. In: *J. Neurosci.* 37.49 (Dec. 2017), pp. 12006–12017.

[75] S J Chew, C Mello, F Nottebohm, E Jarvis, and D S Vicario. "Decrements in auditory responses to a repeated conspecific song are long-lasting and require two periods of protein synthesis in the songbird forebrain". In: *Proceedings of the National Academy of Sciences* 92.8 (1995), pp. 3406–3410.

[76] S J Chew, D S Vicario, and F Nottebohm. "A large-capacity memory system that recognizes the calls and songs of individual birds". en. In: *Proc. Natl. Acad. Sci. U. S. A.* 93.5 (Mar. 1996), pp. 1950–1955.

[77] Patrick Gill, Sarah M N Woolley, Thane Fremouw, and Frédéric E Theunissen. "What's that sound? Auditory area CLM encodes stimulus surprise, not intensity or intensity changes". en. In: *J. Neurophysiol.* 99.6 (June 2008), pp. 2809–2820.

[78] R H R Hahnloser and A Kotowicz. "Auditory representations and memory in birdsong learning". In: *Curr. Opin. Neurobiol.* (2010).

[79] Johan J Bolhuis and Sanne Moorman. "Birdsong memory and the brain: in search of the template". en. In: *Neurosci. Biobehav. Rev.* 50 (Mar. 2015), pp. 41–55.

[80] Alessandro Canopoli, Anja Zai, and Richard Hahnloser. "Lesions of a higher auditory brain area during a sensorimotor period do not impair birdsong learning". In: *Matters* (Mar. 2016).

[81] Alessandro Canopoli, Anja Zai, and Richard H R Hahnloser. "Bilateral neurotoxic lesions in NCM before tutoring onset do not prevent successful tutor song learning". en. In: *Matters* (2017), pp. 12.5/20.

[82] Sarah E London and David F Clayton. "Functional identification of sensory mechanisms required for developmental song learning". en. In: *Nat. Neurosci.* 11.5 (May 2008), pp. 579–586.

[83] Sharon M H Gobes and Johan J Bolhuis. "Birdsong memory: a neural dissociation between song recognition and production". en. In: *Curr. Biol.* 17.9 (May 2007), pp. 789–793.

[84] Eugene Akutagawa and Masakazu Konishi. "New brain pathways found in the vocal control system of a songbird". en. In: *J. Comp. Neurol.* 518.15 (Aug. 2010), pp. 3086–3100.

[85] Jessica A Cardin, Jonathan N Raksin, and Marc F Schmidt. "Sensorimotor nucleus NIf is necessary for auditory processing but not vocal motor output in the avian song system". en. In: *J. Neurophysiol.* 93.4 (Apr. 2005), pp. 2157–2166.

[86] David M Schneider and Richard Mooney. “How Movement Modulates Hearing”. en. In: *Annu. Rev. Neurosci.* 41 (July 2018), pp. 553–572.

[87] Maaya Z Ikeda, Massimo Trusel, and Todd F Roberts. “Memory circuits for vocal imitation”. en. In: *Curr. Opin. Neurobiol.* 60 (Feb. 2020), pp. 37–46.

[88] Todd F Roberts, Katherine A Tschida, Marguerita E Klein, and Richard Mooney. “Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning”. en. In: *Nature* 463.7283 (Feb. 2010), pp. 948–952.

[89] Daniela Vallentin, Georg Kosche, Dina Lipkind, and Michael A Long. “Neural circuits. Inhibition protects acquired song segments during vocal learning in zebra finches”. en. In: *Science* 351.6270 (Jan. 2016), pp. 267–271.

[90] Emily Lambert Mackevicius and Michale Sean Fee. “Building a state space for song learning”. en. In: *Curr. Opin. Neurobiol.* 49 (Apr. 2018), pp. 59–68.

[91] Patrice Adret, C Daniel Meliza, and Daniel Margoliash. “Song tutoring in presinging zebra finch juveniles biases a small population of higher-order song-selective neurons toward the tutor song”. en. In: *J. Neurophysiol.* 108.7 (Oct. 2012), pp. 1977–1987.

[92] Efe Soyman and David S Vicario. “Principles of auditory processing differ between sensory and premotor structures of the songbird forebrain”. en. In: *J. Neurophysiol.* 117.3 (Mar. 2017), pp. 1266–1280.

[93] S F Volman. “Development of neural selectivity for birdsong during vocal learning”. en. In: *J. Neurosci.* 13.11 (Nov. 1993), pp. 4737–4747.

[94] M M Solis and A J Doupe. “Compromised neural selectivity for song in birds with impaired sensorimotor learning”. en. In: *Neuron* 25.1 (Jan. 2000), pp. 109–121.

[95] A J Doupe. “Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development”. en. In: *J. Neurosci.* 17.3 (Feb. 1997), pp. 1147–1167.

[96] A J Doupe and M Konishi. “Song-selective auditory circuits in the vocal control system of the zebra finch”. en. In: *Proc. Natl. Acad. Sci. U. S. A.* 88.24 (Dec. 1991), pp. 11339–11343.

[97] D Margoliash and E S Fortune. “Temporal and harmonic combination-sensitive neurons in the zebra finch’s HVc”. en. In: *J. Neurosci.* 12.11 (Nov. 1992), pp. 4309–4326.

[98] P Janata and D Margoliash. “Gradual emergence of song selectivity in sensorimotor structures of the male zebra finch song system”. en. In: *J. Neurosci.* 19.12 (June 1999), pp. 5108–5118.

[99] Hisataka Fujimoto, Taku Hasegawa, and Dai Watanabe. “Neural coding of syntactic structure in learned vocalizations in the songbird”. en. In: *J. Neurosci.* 31.27 (July 2011), pp. 10023–10033.

[100] Peter L Rauske, Stephen D Shea, and Daniel Margoliash. “State and neuronal class-dependent reconfiguration in the avian song system”. en. In: *J. Neurophysiol.* 89.3 (Mar. 2003), pp. 1688–1701.

[101] Masashi Tanaka, Fangmiao Sun, Yulong Li, and Richard Mooney. “A mesocortical dopamine circuit enables the cultural transmission of vocal behaviour”. en. In: *Nature* 563.7729 (Nov. 2018), pp. 117–120.

[102] Ziqiang Huang, Houda G Khaled, Moritz Kirschmann, Sharon Mh Gobes, and Richard Hr Hahnloser. “Excitatory and inhibitory synapse reorganization immediately after critical sensory experience in a vocal learner”. en. In: *Elife* 7 (Oct. 2018).

[103] Sylvan S Shank and Daniel Margoliash. “Sleep and sensorimotor integration during early vocal learning in a songbird”. en. In: *Nature* 458.7234 (Mar. 2009), pp. 73–77.

[104] A S Dave and D Margoliash. “Song replay during sleep and computational rules for sensorimotor vocal learning”. en. In: *Science* 290.5492 (Oct. 2000), pp. 812–816.

[105] Masakazu Konishi. “The role of auditory feedback in the control of vocalization in the white-crowned sparrow[1]”. en. In: *Z. Tierpsychol.* 22.7 (Jan. 1965), pp. 770–783.

[106] Peter Marler. “A comparative approach to vocal learning: Song development in white-crowned sparrows”. In: *J. Comp. Physiol. Psychol.* 71.2p2 (May 1970), pp. 1–25.

[107] Paul Bloom. *How Children Learn the Meanings of Words*. en. MIT Press, Jan. 2002.

[108] Kelly Jakubowski, Sebastian Finkel, Lauren Stewart, and Daniel Müllensiefen. “Dissecting an earworm: Melodic features and song popularity predict involuntary musical imagery”. en. In: *Psychol. Aesthet. Creat. Arts* 11.2 (May 2017), pp. 122–135.

[109] Asaf Tsoar, Ran Nathan, Yoav Bartan, Alexei Vyssotski, Giacomo Dell’Omo, and Nachum Ulanovsky. “Large-scale navigational map in a mammal”. en. In: *Proc. Natl. Acad. Sci. U. S. A.* 108.37 (Sept. 2011), E718–24.

[110] Lauren J N Brent. “Friends of friends: are indirect connections in social networks important to animal behaviour?” en. In: *Anim. Behav.* 103 (May 2015), pp. 211–222.

[111] Sheena A Josselyn, Stefan Köhler, and Paul W Frankland. “Finding the engram”. en. In: *Nat. Rev. Neurosci.* 16.9 (Sept. 2015), pp. 521–534.

[112] Susumu Tonegawa, Mark D Morrissey, and Takashi Kitamura. “The role of engram cells in the systems consolidation of memory”. en. In: *Nat. Rev. Neurosci.* 19.8 (Aug. 2018), pp. 485–498.

[113] Gaetan de Lavilléon, Marie Masako Lacroix, Laure Rondi-Reig, and Karim Benchenane. “Explicit memory creation during sleep demonstrates a causal role of place cells in navigation”. en. In: *Nat. Neurosci.* 18.4 (Apr. 2015), pp. 493–495.

[114] Mugdha Deshpande, Fakhriddin Pirlepesov, and Thierry Lints. “Rapid encoding of an internal model for imitative learning”. en. In: *Proc. Biol. Sci.* 281.1781 (Apr. 2014), p. 20132630.

[115] Daniela Vallentin and Michael A Long. “Motor origin of precise synaptic inputs onto forebrain neurons driving a skilled behavior”. en. In: *J. Neurosci.* 35.1 (Jan. 2015), pp. 299–307.

[116] Pengcheng Zhou et al. “Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data”. en. In: *Elife* 7 (Feb. 2018).

[117] Aaron S Andalman, Jakob N Foerster, and Michale S Fee. “Control of vocal and respiratory patterns in birdsong: dissection of forebrain and brainstem mechanisms using temperature”. en. In: *PLoS One* 6.9 (Sept. 2011), e25461.

[118] Dmitriy Aronov and Michale S Fee. “Natural changes in brain temperature underlie variations in song tempo during a mating behavior”. en. In: *PLoS One* 7.10 (Oct. 2012), e47856.

[119] Emily L Mackevicius, Shijie Gu, Natalia I Denisenko, and Michale S Fee. “Self-organization of songbird neural sequences during social isolation”. en. In: *Elife* 12 (May 2023).

[120] O Tchernichovski, T Lints, P P Mitra, and F Nottebohm. “Vocal imitation in zebra finches is inversely related to model abundance”. en. In: *Proc. Natl. Acad. Sci. U. S. A.* 96.22 (Oct. 1999), pp. 12901–12904.

[121] O Tchernichovski, F Nottebohm, C E Ho, B Pesaran, and P P Mitra. “A procedure for an automated measurement of song similarity”. en. In: *Anim. Behav.* 59.6 (June 2000), pp. 1167–1176.

[122] Michael A Long, Kalman A Katlowitz, Mario A Svirsky, Rachel C Clary, Tara Mcallister Byun, Najib Majaj, Hiroyuki Oya, Matthew A Howard 3rd, and

Jeremy D W Greenlee. “Functional Segregation of Cortical Regions Underlying Speech Timing and Articulation”. en. In: *Neuron* 89.6 (Mar. 2016), pp. 1187–1193.

[123] Matıas A Goldin, Leandro M Alonso, Jorge A Alliende, Franz Goller, and Gabriel B Mindlin. “Temperature induced syllable breaking unveils nonlinearly interacting timescales in birdsong motor pathway”. en. In: *PLoS One* 8.6 (June 2013), e67814.

[124] Yisi S Zhang, Jason D Wittenbach, Dezhe Z Jin, and Alexay A Kozhevnikov. “Temperature Manipulation in Songbird Brain Implicates the Premotor Nucleus HVC in Birdsong Syntax”. en. In: *J. Neurosci.* 37.10 (Mar. 2017), pp. 2600–2611.

[125] Min Xu, Si-Yu Zhang, Yang Dan, and Mu-Ming Poo. “Representation of interval timing by temporally scalable firing patterns in rat prefrontal cortex”. en. In: *Proc. Natl. Acad. Sci. U. S. A.* 111.1 (Jan. 2014), pp. 480–485.

[126] Tiago Monteiro, Filipe S Rodrigues, Margarida Pexirra, Bruno F Cruz, Ana I Gonçalves, Pavel E Rueda-Orozco, and Joseph J Paton. “Using temperature to analyze the neural basis of a time-based decision”. en. In: *Nat. Neurosci.* 26.8 (Aug. 2023), pp. 1407–1416.

[127] Peter Christian Petersen and György Buzsáki. “Cooling of Medial Septum Reveals Theta Phase Lag Coordination of Hippocampal Cell Assemblies”. en. In: *Neuron* 107.4 (Aug. 2020), 731–744.e3.

[128] Daniel E Okobi Jr, Arkarup Banerjee, Andrew M M Matheson, Steven M Phelps, and Michael A Long. “Motor cortical control of vocal interaction in neotropical singing mice”. en. In: *Science* 363.6430 (Mar. 2019), pp. 983–988.

[129] J W Hastings and B M Sweeney. “ON THE MECHANISM OF TEMPERATURE INDEPENDENCE IN A BIOLOGICAL CLOCK”. en. In: *Proc. Natl. Acad. Sci. U. S. A.* 43.9 (Sept. 1957), pp. 804–811.

[130] Bryan A Reyes, Julie S Pendergast, and Shin Yamazaki. “Mammalian peripheral circadian oscillators are temperature compensated”. en. In: *J. Biol. Rhythms* 23.1 (Feb. 2008), pp. 95–98.

[131] A L Hodgkin and B Katz. “The effect of temperature on the electrical activity of the giant axon of the squid”. en. In: *J. Physiol.* 109.1-2 (Aug. 1949), pp. 240–249.

[132] M Volgushev, T R Vidyasagar, M Chistiakova, and U T Eysel. “Synaptic transmission in the neocortex during reversible cooling”. en. In: *Neuroscience* 98.1 (2000), pp. 9–22.

[133] Lamont S Tang, Marie L Goeritz, Jonathan S Caplan, Adam L Taylor, Mehmet Fisek, and Eve Marder. “Precise temperature compensation of phase in a rhythmic motor pattern”. en. In: *PLoS Biol.* 8.8 (Aug. 2010).

[134] Wafa Soofi, Marie L Goeritz, Tilman J Kispersky, Astrid A Prinz, Eve Marder, and Wolfgang Stein. “Phase maintenance in a rhythmic motor pattern during temperature changes in vivo”. en. In: *J. Neurophysiol.* 111.12 (June 2014), pp. 2603–2613.

[135] R A Pearce and W O Friesen. “Intersegmental coordination of the leech swimming rhythm. I. Roles of cycle period gradient and coupling strength”. en. In: *J. Neurophysiol.* 54.6 (Dec. 1985), pp. 1444–1459.

[136] I Lampl and Y Yarom. “Subthreshold oscillations and resonant behavior: two manifestations of the same mechanism”. en. In: *Neuroscience* 78.2 (May 1997), pp. 325–341.

[137] B L Sabatini and W G Regehr. “Timing of neurotransmission at fast synapses in the mammalian brain”. en. In: *Nature* 384.6605 (Nov. 1996), pp. 170–172.

[138] H A Swadlow, S G Waxman, and T G Weyand. “Effects of variations in temperature on impulse conduction along nonmyelinated axons in the mammalian brain”. en. In: *Exp. Neurol.* 71.2 (Feb. 1981), pp. 383–389.

[139] Abigail S Kalmbach and Jack Waters. “Brain surface temperature under a craniotomy”. en. In: *J. Neurophysiol.* 108.11 (Dec. 2012), pp. 3138–3146.

[140] S M Thompson, L M Masukawa, and D A Prince. “Temperature dependence of intrinsic membrane properties and synaptic potentials in hippocampal CA1 neurons in vitro”. en. In: *J. Neurosci.* 5.3 (Mar. 1985), pp. 817–824.

[141] E Moser, I Mathiesen, and P Andersen. “Association between brain temperature and dentate field potentials in exploring and swimming rats”. en. In: *Science* 259.5099 (Feb. 1993), pp. 1324–1326.

[142] M Volgushev, T R Vidyasagar, M Chistiakova, T Yousef, and U T Eysel. “Membrane properties and spike generation in rat visual cortical cells during reversible cooling”. en. In: *J. Physiol.* 522 Pt 1.Pt 1 (Jan. 2000), pp. 59–76.

[143] A E Dolbear. “The Cricket as a Thermometer”. In: *Am. Nat.* 31.371 (Nov. 1897), pp. 970–971.

[144] Ayako Yamaguchi, David Gooler, Amy Herrold, Shailja Patel, and Winnie W Pong. “Temperature-dependent regulation of vocal pattern generator”. en. In: *J. Neurophysiol.* 100.6 (Dec. 2008), pp. 3134–3143.

[145] Anthony Pires and Ronald R Hoy. “Temperature coupling in cricket acoustic communication”. In: *Journal of Comparative Physiology A* 171.1 (Aug. 1992), pp. 79–92.

[146] Arkarup Banerjee, Robert Egger, and Michael A Long. “Using focal cooling to link neural dynamics and behavior”. en. In: *Neuron* 109.16 (Aug. 2021), pp. 2508–2518.

[147] E I Moser and P Andersen. “Conserved spatial learning in cooled rats in spite of slowing of dentate field potentials”. en. In: *J. Neurosci.* 14.7 (July 1994), pp. 4458–4466.

[148] P Andersen and E I Moser. “Brain temperature and hippocampal function”. en. In: *Hippocampus* 5.6 (1995), pp. 491–498.

[149] T F Roberts, A Roy, J F Prather, and R Mooney. “A synaptic basis for auditory–vocal integration in the songbird”. In: *Journal of* (2008).

[150] Natalia Caporale and Yang Dan. “Spike Timing–Dependent Plasticity: A Hebbian Learning Rule”. In: *Annu. Rev. Neurosci.* 31.1 (July 2008), pp. 25–46.

[151] Peter Dayan and Laurence F Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. en. MIT Press, Aug. 2005.

[152] Logan S James, Angela S Wang, Mila Bertolo, and Jon T Sakata. “Learning to pause: Fidelity of and biases in the developmental acquisition of gaps in the communicative signals of a songbird”. en. In: *Dev. Sci.* 26.5 (Sept. 2023), e13382.

[153] Makoto Araki, M M Bandi, and Yoko Yazaki-Sugiyama. “Mind the gap: Neural coding of species identity in birdsong prosody”. en. In: *Science* 354.6317 (Dec. 2016), pp. 1282–1287.

[154] Timothy M Otchy, Steffen B E Wolff, Juliana Y Rhee, Cengiz Pehlevan, Risa Kawai, Alexandre Kempf, Sharon M H Gobes, and Bence P Ölveczky. “Acute off-target effects of neural circuit manipulations”. en. In: *Nature* 528.7582 (Dec. 2015), pp. 358–363.

[155] Tsai-Wen Chen et al. “Ultrasensitive fluorescent proteins for imaging neuronal activity”. en. In: *Nature* 499.7458 (July 2013), pp. 295–300.

[156] Tatsuo S Okubo, Emily L Mackevicius, and Michale S Fee. “In vivo recording of single-unit activity during singing in zebra finches”. en. In: *Cold Spring Harb. Protoc.* 2014.12 (Oct. 2014), pp. 1273–1283.

[157] James J Jun et al. “Fully integrated silicon probes for high-density recording of neural activity”. en. In: *Nature* 551.7679 (Nov. 2017), pp. 232–236.

[158] Nicholas A Steinmetz et al. “Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings”. en. In: *Science* 372.6539 (Apr. 2021).

[159] Cyrille Rossant et al. “Spike sorting for large, dense electrode arrays”. en. In: *Nat. Neurosci.* 19.4 (Apr. 2016), pp. 634–641.

[160] Kenji Mizuseki, Anton Sirota, Eva Pastalkova, and György Buzsáki. “Theta oscillations provide temporal windows for local circuit computation in the entorhinal-hippocampal loop”. en. In: *Neuron* 64.2 (Oct. 2009), pp. 267–280.

[161] Jorrit S Montijn, Koen Seignette, Marcus H Howlett, J Leonie Cazemier, Maarten Kamermans, Christiaan N Levelt, and J Alexander Heimel. “A parameter-free statistical test for neuronal responsiveness”. en. In: *Elife* 10 (Sept. 2021).

[162] Ana Calabrese and Sarah M N Woolley. “Coding principles of the canonical cortical microcircuit in the avian brain”. en. In: *Proc. Natl. Acad. Sci. U. S. A.* 112.11 (Mar. 2015), pp. 3517–3522.

[163] Yarden Cohen, Jun Shen, Dawit Semu, Daniel P Leman, William A Liberti 3rd, L Nathan Perkins, Derek C Liberti, Darrell N Kotton, and Timothy J Gardner. “Hidden neural states underlie canary song syntax”. en. In: *Nature* 582.7813 (June 2020), pp. 539–544.

[164] William A Liberti 3rd, Jeffrey E Markowitz, L Nathan Perkins, Derek C Liberti, Daniel P Leman, Grigori Guitchounts, Tarciso Velho, Darrell N Kotton, Carlos Lois, and Timothy J Gardner. “Unstable neurons underlie a stable learned behavior”. en. In: *Nat. Neurosci.* 19.12 (Dec. 2016), pp. 1665–1671.

[165] Jeffrey E Markowitz, William A Liberti 3rd, Grigori Guitchounts, Tarciso Velho, Carlos Lois, and Timothy J Gardner. “Mesoscopic patterns of neural activity support songbird cortical sequences”. en. In: *PLoS Biol.* 13.6 (June 2015), e1002158.

[166] Joseph R Scherrer, Galen F Lynch, Jie J Zhang, and Michale S Fee. “An optical design enabling lightweight and large field-of-view head-mounted microscopes”. en. In: *Nat. Methods* 20.4 (Apr. 2023), pp. 546–549.

[167] Hakan Inan, Claudia Schmuckermair, Tugce Tasci, Biafra O Ahanonu, Oscar Hernandez, Jérôme Lecoq, Fatih Dinç, Mark J Wagner, Murat A Erdogdu, and Mark J Schnitzer. "Fast and statistically robust cell extraction from large-scale neural calcium imaging datasets". en. Mar. 2021.

[168] Liron Sheintuch, Alon Rubin, Noa Brande-Eilat, Nitzan Geva, Noa Sadeh, Or Pinchasof, and Yaniv Ziv. "Tracking the Same Neurons across Multiple Days in Ca2+ Imaging Data". en. In: *Cell Rep.* 21.4 (Oct. 2017), pp. 1102–1115.

[169] Jesse H Goldberg and Michale S Fee. "A cortical motor nucleus drives the basal ganglia-recipient thalamus in singing birds". en. In: *Nat. Neurosci.* 15.4 (Feb. 2012), pp. 620–627.

[170] Daniel Deitch, Alon Rubin, and Yaniv Ziv. "Representational drift in the mouse visual cortex". en. In: *Curr. Biol.* 31.19 (Oct. 2021), 4327–4339.e6.

[171] Felix W Moll, Devorah Kranz, Ariadna Corredera Asensio, Margot Elmaleh, Lyn A Ackert-Smith, and Michael A Long. "Thalamus drives vocal onsets in the zebra finch courtship song". en. In: *Nature* 616.7955 (Apr. 2023), pp. 132–136.

[172] Alexei L Vyssotski, Anna E Stepien, Georg B Keller, and Richard H R Hahnloser. "A Neural Code That Is Isometric to Vocal Output and Correlates with Its Sensory Consequences". en. In: *PLoS Biol.* 14.10 (Oct. 2016), e2000317.

[173] G E Vates, B M Broome, C V Mello, and F Nottebohm. "Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches". en. In: *J. Comp. Neurol.* 366.4 (Mar. 1996), pp. 613–642.

[174] E S Fortune and D Margoliash. "Parallel pathways and convergence onto HVc and adjacent neostriatum of adult zebra finches (Taeniopygia guttata)". en. In: *J. Comp. Neurol.* 360.3 (Sept. 1995), pp. 413–441.

[175] Sam E Benezra, Rajeevan T Narayanan, Robert Egger, Marcel Oberlaender, and Michael A Long. "Morphological characterization of HVC projection neurons in the zebra finch (Taeniopygia guttata)". en. In: *J. Comp. Neurol.* 526.10 (July 2018), pp. 1673–1689.

[176] Yoonseob Lim, Ryan Lagoy, Barbara G Shinn-Cunningham, and Timothy J Gardner. "Transformation of temporal sequences in the zebra finch auditory system". en. In: *Elife* 5 (Nov. 2016).

[177] Gesa Feenders, Miriam Liedvogel, Miriam Rivas, Manuela Zapka, Haruhito Horita, Erina Hara, Kazuhiro Wada, Henrik Mouritsen, and Erich D Jarvis. "Molecular mapping of movement-associated areas in the avian brain: a motor theory for vocal learning origin". en. In: *PLoS One* 3.3 (Mar. 2008), e1768.

[178] D Marr. "A theory of cerebellar cortex". en. In: *J. Physiol.* 202.2 (June 1969), pp. 437–470.

[179] N Alex Cayco-Gajic and R Angus Silver. "Re-evaluating Circuit Mechanisms Underlying Pattern Separation". en. In: *Neuron* 101.4 (Feb. 2019), pp. 584–602.

[180] Glenn C Turner, Maxim Bazhenov, and Gilles Laurent. "Olfactory representations by Drosophila mushroom body neurons". en. In: *J. Neurophysiol.* 99.2 (Feb. 2008), pp. 734–746.

[181] Kyle S Honegger, Robert A A Campbell, and Glenn C Turner. "Cellular-resolution population imaging reveals robust sparse coding in the Drosophila mushroom body". en. In: *J. Neurosci.* 31.33 (Aug. 2011), pp. 11772–11785.

[182] Sophie J C Caron, Vanessa Ruta, L F Abbott, and Richard Axel. "Random convergence of olfactory inputs in the Drosophila mushroom body". en. In: *Nature* 497.7447 (May 2013), pp. 113–117.

[183] Javier Perez-Orive, Ofer Mazor, Glenn Turner, Stijn Cassenaer, Rachel I Wilson, and Gilles Laurent. “Oscillations and Sparsening of Odor Representations in the”. In: *Science* ().

[184] Ashok Litwin-Kumar, Kameron Decker Harris, Richard Axel, Haim Sompolinsky, and L F Abbott. “Optimal Degrees of Synaptic Connectivity”. en. In: *Neuron* 93.5 (Mar. 2017), 1153–1164.e7.

[185] Peter Y Wang, Yi Sun, Richard Axel, L F Abbott, and Guangyu Robert Yang. “Evolving the olfactory system with machine learning”. en. In: *Neuron* 109.23 (Dec. 2021), 3879–3892.e5.

[186] Andrew C Lin, Alexei M Bygrave, Alix de Calignon, Tzumin Lee, and Gero Miesenböck. “Sparse, decorrelated odor coding in the mushroom body enhances learned odor discrimination”. en. In: *Nat. Neurosci.* 17.4 (Apr. 2014), pp. 559–568.

[187] Maria Ahmed et al. “Input density tunes Kenyon cell sensory responses in the Drosophila mushroom body”. en. In: *Curr. Biol.* 33.13 (July 2023), 2742–2760.e12.

[188] Jill Soha. “The auditory template hypothesis: a review and comparative perspective”. In: *Anim. Behav.* 124 (Feb. 2017), pp. 247–254.

[189] Masakazu Konishi. “Auditory Environment and Vocal Development in Birds”. In: *Perception and Experience*. Ed. by Richard D Walk and Herbert L Pick. Boston, MA: Springer US, 1978, pp. 105–118.

[190] Anthony Leonardo. “Degenerate coding in neural systems”. en. In: *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 191.11 (Nov. 2005), pp. 995–1010.

[191] Susan Carey and Elsa Bartlett. “Acquiring a Single New Word”. en. In: (Aug. 1978).

[192] Ellen M Markman. *Categorization and Naming in Children: Problems of Induction*. en. MIT Press, 1989.

[193] Fei Xu and Joshua B Tenenbaum. “Word learning as Bayesian inference”. en. In: *Psychol. Rev.* 114.2 (Apr. 2007), pp. 245–272.

[194] Jordan M Moore and Sarah M N Woolley. “Emergent tuning for learned vocalizations in auditory cortex”. en. In: *Nat. Neurosci.* 22.9 (Sept. 2019), pp. 1469–1476.

[195] Noopur Amin, Michael Gastpar, and Frédéric E Theunissen. “Selective and efficient neural coding of communication signals depends on early acoustic and social environment”. en. In: *PLoS One* 8.4 (Apr. 2013), e61417.

[196] Vanessa C Miller-Sims and Sarah W Bottjer. “Development of neural responsivity to vocal sounds in higher level auditory cortex of songbirds”. en. In: *J. Neurophysiol.* 112.1 (July 2014), pp. 81–94.

[197] Jelena Katic, Yuichi Morohashi, and Yoko Yazaki-Sugiyama. “Neural circuit for social authentication in song learning”. en. In: *Nat. Commun.* 13.1 (Aug. 2022), p. 4442.

[198] Mimi L Phan, Carolyn L Pytte, and David S Vicario. “Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds”. en. In: *Proc. Natl. Acad. Sci. U. S. A.* 103.4 (Jan. 2006), pp. 1088–1093.

[199] Shin Yanagihara and Yoko Yazaki-Sugiyama. “Auditory experience-dependent cortical circuit shaping for memory formation in bird song learning”. en. In: *Nat. Commun.* 7 (June 2016), p. 11946.

[200] Jerry A Fodor. *The Modularity of Mind*. en. MIT Press, Apr. 1983.

[201] P K Kuhl, K A Williams, F Lacerda, K N Stevens, and B Lindblom. “Linguistic experience alters phonetic perception in infants by 6 months of age”. en. In: *Science* 255.5044 (Jan. 1992), pp. 606–608.

[202] J S Johnson and E L Newport. “Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language”. en. In: *Cogn. Psychol.* 21.1 (Jan. 1989), pp. 60–99.

[203] Janet F Werker and Richard C Tees. “Cross-language speech perception: Evidence for perceptual reorganization during the first year of life”. In: *Infant Behav. Dev.* 7.1 (Jan. 1984), pp. 49–63.

[204] P Belin, R J Zatorre, P Lafaille, P Ahad, and B Pike. “Voice-selective areas in human auditory cortex”. en. In: *Nature* 403.6767 (Jan. 2000), pp. 309–312.

[205] R Näätänen et al. “Language-specific phoneme representations revealed by electric and magnetic brain responses”. en. In: *Nature* 385.6615 (Jan. 1997), pp. 432–434.

[206] Nima Mesgarani, Connie Cheung, Keith Johnson, and Edward F Chang. “Phonetic feature encoding in human superior temporal gyrus”. en. In: *Science* 343.6174 (Feb. 2014), pp. 1006–1010.

[207] L I Zhang, S Bao, and M M Merzenich. “Persistent and specific influences of early acoustic environments on primary auditory cortex”. en. In: *Nat. Neurosci.* 4.11 (Nov. 2001), pp. 1123–1130.

[208] Edward F Chang and Michael M Merzenich. “Environmental noise retards auditory cortical development”. en. In: *Science* 300.5618 (Apr. 2003), pp. 498–502.

[209] Anne E Takesian, Luke J Bogart, Jeff W Lichtman, and Takao K Hensch. “Inhibitory circuit gating of auditory critical-period plasticity”. en. In: *Nat. Neurosci.* 21.2 (Feb. 2018), pp. 218–227.

[210] Robert C Froemke, Michael M Merzenich, and Christoph E Schreiner. “A synaptic memory trace for cortical receptive field plasticity”. en. In: *Nature* 450.7168 (Nov. 2007), pp. 425–429.

[211] Jennifer K Schiavo, Silvana Valtcheva, Chloe J Bair-Marshall, Soomin C Song, Kathleen A Martin, and Robert C Froemke. “Innate and plastic mechanisms for maternal behaviour in auditory cortex”. en. In: *Nature* 587.7834 (Nov. 2020), pp. 426–431.

[212] Johannes C Dahmen and Andrew J King. “Learning to hear: plasticity of auditory cortical processing”. en. In: *Curr. Opin. Neurobiol.* 17.4 (Aug. 2007), pp. 456–464.

[213] Ann B Butler, Anton Reiner, and Harvey J Karten. “Evolution of the amniote pallium and the origins of mammalian neocortex”. en. In: *Ann. N. Y. Acad. Sci.* 1225 (Apr. 2011), pp. 14–27.

[214] Yuan Wang, Agnieszka Brzozowska-Prechtl, and Harvey J Karten. “Laminar and columnar auditory cortex in avian brain”. en. In: *Proc. Natl. Acad. Sci. U. S. A.* 107.28 (July 2010), pp. 12676–12681.

[215] Sarah C Woolley and Sarah M N Woolley. “Integrating Form and Function in the Songbird Auditory Forebrain”. In: *The Neuroethology of Birdsong*. Ed. by Jon T Sakata, Sarah C Woolley, Richard R Fay, and Arthur N Popper. Cham: Springer International Publishing, 2020, pp. 127–155.

[216] Sidney R Lehky, Terrence J Sejnowski, and Robert Desimone. “Selectivity and sparseness in the responses of striate complex cells”. en. In: *Vision Res.* 45.1 (Jan. 2005), pp. 57–73.

[217] Jesse H Goldberg and Michale S Fee. “Singing-related neural activity distinguishes four classes of putative striatal neurons in the songbird basal ganglia”. en. In: *J. Neurophysiol.* 103.4 (Apr. 2010), pp. 2002–2014.

[218] Mimi H Kao, Brian D Wright, and Allison J Doupe. “Neurons in a forebrain nucleus required for vocal plasticity rapidly switch between precise firing and variable bursting depending on social context”. en. In: *J. Neurosci.* 28.49 (Dec. 2008), pp. 13232–13247.

[219] Jessica A Cardin and Marc F Schmidt. “Auditory responses in multiple sensorimotor song system nuclei are co-modulated by behavioral state”. en. In: *J. Neurophysiol.* 91.5 (May 2004), pp. 2148–2163.

[220] D Margoliash. “Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow”. en. In: *J. Neurosci.* 3.5 (May 1983), pp. 1039–1057.

[221] D Margoliash. “Preference for autogenous song by auditory neurons in a song system nucleus of the white-crowned sparrow”. en. In: *J. Neurosci.* 6.6 (June 1986), pp. 1643–1661.

[222] Joshua T Trachtenberg. “Competition, inhibition, and critical periods of cortical plasticity”. en. In: *Curr. Opin. Neurobiol.* 35 (Dec. 2015), pp. 44–48.

[223] Sandra J Kuhlman, Nicholas D Olivas, Elaine Tring, Taruna Ikrar, Xiangmin Xu, and Joshua T Trachtenberg. “A disinhibitory microcircuit initiates critical-period plasticity in the visual cortex”. en. In: *Nature* 501.7468 (Sept. 2013), pp. 543–546.

[224] Tommaso Pizzorusso, Paolo Medini, Nicoletta Berardi, Sabrina Chierzi, James W Fawcett, and Lamberto Maffei. “Reactivation of ocular dominance plasticity in the adult visual cortex”. en. In: *Science* 298.5596 (Nov. 2002), pp. 1248–1251.

[225] Nadine Gogolla, Pico Caroni, Andreas Lüthi, and Cyril Herry. “Perineuronal nets protect fear memories from erasure”. en. In: *Science* 325.5945 (Sept. 2009), pp. 1258–1261.

[226] Timothy S Balmer, Vanessa M Carels, Jillian L Frisch, and Teresa A Nick. “Modulation of perineuronal nets and parvalbumin with developmental song learning”. en. In: *J. Neurosci.* 29.41 (Oct. 2009), pp. 12878–12885.

[227] Gilles Cornez, Shelley Valle, Ednei Barros Dos Santos, Ioana Chiver, Wendt Müller, Gregory F Ball, Charlotte A Cornil, and Jacques Balthazart. "Perineuronal nets in HVC and plasticity in male canary song". en. In: *PLoS One* 16.8 (Aug. 2021), e0252560.

[228] C V Mello, G E Vates, S Okuhata, and F Nottebohm. "Descending auditory pathways in the adult male zebra finch (Taeniopygia guttata)". en. In: *J. Comp. Neurol.* 395.2 (June 1998), pp. 137–160.

[229] Jörgen Kornfeld, Sam E Benezra, Rajeevan T Narayanan, Fabian Svara, Robert Egger, Marcel Oberlaender, Winfried Denk, and Michael A Long. "EM connectomics reveals axonal target variation in a sequence-generating network". en. In: *Elife* 6 (Mar. 2017).

[230] Daniel E Wilson, David E Whitney, Benjamin Scholl, and David Fitzpatrick. "Orientation selectivity and the functional clustering of synaptic inputs in primary visual cortex". en. In: *Nat. Neurosci.* 19.8 (Aug. 2016), pp. 1003–1009.

[231] Niansheng Ju, Yang Li, Fang Liu, Hongfei Jiang, Stephen L Macknik, Susana Martinez-Conde, and Shiming Tang. "Spatiotemporal functional organization of excitatory synaptic inputs onto macaque V1 neurons". en. In: *Nat. Commun.* 11.1 (Feb. 2020), p. 697.

[232] Michael D Adoff, Jason R Climer, Heydar Davoudi, Jonathan S Marvin, Loren L Looger, and Daniel A Dombeck. "The functional organization of excitatory synaptic input to place cells". en. In: *Nat. Commun.* 12.1 (June 2021), p. 3558.

[233] Rongwen Lu et al. "Video-rate volumetric functional imaging of the brain at synaptic resolution". en. In: *Nat. Neurosci.* 20.4 (Apr. 2017), pp. 620–628.

[234] Eve Armstrong and Henry D I Abarbanel. "Model of the songbird nucleus HVC as a network of central pattern generators". en. In: *J. Neurophysiol.* 116.5 (Nov. 2016), pp. 2405–2419.

[235] Jonathan Cannon, Nancy Kopell, Timothy Gardner, and Jeffrey Markowitz. “Neural Sequence Generation Using Spatiotemporal Patterns of Inhibition”. en. In: *PLoS Comput. Biol.* 11.11 (Nov. 2015), e1004581.

[236] Rich Pang, Alison Duffy, David Bell, Zsofia Torok, and Adrienne Fairhall. “Precision motor timing via scalar input fluctuations”. en. May 2022.

[237] Mikhail I Rabinovich, Pablo Varona, Allen I Selverston, and Henry D I Abarbanel. “Dynamical principles in neuroscience”. In: *Rev. Mod. Phys.* 78.4 (Nov. 2006), pp. 1213–1265.

[238] Christopher M Bishop. *Pattern Recognition and Machine Learning*. en. Springer, Aug. 2006.

[239] Daniel Margoliash and Marc F Schmidt. “Sleep, off-line processing, and vocal learning”. en. In: *Brain Lang.* 115.1 (Oct. 2010), pp. 45–58.

[240] M F Schmidt and M Konishi. “Gating of auditory responses in the vocal control system of awake songbirds”. en. In: *Nat. Neurosci.* 1.6 (Oct. 1998), pp. 513–518.

[241] Sébastien Derégnaucourt, Partha P Mitra, Olga Fehér, Carolyn Pytte, and Ofer Tchernichovski. “How sleep affects the developmental learning of bird song”. en. In: *Nature* 433.7027 (Feb. 2005), pp. 710–716.

[242] Shane R Crandall, Murtaza Adam, Amanda K Kinnischtzke, and Teresa A Nick. “HVC neural sleep activity increases with development and parallels nightly changes in song behavior”. en. In: *J. Neurophysiol.* 98.1 (July 2007), pp. 232–240.

[243] Richard H R Hahnloser, Alexay A Kozhevnikov, and Michale S Fee. “Sleep-related neural activity in a premotor and a basal-ganglia pathway of the songbird”. en. In: *J. Neurophysiol.* 96.2 (Aug. 2006), pp. 794–812.

[244] Richard H R Hahnloser and Michale S Fee. “Sleep-related spike bursts in HVC are driven by the nucleus interface of the nidopallium”. en. In: *J. Neurophysiol.* 97.1 (Jan. 2007), pp. 423–435.

[245] Nathaniel B Sawtell. “Neural Mechanisms for Predicting the Sensory Consequences of Behavior: Insights from Electrosensory Systems”. en. In: *Annu. Rev. Physiol.* 79 (Feb. 2017), pp. 381–399.

[246] Daisuke Hattori, Yoshinori Aso, Kurtis J Swartz, Gerald M Rubin, L F Abbott, and Richard Axel. “Representations of Novelty and Familiarity in a Mushroom Body Compartment”. en. In: *Cell* 169.5 (May 2017), 956–969.e17.

[247] Toshihiko Hosoya, Stephen A Baccus, and Markus Meister. “Dynamic predictive coding by the retina”. en. In: *Nature* 436.7047 (July 2005), pp. 71–77.

[248] David M Schneider, Anders Nelson, and Richard Mooney. “A synaptic and circuit basis for corollary discharge in the auditory cortex”. en. In: *Nature* 513.7517 (Sept. 2014), pp. 189–194.

[249] David M Schneider, Janani Sundararajan, and Richard Mooney. “A cortical filter that learns to suppress the acoustic consequences of movement”. en. In: *Nature* 561.7723 (Sept. 2018), pp. 391–395.

[250] Anders Nelson, David M Schneider, Jun Takatoh, Katsuyasu Sakurai, Fan Wang, and Richard Mooney. “A circuit for motor cortical modulation of auditory cortical activity”. en. In: *J. Neurosci.* 33.36 (Sept. 2013), pp. 14342–14353.

[251] Shobhit Singla, Conor Dempsey, Richard Warren, Armen G Enikolopov, and Nathaniel B Sawtell. “A cerebellum-like circuit in the auditory system cancels responses to self-generated sounds”. In: *Nat. Neurosci.* 20.7 (2017), pp. 943–950.

[252] N Suga and T Shimozawa. “Site of neural attenuation of responses to self-vocalized sounds in echolocating bats”. en. In: *Science* 183.130 (Mar. 1974), pp. 1211–1213.

[253] James F A Poulet and Berthold Hedwig. “A corollary discharge maintains auditory sensitivity during sound production”. en. In: *Nature* 418.6900 (Aug. 2002), pp. 872–876.

[254] Trinity B Crapse and Marc A Sommer. “Corollary discharge across the animal kingdom”. en. In: *Nat. Rev. Neurosci.* 9.8 (Aug. 2008), pp. 587–600.

[255] Emily L Mackevicius, Michael T L Happ, and Michale S Fee. “An avian cortical circuit for chunking tutor song syllables into simple vocal-motor units”. en. In: *Nat. Commun.* 11.1 (Oct. 2020), p. 5029.

[256] Theda H Heinks-Maldonado, Daniel H Mathalon, Max Gray, and Judith M Ford. “Fine-tuning of auditory cortex during speech production”. en. In: *Psychophysiology* 42.2 (Mar. 2005), pp. 180–190.

[257] Steven J Eliades and Xiaoqin Wang. “Neural substrates of vocalization feedback monitoring in primate auditory cortex”. en. In: *Nature* 453.7198 (June 2008), pp. 1102–1106.

[258] A H Seidl. “Regulation of conduction time along axons”. en. In: *Neuroscience* 276 (Sept. 2014), pp. 126–134.

[259] C E Carr and M Konishi. “A circuit for detection of interaural time differences in the brain stem of the barn owl”. en. In: *J. Neurosci.* 10.10 (Oct. 1990), pp. 3227–3246.

[260] L A Jeffress. “A place theory of sound localization”. en. In: *J. Comp. Physiol. Psychol.* 41.1 (Feb. 1948), pp. 35–39.

[261] Abigail A Russo et al. “Motor Cortex Embeds Muscle-like Commands in an Untangled Population Response”. en. In: *Neuron* 97.4 (Feb. 2018), 953–966.e8.

[262] Abigail A Russo, Ramin Khajeh, Sean R Bittner, Sean M Perkins, John P Cunningham, L F Abbott, and Mark M Churchland. “Neural Trajectories in the Supplementary Motor Area and Motor Cortex Exhibit Distinct Geometries, Compatible with Different Classes of Computation”. en. In: *Neuron* 107.4 (Aug. 2020), 745–758.e6.

[263] Tom Michael Mitchell. “The need for biases in learning generalizations”. In: *Technical Report CBM-TR-117* (1980).

[264] Eric B Baum. *What is Thought?* en. MIT Press, 2004.

[265] S Gershman. *What Makes Us Smart*. Princeton University Press, Oct. 2021.

[266] Jean-Nicolas Audet, Mélanie Couture, and Erich D Jarvis. "Songbird species that display more-complex vocal learning are better problem-solvers and have larger brains". en. In: *Science* 381.6663 (Sept. 2023), pp. 1170–1175.

[267] Clare Walton, Eben Pariser, and Fernando Nottebohm. "The zebra finch paradox: song is little changed, but number of neurons doubles". en. In: *J. Neurosci.* 32.3 (Jan. 2012), pp. 761–774.

[268] T A Nick and M Konishi. "Dynamic control of auditory activity during sleep: correlation between song response and EEG". en. In: *Proc. Natl. Acad. Sci. U. S. A.* 98.24 (Nov. 2001), pp. 14012–14016.

[269] Melissa J Coleman, Arani Roy, J Martin Wild, and Richard Mooney. "Thalamic gating of auditory responses in telencephalic song control nuclei". en. In: *J. Neurosci.* 27.37 (Sept. 2007), pp. 10024–10036.

[270] Stephen D Shea and Daniel Margoliash. "Basal forebrain cholinergic modulation of auditory activity in the zebra finch song system". en. In: *Neuron* 40.6 (Dec. 2003), pp. 1213–1226.

[271] R Mooney. "Different subthreshold mechanisms underlie song selectivity in identified HVc neurons of the zebra finch". en. In: *J. Neurosci.* 20.14 (July 2000), pp. 5420–5436.

[272] Alan Bush, Juan F Döppler, Franz Goller, and Gabriel B Mindlin. "Syringeal EMGs and synthetic stimuli reveal a switch-like activation of the songbird's vocal motor program". In: *Proceedings of the National Academy of Sciences* 115.33 (2018), pp. 8436–8441.

[273] Andreas Christian Kotowicz. “Characterization of auditory responses recorded in the caudal medial nidopallium in the freely behaving, awake juvenile zebra finch”. PhD thesis. ETH Zurich, 2014.

[274] M M Solis and A J Doupe. “Contributions of tutor and bird’s own song experience to neural selectivity in the songbird anterior forebrain”. en. In: *J. Neurosci.* 19.11 (June 1999), pp. 4559–4584.

[275] C J Price, R J Wise, E A Warburton, C J Moore, D Howard, K Patterson, R S Frackowiak, and K J Friston. “Hearing and saying. The functional neuro-anatomy of auditory word processing”. en. In: *Brain* 119 (Pt 3) (June 1996), pp. 919–931.

[276] Ghislaine Dehaene-Lambertz, Lucie Hertz-Pannier, and Jessica Dubois. “Nature and nurture in language acquisition: anatomical and functional brain-imaging studies in infants”. en. In: *Trends Neurosci.* 29.7 (July 2006), pp. 367–373.

[277] J Cynx, H Williams, and F Nottebohm. “Hemispheric differences in avian song discrimination”. en. In: *Proc. Natl. Acad. Sci. U. S. A.* 89.4 (Feb. 1992), pp. 1372–1375.

[278] Sanne Moorman, Sharon M H Gobes, Maaike Kuijpers, Amber Kerkhofs, Matthijs A Zandbergen, and Johan J Bolhuis. “Human-like brain hemispheric dominance in birdsong learning”. en. In: *Proc. Natl. Acad. Sci. U. S. A.* 109.31 (July 2012), pp. 12782–12787.

[279] Sanne Moorman, Sharon M H Gobes, Ferdinand C van de Kamp, Matthijs A Zandbergen, and Johan J Bolhuis. “Learning-related brain hemispheric dominance in sleeping songbirds”. en. In: *Sci. Rep.* 5 (Mar. 2015), p. 9041.

[280] H Williams, L A Crane, T K Hale, M A Esposito, and F Nottebohm. “Right-side dominance for song control in the zebra finch”. en. In: *J. Neurobiol.* 23.8 (Oct. 1992), pp. 1006–1020.

[281] Claude Z H Wang, Joshua A Herbst, Georg B Keller, and Richard H R Hahnloser. "Rapid interhemispheric switching during vocal production in a songbird". en. In: *PLoS Biol.* 6.10 (Oct. 2008), e250.

[282] Massimo Trusel, Danyal Alam, and Todd F Roberts. "Synaptic Connectivity of Afferent Inputs onto Projection Neurons in the Songbird HVC". en. Nov. 2022.

[283] Bradley M Colquitt, Devin P Merullo, Genevieve Konopka, Todd F Roberts, and Michael S Brainard. "Cellular transcriptomics reveals evolutionary identities of songbird vocal circuits". en. In: *Science* 371.6530 (Feb. 2021).

[284] E D Jarvis and F Nottebohm. "Motor-driven gene expression". en. In: *Proc. Natl. Acad. Sci. U. S. A.* 94.8 (Apr. 1997), pp. 4097–4102.

[285] Erich D Jarvis et al. "Global view of the functional molecular organization of the avian cerebrum: mirror images and functional columns". en. In: *J. Comp. Neurol.* 521.16 (Nov. 2013), pp. 3614–3665.

[286] F Duarte Ortiz and R Mooney. "Imaging Auditory Cortical Neuron Activity in the Singing Songbird". In: 2022.

[287] D B Kelley and F Nottebohm. "Projections of a telencephalic auditory nucleus-field L-in the canary". en. In: *J. Comp. Neurol.* 183.3 (Feb. 1979), pp. 455–469.

[288] Edmund A Arbas and Ronald L Calabrese. "Rate modification in the heartbeat central pattern generator of the medicinal leech". In: *Journal of Comparative Physiology A* 155.6 (Nov. 1984), pp. 783–794.

[289] Sarah C Woolley, Raghav Rajan, Mati Joshua, and Allison J Doupe. "Emergence of context-dependent variability across a basal ganglia network". en. In: *Neuron* 82.1 (Apr. 2014), pp. 208–223.

[290] Lei Xiao, Gaurav Chattree, Francisco Garcia Oscos, Mou Cao, Matthew J Wanat, and Todd F Roberts. "A Basal Ganglia Circuit Sufficient to Guide Birdsong Learning". en. In: *Neuron* 98.1 (Apr. 2018), 208–221.e5.

[291] Shanglin Zhou, Sotiris C Masmanidis, and Dean V Buonomano. “Neural Sequences as an Optimal Dynamical Regime for the Readout of Time”. en. In: *Neuron* 108.4 (Nov. 2020), 651–658.e5.

[292] Ana Amador, Yonatan Sanz Perl, Gabriel B Mindlin, and Daniel Margoliash. “Elemental gesture dynamics are encoded by song premotor cortical neurons”. en. In: *Nature* 495.7439 (Mar. 2013), pp. 59–64.

[293] Ann Kennedy, Greg Wayne, Patrick Kaifosh, Karina Alviña, L F Abbott, and Nathaniel B Sawtell. “A temporal basis for predicting the sensory consequences of motor commands in an electric fish”. en. In: *Nat. Neurosci.* 17.3 (Mar. 2014), pp. 416–422.

[294] Mu Zhou, Feixue Liang, Xiaorui R Xiong, Lu Li, Haifu Li, Zhongju Xiao, Huizhong W Tao, and Li I Zhang. “Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex”. en. In: *Nat. Neurosci.* 17.6 (June 2014), pp. 841–850.

[295] Richard J Gardner, Erik Hermansen, Marius Pachitariu, Yoram Burak, Nils A Baas, Benjamin A Dunn, May-Britt Moser, and Edvard I Moser. “Toroidal topology of population activity in grid cells”. en. In: *Nature* 602.7895 (Feb. 2022), pp. 123–128.

[296] Steffen Schneider, Jin Hwa Lee, and Mackenzie Weygandt Mathis. “Learnable latent embeddings for joint behavioural and neural analysis”. en. In: *Nature* 617.7960 (May 2023), pp. 360–368.

[297] Rishidev Chaudhuri, Berk Gerçek, Biraj Pandey, Adrien Peyrache, and Ila Fiete. “The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep”. en. In: *Nat. Neurosci.* 22.9 (Sept. 2019), pp. 1512–1520.

[298] Daniel Margoliash. “Sleep, learning, and birdsong”. en. In: *ILAR J.* 51.4 (2010), pp. 378–386.

[299] Emily L Mackevicius, Shijie Gu, Natalia I Denisenko, and Michale S Fee. "Self-organization of songbird neural sequences during social isolation". en. Feb. 2022.

[300] Shijie Gu, Emily L Mackevicius, Michale S Fee, and Pengcheng Zhou. "Spatial Tracking Across Time (STAT): Tracking Neurons Across In-Vivo Imaging Sessions through Optimizing Local Neighborhood Motion Consistency". en. May 2023.

[301] I Winkler, T Kujala, H Tiitinen, P Sivonen, P Alku, A Lehtokoski, I Czigler, V Csépe, R J Ilmoniemi, and R Näätänen. "Brain responses reveal the learning of foreign language phonemes". en. In: *Psychophysiology* 36.5 (Sept. 1999), pp. 638–642.

[302] Dmitriy Aronov, Aaron S Andalman, and Michale S Fee. "A specialized forebrain circuit for vocal babbling in the juvenile songbird". en. In: *Science* 320.5876 (May 2008), pp. 630–634.

[303] Sarita S Shaevitz and Frédéric E Theunissen. "Functional connectivity between auditory areas field L and CLM and song system nucleus HVC in anesthetized zebra finches". en. In: *J. Neurophysiol.* 98.5 (Nov. 2007), pp. 2747–2764.

[304] Xiaoxuan Jia, Joshua H Siegle, Corbett Bennett, Samuel D Gale, Daniel J Denman, Christof Koch, and Shawn R Olsen. "High-density extracellular probes reveal dendritic backpropagation and facilitate neuron classification". en. In: *J. Neurophysiol.* 121.5 (May 2019), pp. 1831–1847.

[305] Patrice Adret, C Daniel Meliza, and Daniel Margoliash. "Song tutoring in presinging zebra finch juveniles biases a small population of higher-order song-selective neurons toward the tutor song". en. In: *J. Neurophysiol.* 108.7 (Oct. 2012), pp. 1977–1987.

[306] Erik Rybakken, Nils Baas, and Benjamin Dunn. "Decoding of Neural Data Using Cohomological Feature Extraction". en. In: *Neural Comput.* 31.1 (Jan. 2019), pp. 68–93.

[307] Husain H Danish, Dmitriy Aronov, and Michale S Fee. “Rhythmic syllable-related activity in a songbird motor thalamic nucleus necessary for learned vocalizations”. en. In: *PLoS One* 12.6 (June 2017), e0169568.

[308] E T Vu, M E Mazurek, and Y C Kuo. “Identification of a forebrain motor programming network for the learned song of zebra finches”. en. In: *J. Neurosci.* 14.11 Pt 2 (Nov. 1994), pp. 6924–6934.

[309] Niangui Wang, Patrick Hurley, Carolyn Pytte, and John R Kirn. “Vocal control neuron incorporation decreases with age in the adult zebra finch”. en. In: *J. Neurosci.* 22.24 (Dec. 2002), pp. 10864–10870.

[310] Dmitriy Aronov, Lena Veit, Jesse H Goldberg, and Michale S Fee. “Two distinct modes of forebrain circuit dynamics underlie temporal patterning in the vocalizations of young songbirds”. en. In: *J. Neurosci.* 31.45 (Nov. 2011), pp. 16353–16368.

[311] Richard Mooney and Jonathan F Prather. “The HVC microcircuit: the synaptic basis for interactions between song motor and vocal plasticity pathways”. en. In: *J. Neurosci.* 25.8 (Feb. 2005), pp. 1952–1964.

[312] A Hanuschkin, S Ganguli, and R H R Hahnloser. “A Hebbian learning rule gives rise to mirror neurons and links them to control theoretic inverse models”. en. In: *Front. Neural Circuits* 7 (June 2013), p. 106.

[313] Richard H R Hahnloser, Surya Ganguli, Rodrigo Quian Quiroga, and Stefano Panzeri. “Vocal learning with inverse models”. en. In: *Principles of Neural Coding*. Ed. by Rodrigo Quian Quiroga and Stefano Panzeri. Principles of Neural Coding. CRC Press: CRC Press, 2013, pp. 547–564.

[314] Jonathan Garst-Orozco, Baktash Babadi, and Bence P Ölveczky. “A neural circuit mechanism for regulating vocal variability during song learning in zebra finches”. en. In: *Elife* 3 (Dec. 2014), e03697.

[315] Emily L Mackevicius, Andrew H Bahle, Alex H Williams, Shijie Gu, Natalia I Denisenko, Mark S Goldman, and Michale S Fee. "Unsupervised discovery of temporal sequences in high-dimensional datasets, with applications to neuroscience". en. In: *Elife* 8 (Feb. 2019).

[316] Emily L Mackevicius, Shijie Gu, Natalia I Denisenko, and Michale S Fee. "Self-organization of songbird neural sequences during social isolation". en. In: *Elife* 12 (May 2023).

[317] Emily L Mackevicius, Shijie Gu, Natalia I Denisenko, and Michale S Fee. "Self-organization of songbird neural sequences during social isolation". en. In: *Elife* 12 (May 2023).

[318] Emily L Mackevicius, Shijie Gu, Natalia I Denisenko, and Michale S Fee. "Self-organization of songbird neural sequences during social isolation". en. In: *Elife* 12 (May 2023).

[319] Emily L Mackevicius, Shijie Gu, Natalia I Denisenko, and Michale S Fee. "Self-organization of songbird neural sequences during social isolation". en. In: *Elife* 12 (May 2023).

[320] Emily L Mackevicius, Shijie Gu, Natalia I Denisenko, and Michale S Fee. "Self-organization of songbird neural sequences during social isolation". en. In: *Elife* 12 (May 2023).

[321] Emily L Mackevicius, Shijie Gu, Natalia I Denisenko, and Michale S Fee. "Self-organization of songbird neural sequences during social isolation". en. In: *Elife* 12 (May 2023).

[322] Mattia Chini, Thomas Pfeffer, and Ileana Hanganu-Opatz. "An increase of inhibition drives the developmental decorrelation of neural activity". en. In: *Elife* 11 (Aug. 2022).

[323] Alex H Williams, Anthony Degleris, Yixin Wang, and Scott W Linderman. “Point process models for sequence detection in high-dimensional neural spike trains”. en. In: *Adv. Neural Inf. Process. Syst.* 33 (Dec. 2020), pp. 14350–14361.

[324] Chethan Pandarinath, Daniel J O’Shea, Jasmine Collins, Rafal Jozefowicz, Sergey D Stavisky, Jonathan C Kao, Eric M Trautmann, Matthew T Kaufman, Stephen I Ryu, Leigh R Hochberg, et al. “Inferring single-trial neural population dynamics using sequential auto-encoders”. In: *Nat. Methods* 15.10 (2018), pp. 805–815.