Neural circuits for socially guided vocal learning

Linda Bistere

Graduate School of
Systemic Neurosciences
LMU Munich

Dissertation der Graduate School of Systemic
Neurosciences der Ludwig-Maximilians-Universität München

April, 2024

Supervisor:
Dr. Daniela Vallentin
Max-Planck-Institut für Biologische Intelligenz
Eberhard-Gwinner-Str. 11, 82319 Seewiesen, Starnberg

Second Reviewer:
Prof. Dr. Manfred Gahr
Max-Planck-Institut für Biologische Intelligenz
Eberhard-Gwinner-Str. 4, 82319 Seewiesen, Starnberg

Third Reviewer:
Prof. Dr. Steffen Hage
Eberhard-Karls-Universität Tübingen
Elfriede-Aulhorn-Str. 5, 72076 Tübingen

Date of Submission: April 25th, 2024
Date of Defense: October 7th, 2024

CONTENTS

1 INTRODUCTION

Language is one of the most prominent examples of vocal learning. Understanding the written words of this paragraph requires the acquired skill of language. This process begins early in human life, by hearing parents or caregivers talk and gradually learning the language by interaction and imitation. Initially, the pronunciation of words is not accurate, but over time, with a lot of practice, the pronunciation improves as do the communication skills.

The acquisition of the first language occurs during the first few years of life. During this period, language learning takes place as a natural process through interactions with parents and other caregivers. Acquiring a second language later in life is a much more tedious process, often involving language lessons and the challenge of memorizing vocabulary, leading to the gradual improvement in language comprehension and conversation with others. Although both processes are vocal learning, the acquisition of the first language early in life takes place during a critical period. Once this period has passed, learning any language becomes considerably more challenging[1]. Thus, social interactions with parents and caregivers during the time of critical period are crucial, as they enable and facilitate the vocal learning process in humans.

Vocal learning is not limited to humans. Among other species, whales, harbor seals and songbirds also learn vocalizations[2]. In this thesis I am exploring the impact of social interactions on vocal learning in the zebra finch, a gregarious songbird. Zebra finches live in large colonies and communicate with each other by using vocalizations thousands of times every day. One type of such vocalizations is the courtship song. In zebra finches, only male birds sing and young juvenile birds learn their song from their fathers or other adult males (tutors) during the critical period. In contrast, female birds do not sing, but they produce short, innate vocalizations known as “calls”. Although much of the research in vocal learning of zebra finches has focused on how the juveniles learn song from a male tutor bird[3–12], the contribution of social interactions between females and juveniles has so far been largely neglected. A few studies have provided evidence, that social interactions with females may influence song learning. For example, displaying a video of a female bird visual display during juvenile song practice led to juveniles learning a better copy of the tutor song, than if juveniles were exposed to the same video outside of the singing context[13]. Furthermore, pairing a blindfolded male juvenile with a female sibling during the song learning phase led to a more accurate copy of learned song to tutor song than if the female sibling was absent[14]. How social interactions between juvenile and an adult female may influence song learning has so far remained unknown.

In the first chapter, I investigate the role of social interactions between juveniles and female birds on song learning by training juvenile birds to copy a tutor song playback. I analyze the influence of female presence on juvenile vocal learning by comparing the similarity of learned song to tutor song playback depending on whether the female was present or absent during the song learning phase. To investigate, whether female birds are providing juveniles with vocal feedback during song practice, I track both female and juvenile vocalizations during the song learning phase. I perform intracellular recordings in the song production pathway of male juvenile birds while presenting female calls during juveniles' song, to explore whether female vocalizations have any impact on the song learning circuitry in juveniles.

In the second chapter, I investigate the temporal dynamics of song crystallization. At the onset of the song learning phase, juvenile birds produce a subsong with highly variable features[15]. As the learning phase progresses, subsong transforms into a plastic song with syllables (elements of song) that are repeated across the song renditions[7,15]. Although the song at this point has less variability than the subsong, the spectral and temporal features of the syllables are continuously refined[7]. The neural dynamics during song production have been largely described in adult birds[16–18] and in juveniles during the early stages of song learning[19–21], but how song crystallization is reflected in the song system remains unknown. I describe the change in temporal dynamics of song and the neural activity in the song production pathway during the crystallization phase.

1.1 BEHAVIORAL BASIS FOR VOCAL LEARNING

Similar to humans using language as a form of vocal communication, other species within the animal kingdom use vocalizations to communicate too. Vocalizations serve numerous purposes, including warnings of approaching predators[22], expressions of distress[23], and the formation of social bonds[24]. In many species, the vocalizations employed are largely innate. However, a few species learn their vocalizations by imitation, including cetaceans, bats, parrots, hummingbirds and songbirds (Oscine passerines)[25,26]. This form of vocal production learning is defined as modifying individuals own vocalizations by exposure to vocalizations from other individuals[27].

1.1.1 Zebra finch as a model organism for vocal learning

One of the most popular songbirds used to study vocal learning is the zebra finch (*Taeniopygia guttata*). This is attributed to the highly stereotyped song of zebra finch males and the short critical period (approximately 90 days) for song learning. Apart from easily traceable song throughout development, the advantage for studying zebra finches in a lab setting arises from their breeding habits. Zebra finches are opportunistic breeders, that will readily produce offspring throughout the year, facilitating the studying of vocal learning. During the song learning phase, the song changes from initially highly variable vocalizations to stereotyped renditions of the same syllables[7]. The development of song and its acoustic features can be easily traced by recording the vocalizations from male juveniles in sound-attenuated boxes. Young juvenile males need to learn their song from other adult males to court a female once they reach adulthood[15]. Throughout their life, zebra finch males learn and sing only one song[15]. Each repetition of the song is very stereotyped[15], which facilitates studying their singing behavior from a behavioral and neural perspective. Additionally, contrary to other animals like mice and rats that need to be trained to perform a specific skilled motor task, young zebra finches will readily learn a song from conspecifics without any external reinforcement (i.e. food or water reward)[28]. Even in the absence of an appropriate song model, the juvenile birds will develop a song that they will sing as adults[29]. Lastly, learning a song and singing to court a female falls within the behavioral repertoire also observed in zebra finches in the wild[15]. Therefore, zebra finches are excellent candidates for studying vocal learning and vocal production behavior from a neuroethological perspective.

1.1.2 Vocalizations in zebra finches

In zebra finches, only males sing while females choose their mates based on their song[30]. Zebra finch song is a 500 – 1200 ms long vocalization (a song motif) produced by males (Figure 1A) [31]. Right before the onset of the song (first motif), males produce introductory notes[32]. Each following motif consists of multiple syllables and during one singing period males repeat the motif several times (song bout). The song of each adult male is unique, yet highly repetitive across renditions[15]. Adult male birds sing several hundred times every day[33]. Some of the songs are directly addressed to a female (directed song) whereas others do not have an addressee and are used for song maintenance in adulthood (undirected song)[34]. The directed song produced during courtship is highly stereotyped, while the undirected song is slightly more variable[34].

Apart from song, both male and female zebra finches also use short vocalizations known as “calls” to communicate with each other[15,35]. As zebra finches live in large flocks of hundreds of individuals in their native Australia, calls are often used to establish a social bond within the flock[36]. There are several different call types, including a frequently used stack call and a tet call (Figure 1B). Some of the other types of calls include distance calls to locate a conspecific (Figure 1B), whine calls and nest calls related to pair bond, as well as aggressive calls like the wsst call[35]. While the song is learned, most of the calls, apart from the distance call in males[37], are innate and readily used by young birds. Additionally, calling behavior in zebra finches is not random but occurs at a certain latency to avoid overlap with other conspecifics[38]. The rich vocal repertoire and its specificity of use reflects the complexity of social behavior within the flock of zebra finches.

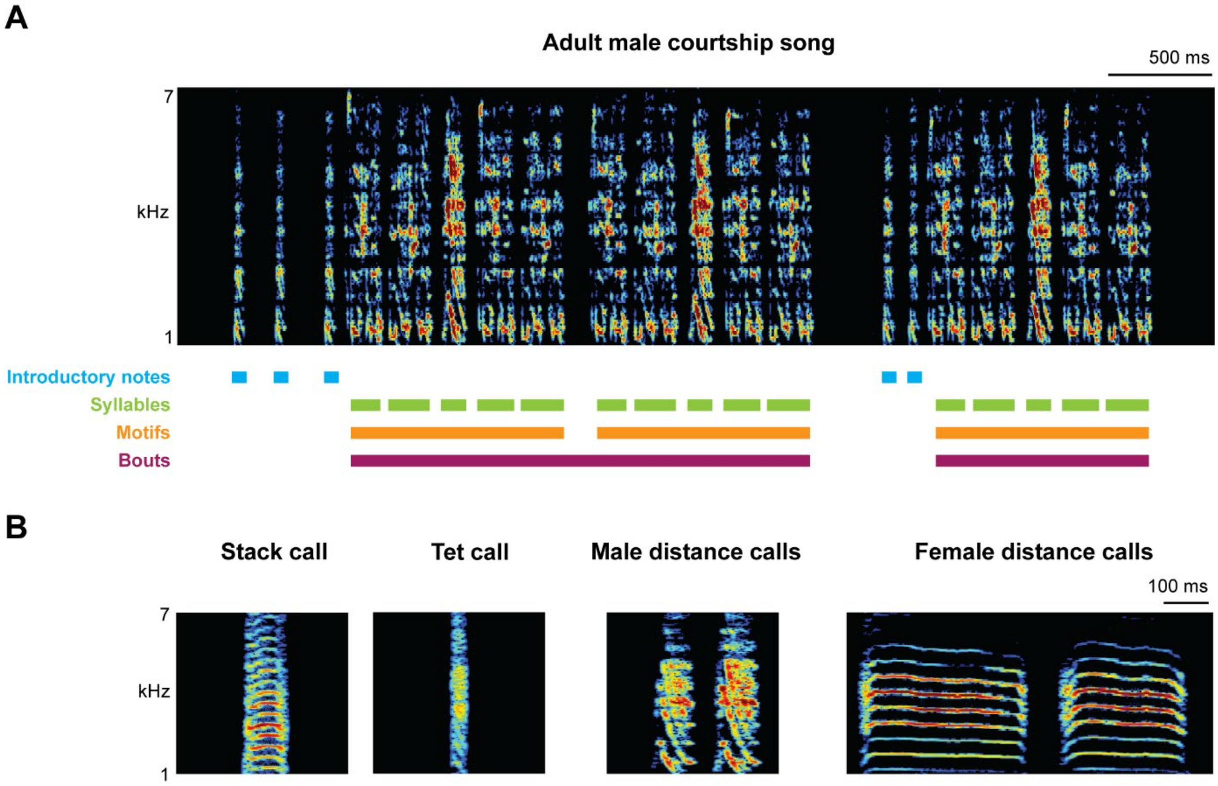


Figure 1: Vocalizations in zebra finches. A) A spectrogram of zebra finch courtship song produced by an adult male. Colorful bars below highlight each part of the song with its respective terminology. B) Examples of three widely used calls in zebra finches. While stack and tet calls are similar in females and males, male distance calls exhibit a higher degree of frequency modulation than female calls.

1.1.3 Vocal learning in zebra finches

To attract a female as adults, young juvenile males learn a song by imitating the song of a father or a tutor bird[5] (Figure 2A,B). The song learning process occurs during a critical period during the first 90 days of the juveniles' life[9]. The first vocalizations produced by young males during song practice resemble babbling and have a very variable acoustic structure[7]. These vocalizations are necessary for tutor song memorization[39]. As the vocal learning phase progresses, the song can be recognized by consecutive early syllables (Figure 2A). The early syllables are often similar to each other acoustically and the juvenile bird is repeating them during song production[7]. With more practice, these syllables undergo local refinement of the acoustic structure and acoustically distinct syllables emerge. When challenged with prioritizing either first learning the correct syllable sequence in song or spectral features, zebra finch juveniles will master the spectral features of their song first[40]. After ~65 days post hatch, the tempo of the produced song increases, which is attributed to shorter gap durations between syllables[41]. The durations of the syllables are not affected by this change[41]. During this time of the learning phase, the juvenile birds produce more songs per day than during the early learning phase or adulthood[33]. Towards the end of the learning phase, the song is crystalized as the acoustic properties of each syllable are refined by the juvenile bird locally[42] and the song becomes increasingly stereotyped, until a final version of the adult song is achieved[7] (Figure 2 A,B). Once the song is learned, it will remain the same over the rest of the birds' life with little variation.

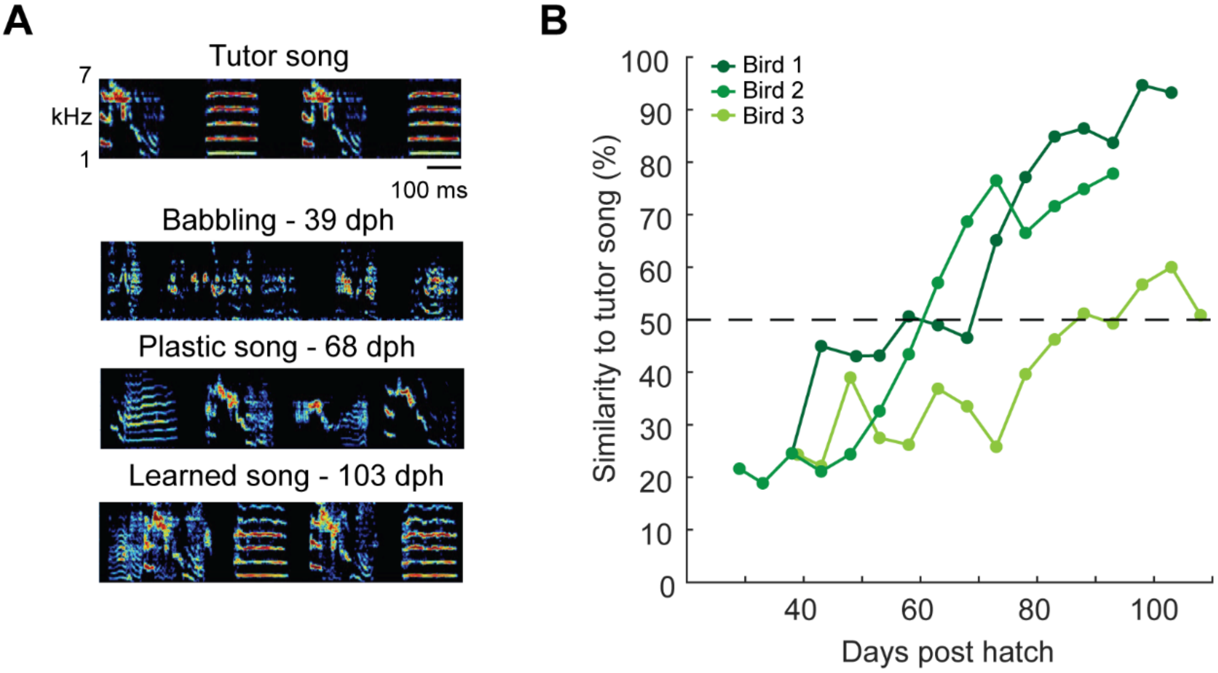


Figure 2: Song learning in zebra finches. A) Example of song learning from Bird 1, top: tutor song, below: three examples of birds own song from Bird 1 during the song learning phase with the respective age. B) Song learning trajectories of three birds during the song learning phase.

1.1.4 Constraints of vocal learning in zebra finches

How flexible are zebra finches in learning their song? Several vocal learners, like canaries and parrots, are capable of learning new vocalizations throughout their life and are therefore called open-ended learners. In contrast, zebra finches are close-ended learners. This means, that zebra finches only learn song during a critical period (approximately first 90 days of life). If juvenile males are not exposed to tutor song during this time, they will develop a song with atypical features (long syllable durations)[5]. These songs are considered less attractive by female birds and lead to a lower mating success[29]. An interesting study was conceived by tracking how the song of the offspring of such song-isolated birds develops over several generations[43]. The study found, that after three to four generations, the song of the offspring had species-typical features. This experiment shows, that there is a genetic bias to learn species-specific song. This finding is further supported when analyzing the learned song of cross-fostered zebra finches. In the study by Araki et al. (2016), juvenile zebra finches were cross-fostered with Bengalese finch tutors, whose song is acoustically different from the zebra finch song. Although the acoustic features of the learned song resembled that of the Bengalese tutor, the gap duration remained zebra-finch specific[44], suggesting a genetic bias in song timing. Indeed, even when zebra finch juveniles were tutored with variable gap durations, the learned song consisted of stereotyped gap durations across song motifs[45]. To summarize, although zebra finches are able to learn song from genetically unrelated birds with different song than the genetic fathers', the learned song will portray species-specific song features and it must be learned during the critical period.

1.1.5 Social aspects of vocal learning in zebra finches

The learning success (how similar the learned song is to a tutor song) also depends on the social situation during song learning. Juvenile zebra finches will copy the tutor song most accurately, if tutored by a conspecific adult male bird[3]. It is also possible to tutor zebra finches in social isolation with a playback of an adult male's song[3]. In this case, juvenile zebra finches will learn a more similar song copy to tutor song in an operant paradigm, where juveniles can elicit the playback of tutor song by pecking a key, than if they are exposed to passive playbacks of the tutor song[3]. The number of playbacks per day should be optimized to increase the goodness of a copy, as too many playbacks lead to poorer learning results[46].

When juveniles are exposed to several adult birds, they tend to learn from the adult with whom they have the most social interactions with[6,8,10]. The most accurate copy of song is achieved, if juveniles can see and interact with the tutor bird. Preventing juvenile birds from visual or visual and auditory contact to the tutor leads to fewer syllables copied from the tutor song[47]. In this scenario, juvenile zebra finches will copy the song that they have heard from their father before reaching 35 days post hatch. Indeed, the memory of the tutor song to learn from can form before reaching nutritional independence[11]. The extent to which the juveniles copy tutor song is also influenced by the number of male siblings in the clutch – the more male siblings within the same clutch, the poorer the song copy[12]. The juveniles, that are surrounded with more male siblings copy fewer elements from their fathers' song leading to the production of shorter song motifs. This observation indicates that both the social and auditory experience play a role in song learning[12]. Siblings that learn the song earlier than other siblings in a clutch are less affected than later learners[12]. Interestingly, the early learners were typically among the younger siblings of the clutch[12].

Another study investigated the role of the company of a female sibling during the song learning phase[14]. Juvenile males were blindfolded and paired with a live tutor during the song learning phase[14]. In one of the groups, the male juvenile was additionally accompanied by a young female sibling[14]. Analysis of the learned song indicated, that male juveniles, that were accompanied with the female sibling copied more from the tutor song than males where the female sibling was absent[14]. Because the male juveniles were blindfolded, these results suggest that social interactions (auditory and physical contact) between the male and the female sibling were sufficient to improve song learning[14]. In contrast, a more recent study investigated the influence of a visual female display ('fluff-up'), that is commonly used to signal song preference[48], on song learning in juveniles[13]. During part of the song learning phase, one group of male juveniles was presented with a video of a female visual display during song practice, while the other group was exposed to the same video in a non-contingent manner[13]. After the learning phase, the learned song from the juveniles was compared to the tutor song - juveniles receiving a contingent visual display learned a more accurate copy to the tutor song than juveniles, where the visual display appeared randomly[13]. This study highlights, that visual feedback from a female bird can also support juvenile song learning.

Although both of these studies investigated female influence on song learning, they do not address the question of how an adult female might support the song learning of a juvenile in the entire repertoire of social interactions. In the first study, juvenile males could interact with the female sibling but they had no visual contact[14], whereas in the more recent study the possible female reinforcement was reduced to a video of a female visual display, without the juvenile interacting with an actual female[13]. To bridge this gap, I first examine whether female company promotes song learning by pairing juveniles with adult females with unrestricted access to social interactions. To investigate, whether the vocal interactions between a juvenile and a female could provide an insight on how females may aid the song learning process, I track the vocalizations from females and juveniles and analyze their co-occurrence throughout the song learning phase.

1.1.6 The role of females in song learning

Since vocal learning is one of the main aspects studied in zebra finches and female zebra finches do not sing, their role in song learning remains understudied. This sexual dimorphism is thought to be based on differences in the brain structure of males and females. While male brains exhibit a well described song system, female brains share some of the brain areas involved in song production in males, but these are smaller in size[49–51]. Despite this differentiation, female birds are capable of distinguishing subtle acoustic features of adult male songs[52]. For example, when exposed to a directed song and an undirected song of the same male, female birds will show preference towards the directed song[52]. The preference can be assessed either by observing which speaker (directed versus undirected song playback) the female spends more time at or by observing an increase in female's calling behavior as a response to a song playback[53]. Additionally, female birds are able to memorize a male song that they were exposed to during the early development[30,54–56] and to distinguish preferable features of the song in a gradient fashion – the more parts of a song resemble a females' fathers' song, the more females will prefer it[55]. How female song preferences may influence juvenile song learning process remains an interesting question.

1.1.6.1 Auditory feedback during song production

During singing, the auditory feedback of the produced sounds is crucial for song maintenance in adult birds. If adult male birds are deafened, the song they produce gradually deteriorates[57,58]. In juvenile birds deafening leads to a more rapid deterioration of song than in adults[58], likely due to less song practice at the time of deafening and the decreased amount of song plasticity in adults. Distorting the auditory feedback by exposing adult birds to playback of their own song in a delayed manner respective to the song production also leads to a deterioration of song[59], once more highlighting the importance of intact auditory feedback of the produced song. However, targeted distorted auditory perturbations can also lead to subtle acoustic changes in song even in adult male zebra finches with crystalized song. In this case, adult birds that are exposed to a white noise blast targeted at a particular syllable are able to adapt the pitch of that syllable to escape the aversive white noise stimulus[60–62]. In juvenile birds, such auditory feedback leads to changes in the acoustic structure of song syllables, stuttering and repetitions of introductory notes[63]. On the contrary side, song performance could also be perturbed by a female producing calls – a positive reinforcement signal in adults[64,65]. How female calls may affect juvenile song learning and the underlying neural circuitry for song production remains unknown.

1.2 BRAIN CIRCUITRY FOR VOCAL LEARNING IN SONGBIRDS

Vocalizations in songbirds are tightly controlled by the underlying brain circuitry – the song system. The brain areas of songbirds are organized in separate nuclei, that are connected to each other by projection neurons. The first attempts of exploring which brain areas are involved in singing were made by lesioning forebrain areas and observing whether the singing behavior changed, while staining experiments revealed the connectivity of these brain areas[66–68]. For example, bilaterally lesioning HVC left canaries unable to produce learned songs, yet these birds would still assume singing posture[66]. These early findings highlighted the involvement of brain area HVC (proper name) in song production but its function during singing was yet to be explored (Figure 3).

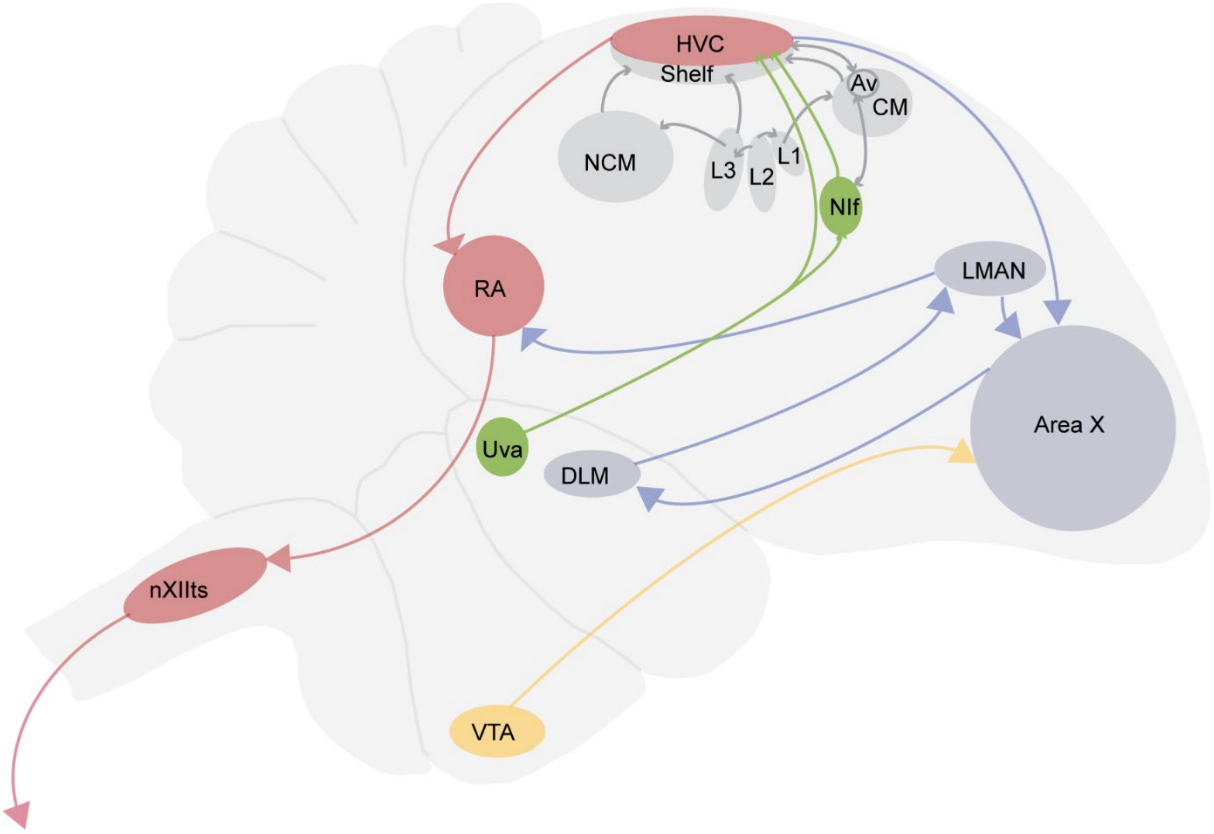


Figure 3: Neural circuitry for song learning and production. In red: song production pathway, comprised of HVC, RA and nXIIts; in grey: auditory pathway with projections to HVC; in green: thalamic nucleus Uva and sensorimotor nucleus NIf; in blue: anterior forebrain pathway, consisting of HVC, Area X, DLM and LMAN; in yellow: dopaminergic ventral tegmental area (VTA) projects to Area X in the anterior forebrain pathway.

1.2.1 Two pathways of the song network

On a large scale, the song system consists of two pathways involved in song learning, production and maintenance – the song output pathway and the anterior forebrain pathway. The song output pathway consists of the premotor area HVC, robust nucleus of arcopallium (RA) with projections to dorsal medial nucleus (DM) and hypoglossal motoneurons (nXIIts), that ultimately innervate the syringeal muscles[69] (Figure 3). The anterior forebrain pathway consists of HVC, Area X of the medial striatum (Area X), dorsal lateral nucleus of medial thalamus (DLM) and lateral magnocellular nucleus of the anterior nidopallium (LMAN) (Figure 3).

1.2.2 Song output pathway

When an adult bird sings, premotor neurons in HVC produce a sparse bursting activity that spans throughout the song motif[16,17,70], while inhibitory interneurons control the precision of firing of HVC projection neurons[71]. The current view in the birdsong research field, is that by producing such sparse bursts at unique timepoints during song[16], HVC encodes the temporal sequence of song[72]. Within HVC, there are four different neuron types – premotor HVC_{RA} projection neurons, $HVC_{Area\ X}$ projection neurons, that project to the anterior forebrain pathway, HVC_{Av} projection neurons that send vocal motor signals to the auditory areas[73] and inhibitory interneurons. While HVC_{RA} neurons send information about the timing of the song to the downstream motor nucleus RA, the anterior forebrain pathway receives timing information about the ongoing song from $HVC_{Area\ X}$ neurons.

Downstream from HVC is the motor nucleus RA, which orchestrates the acoustic features of syllables[70]. During song production, neurons in RA exhibit precise bursts that are associated with subsyllabic features of song[70]. These bursts are likely initiated by the a precise bursting pattern of HVC_{RA} neurons[16], however inhibitory network within RA from surrounding interneurons could also contribute to the bursting activity[74]. Stimulating RA during song production leads to a distortion of the acoustic features of the syllables, without affecting the timing of song[75]. RA further projects to respiratory areas in the brainstem[76] and hypoglossal motoneurons (nXIIts)[69] that directly innervate syrinx.

1.2.3 Anterior forebrain pathway

An efference copy of HVC_{RA} projection neurons is also sent to Area X by $HVC_{Area\ X}$ projection neurons and passes through the anterior forebrain pathway loop ultimately reconnecting back to RA (Figure 3). During song production, neural activity in both LMAN and Area X increases[77,78]. Since Area X receives information about song timing from HVC, dopaminergic input from ventral tegmental area (VTA) and information about LMAN activity that is involved in producing variability during song production[79], it has been hypothesized that Area X then biases song production towards successful renditions[80]. Lesions of Area X do not affect singing in adult birds, indicating, that it is not necessary for the motor production of song[81,82].

Neurons downstream from Area X in LMAN change their firing properties during singing depending on the social context[83] – when an adult male performs to a female and produces a stereotyped song, neurons in LMAN exhibit firing characteristics that are tightly linked to song production[84]. When an adult male is singing a more variable, undirected song by himself, spiking activity of neurons in LMAN becomes more variable on a trial to trial basis[84]. These observations highlight, that neural activity in LMAN is connected with the degree of variability in song features. In line with this idea, lesions to LMAN lead to a loss of variability of acoustic features during production of undirected song, making it as stereotyped as directed song[85]. Lesioning LMAN in deafened adult birds protects the song structure from the deterioration associated with deafening[86], indicating that anterior forebrain pathway is involved in matching auditory feedback during singing with variations in song production.

Taken together, these findings highlight that anterior forebrain pathway is not necessary for motor production of song, but is involved in song learning and song maintenance in adult birds.

1.2.4 Auditory pathway

Auditory inputs to the song system are mainly provided by a projection from the sensorimotor interfacial nucleus of the nidopallium (NIf) to HVC[87–89] (Figure 3). NIf in turn receives projections from the claudal mesopallium (CM), which further receives input from Field L (areas L1, L2 and L3) that also projects to caudal medial nidopallium (NCM). Areas NCM, CM and Field L also project to HVC Shelf – an auditory area ventral to HVC[90]. Within this network of auditory areas, the neural responses to auditory stimuli vary between areas – a population of neurons in NCM show responses to conspecific songs despite induced background noise, suggesting a mechanism for recognizing socially relevant auditory stimuli[91,92], while neurons in L2 and L1 of the Field L are less selective to auditory stimuli with the selectivity gradually increased in L3 and NCM[93].

Most of the auditory information to HVC is relayed through NIf, that receives a major auditory input from a subset of the auditory area CM – the Avalanche[94], and an upstream thalamic nucleus uvaeformis (Uva). Neurons in Uva also project to HVC and exhibit increased activity at the onset of song syllables during singing, potentially driving each syllable within the song sequence[95], whereas the neurons projecting to Avalanche from HVC are necessary for song learning and song plasticity in adulthood[73]. Since Avalanche has reciprocal connections with both HVC and NIf, and NIf also projects to HVC, these areas are considered to be the sensorimotor loop that connects auditory pathway to the song system[96].

Activity in NIf and HVC is highly correlated, as auditory responses in NIf closely precede auditory responses in HVC[88,97]. While HVC exhibits replay of neural activity during sleep, the initiation of this activity occurs in NIf[98]. In contrast, inactivation of NIf greatly decreases auditory responses in HVC[88], but weak auditory responses in HVC remain likely due to a comparatively small auditory input to HVC provided by Avalanche[94,99]. Neurons in NIf show auditory activity and vocal-motor activity that increases preceding vocalizations and remains suppressed after vocalizations in adults[100]. In juveniles, NIf exhibits onset responses during babbling, and some neurons are specific to a particular syllable identity[101]. Furthermore, NIf also shows auditory responses to tutor song in listening juveniles[101] and to conspecific calls in listening adults[102], suggesting that NIf may relay socially relevant auditory information to the song system. Song learning and production are affected by the temporary inactivation of NIf[103,104]. However, bilateral lesions of NIf do not affect song maintenance which relies on the auditory feedback, suggesting that song maintenance may also be supported by auditory inputs from CM to HVC[105].

Taken together, these findings suggest that the song system is receiving auditory input that is socially relevant and crucial for song learning from the auditory pathway.

1.2.5 Neural activity in the song system during song learning

The sensorimotor nucleus HVC is involved in song production[16,66] and song learning[104,106]. During song learning the chain-like bursting pattern of HVC develops, as it is not present at the onset of the song learning phase[19,107]. At this time, the vocal babbling phase is orchestrated by LMAN[20] projecting to RA, since the projections between HVC and RA have not been established yet[108]. LMAN controls the production of subsong, which is highly variable in its spectral and temporal features[109]. The temporal variability of song is additionally accompanied by irregular breathing patterns in singing juveniles[110,111]. During the course of song learning as the song becomes more stereotyped, the main input to motor area RA gradually switches from LMAN to HVC[21,110]. The projection neurons in HVC produce rhythmic bursts during singing, as the juveniles produce acoustically similar, rhythmic ~100 ms long syllables (protosyllables)[19]. As the plastic song phase begins, syllables become acoustically diverse and bursting activity in HVC begins to differentiate towards bursts during the production of specific syllables[7,19]. The song exhibits variable gaps at this point, that gradually become shorter in duration to match the adult-like gaps[41]. During the song crystallization phase, the spectral features of each syllable are refined and a chain-like bursting activity of HVC projection neurons throughout the song motif develops. This process is accompanied by increasing inhibition from the HVC interneurons, which increase their firing during the learned syllables and decrease the firing activity during syllables, that still need practice[112]. The inhibitory neurons therefore preserve the precise bursting pattern of the HVC projection neurons[71]. Interestingly, the gradually increasing inhibition also controls auditory responses to tutor song in HVC. During the learning phase, premotor $\mathrm{HVC_{RA}}$ projection neurons will exhibit auditory responses during tutor song playback, that are absent in the adulthood[112]. This finding indicates, that inhibitory interneurons in HVC additionally also control the input to HVC from surrounding auditory areas.

How does the activity of HVC projection neurons form, if juvenile birds are not exposed to tutor song? In an interesting study, juvenile birds were kept from exposure to tutor song until the production of a plastic song. Calcium imaging of HVC revealed, that also in the absence of tutor song HVC projection neurons had already formed short, chain-like neural sequences during song production[107]. After exposure to the tutor song, these sequences became linked to newly learned syllables and integrated during song production[107].

HVC projects to the downstream motor nucleus RA, that exhibits very variable spiking patterns at the onset of vocal development. As the juvenile bird is practicing, the spiking patterns in RA become more precise until the formation of bursts develops[21]. Inactivating the input to RA from LMAN in juveniles leads to a stereotyped, adult-like neural activity in RA before the maturation of song[21], indicating that LMAN is necessary for vocal exploration during song learning. In contrast, inactivation of HVC input to RA during song learning leads to abolished stereotypy in neural activity of RA and song performance[21].

Simultaneously, the song learning process is also largely influenced by the anterior forebrain pathway. While LMAN is introducing variability in each song rendition, Area X evaluates the goodness of the performance and biases the neural dynamics in LMAN to produce the best achieved song version so far. Lesions of LMAN during early development lead to reduced variability in song production[82,109], whereas juvenile birds with lesioned Area X produce highly variable songs with long syllables in adulthood[82]. In another study, lesioning Area X in juveniles did not cease the exploratory behavior, but lesions to DLM increased the stereotypy of produced song[113]. Area X also receives dopaminergic inputs from VTA. Although dopaminergic signaling from VTA has been shown to occur during production of favorable song copy in adults[114], how the dopaminergic signaling might influence the neural dynamics in Area X during song learning in juveniles remains unknown.

1.2.6 Auditory feedback during song production

Adult birds are not able to learn new syllables after the critical period for song learning has passed. The auditory feedback remains important also in adulthood and is used for song maintenance. In deafened adult birds, the learned song gradually deteriorates, as the auditory feedback during song production is missing[57], but this effect can be prevented by lesions to LMAN[86]. Although adult birds are unable to learn new syllables, if exposed to a distorted auditory feedback paradigm with aversive white-noise blasts during singing, adult birds are capable of modifying some acoustic features of already learned song syllables, like pitch[61,62]. Neurons in LMAN do not exhibit auditory responses to distorted auditory feedback[115], but pharmacological inactivation of LMAN reverses the learned adjustment of pitch, suggesting that the basal ganglia loop plays a crucial role in this behavior[61]. As mentioned previously, a core nucleus in the basal ganglia loop, Area X, receives dopaminergic input from VTA. During song perturbations with white noise, the dopaminergic neurons in VTA decrease in activity during song production if the song outcome was worse than predicted (syllable was distorted with white noise) and increase the firing activity if the outcome was better than predicted (catch-trial without white noise)[60]. The modulation of firing activity in dopaminergic neurons in VTA is absent during passive listening of birds own song playback, indicating that the error-signal is performance-related and not purely auditory[60]. Dopamine release in Area X can also be triggered by social feedback during song performance in form of female calls[64]. Apart from auditory feedback, songbirds are able to modify their song based on other feedback modalities[116,117]. For example, adult Bengalese finches are able to modify their song based on the color of a light cue[116]. After initial training using white noise blasts paired with a light cue to encourage song modification, Bengalese finches would adapt their song even in the absence of white noise blasts by relying only on the light cue[116]. In another study, Bengalese finches were given mild cutaneous electric stimulation, that was targeted at a specific song syllable during singing[117]. The adult birds were capable of adjust the pitch of the targeted syllable to escape the electric stimulation[117].

1.2.7 Can we expect auditory responses in HVC?

Since HVC receives projections from auditory areas and at the same time projects to the motor area RA for song production, it is a promising candidate to investigate whether auditory feedback shapes its neural activity during song learning in juveniles. In adults, a study reported that HVC projection neurons do not show a change in activity to tutor song playback or female calls that occurred during song production[65,112]. In contrast, in anesthetized adult songbirds, HVC displays auditory responses to birds own song[118–122]. In anesthetized white-crowned sparrows, neurons in HVC exhibit responses specific to sequences of parts of birds own song, but not when these parts of song were presented separately, when spectral and temporal features of the song were altered or to a conspecific song[118,119]. A similar observation was later made in HVC of anesthetized adult zebra finches – neurons were responsive to birds own song, but not to a reversed birds own song, conspecific song or isolated parts of syllables of the birds own song[120,123]. The auditory responses to birds own song in anesthetized adults were not exclusive to one group of neurons, but rather were relayed by both $HVC_{Area\ X}$ and HVC_{RA} projecting neurons and interneurons[122]. Such auditory responses in adult HVC are state-dependent and absent during wakefulness[97]. In contrast, HVC projection neurons in juveniles show auditory responses to tutor song playback while no auditory responses are exhibited to white noise blasts during singing[63,112]. This developmental switch likely occurs due to an increase of inhibition within HVC, that also gates the input from auditory areas. Additionally, HVC neurons receive dopaminergic projections from the midbrain nucleus periaqueductal grey (PAG). During the exposure to a singing tutor these PAG neurons increase their activity which leads to release of dopamine in HVC, facilitating the encoding of tutor song[4]. This finding suggests, that HVC may not only exhibit auditory responses but also encode socially valent signals. Taken together, these findings suggest that juvenile HVC could be the interface where auditory, socially relevant feedback is implemented to guide song practice.

2 FEMALE VOCAL FEEDBACK PROMOTES SONG LEARNING IN MALE JUVENILE ZEBRA FINCHES

Female vocal feedback promotes song learning in male juvenile zebra finches

Linda Bistere[1,2], Carlos M. Gomez-Guzman[1,2], Yirong Xiong[3], Daniela Vallentin[1*]

[1] Max Planck Institute for Biological Intelligence, 82319 Seewiesen, Germany

2 Graduate School of Systemic Neurosciences, 82152 Planegg, Germany

[3] University of Tübingen, 72076 Tübingen, Germany

Author Contributions: L.B. and D.V. conceived the study and designed the experiments, L.B. and C.M.G.G. conducted the experiments, L.B., Y.X., C.M.G.G. and D.V. analyzed the data, L.B. and D.V. wrote the manuscript, D.V. acquired funding and supervised the project.

* To whom correspondence may be addressed:

Daniela Vallentin
Max Planck Institute for Biological Intelligence
Eberhard-Gwinner-Str.
82319 Seewiesen, Germany
+49 (0) 8157-932-460
daniela.vallentin@bi.mpg.de

Competing Interest Statement: No competing interests.

Keywords: animal behavior, vocal learning, social feedback, songbirds, HVC, electrophysiology

This Word document includes:

Main Text including Figures and Figure Legends

Acknowledgements and Funding Sources

Methods

References

Abstract

Social interactions promote vocal learning but little is known about how social feedback affects the vocal learning process and its underlying neural circuitry. To address this issue, we explored song imitation in juvenile male zebra finches raised either in the presence or absence of females. By tutoring juvenile zebra finches with a controlled tutor song and simultaneously tracking their song learning progress in relation to co-occurring female vocalizations, we found that the company of an adult female leads to a more accurate spectral and temporal copy of the tutor song. In cases when the number of female calls correlated with song practice throughout development, juveniles exhibited a tendency to sing a song more closely resembling the tutor song hinting towards the possibility that females may provide practice-specific feedback. To decipher whether female vocal feedback has an impact on the neural activity within the song learning pathway, we performed intracellular recordings of HVC projection neurons in listening and singing zebra finches. In juvenile zebra finches, we found that female vocalizations can modulate neural activity in HVC during passively listening and singing. In contrast, in singing adult zebra finches female calls do not have an impact on the singing-related neural activity pattern. Interestingly, we found female call-evoked responses outside of the context of singing to persist after development suggesting an age-independent mechanism for the representation of behaviorally relevant vocal feedback. These results highlight the contribution of female vocal feedback to developmental song learning and how vocal input from a conspecific other than the tutor song can influence the neural circuit involved in song learning and production.

Introduction

Vocal learning is a complex skill exhibited by only a small subset of all vocal species[1,2]. Through observation and imitation of a role model, vocal learners practice their vocalizations to achieve a good approximation of the model[3,4]. This learning process is often influenced by the social environment[5–7], including parental care[8,9]. Previous work in songbirds has investigated vocal learning in the presence of an adult male song model (tutor)[10–16] but the extent to which social factors, specifically vocal interactions with a female bird, a companion that is not able to sing[17], influence song development remains largely unknown.

Social feedback during vocal learning in zebra finches

Vocal learning in zebra finches occurs during a developmental critical period – a limited time interval during which a juvenile tutee learns to reproduce the vocalizations of an adult tutor[18,19]. During this period, the physical presence of a tutor compared to a song playback improves the learning outcome[20], which is likely due to the social interaction between tutee and tutor[6,10,14,16]. Whether the physical presence of another conspecific is sufficient to improve song learning cannot be disentangled in the previously described scenario since the male adult zebra finch will also function as the tutor. In contrast, female zebra finches do not sing and would therefore not provide a model courtship song. A previous study has shown that song-contingent visual feedback from female zebra finches to practicing juveniles results in an increase in the spectral similarity of learned song to tutor song[21]. However, not only visual cues can aid juvenile song learning[22], - females also frequently produce short, innate vocalizations (‘calls’)[17,23–25] which can serve as vocal feedback[26]. It has been demonstrated that female birds produce these calls in response to preferred adult songs[27] indicating that female birds are capable of distinguishing different male courtship songs[28–30] and can have preferences for distinct features of these songs[29,31–34]. Thus, females are well equipped to also provide behaviorally relevant vocalizations to developing males. Therefore, by raising juvenile males in the presence or absence of females and simultaneously tutoring them with a controlled song model, we can quantify the impact of female vocalizations on song learning.

HVC is involved in song learning

During the critical period, the song learning and production pathway is in a state of increased plasticity[35–37]. Listening to a tutor song is sufficient to induce activity changes in the premotor nucleus HVC (proper name)[36,38,39], a brain area essential for song production in adults[40,41]. Throughout the song learning period, HVC premotor neurons develop a sparse spiking pattern that ultimately drives the production of the learned stereotyped song[35,36,40]. Tutor song-evoked responses in HVC premotor neurons are then suppressed by HVC inhibitory interneurons in adult zebra finches, once learning is complete[36]. In singing adult males this song-related neural activity remains unaffected when females call[42]. Whether vocal responses from females can influence song outcome by modulating the activity of HVC projection neurons in juvenile birds is not known.

We used behavioral assays to determine if the presence of a non-singing female increases the similarity between a learned song and its model. We recorded the vocalizations from female birds while juveniles were practicing their songs and assessed how these vocal reactions change during the learning phase. Then, to determine whether female vocal input can impact the underlying song learning circuitry, we performed intracellular and extracellular recordings in HVC of awake and listening juvenile and adult birds while presenting playbacks of female calls or playbacks of the bird's own song with overlapping female calls. Last, we also recorded intracellularly in freely moving and singing juvenile birds to decipher whether female calls have an impact on the ongoing activity within this premotor nucleus.

We found that raising juvenile males with non-singing females increased the similarity of learned song to tutor song and was associated with increased female calling as learning progressed. We also found that auditory input from females changed the activity of premotor neurons of HVC. These findings suggest a social feedback mechanism by which a non-singing instructor can guide developmental song learning.

Results

Female presence leads to improved song learning in juvenile male zebra finches

To assess whether the presence of a female zebra finch has an impact on the song learning outcome of juvenile male birds, we raised juveniles in two different social contexts: alone or with a female bird (Figure 1A). In both conditions juvenile birds were presented with the same tutor song playback consisting of two synthetized syllables repeated twice (ABAB tutor song, Figure 1B) which could be elicited by pecking a key. The tutor song playback had a predefined duration and spectral features to which the juvenile song could be compared to (see Methods). Once the birds reached adulthood we quantified the similarity between the learned song and the tutor song (Figure 1C) using a feature-based approach[43]. We found that songs produced by birds raised with an adult female were more similar to the tutor song (median similarity=46.17 %, n=18 birds) than songs produced by birds raised without a female (median similarity=28.44 %, n=15 birds, Wilcoxon rank sum test, p=0.02, Figure 1D).

When analyzing the syllable rate, i.e. the number of syllables (syll) sung per second (excluding the silent gaps) we found that birds raised with a female had a lower syllable rate (with female: median syllable rate=5.22 syll/sec, alone: median syllable rate=6.88 syll/sec, p=0.0029, Wilcoxon rank sum test, Figure 1E). Interestingly, birds copied the syllable rate of the tutor song playback (5.87 syll/sec) when raised with a female (5.44±1.17 syll/sec, p=0.137, one sample t-test) but not when raised alone (7.37±1.91 syll/sec, p=0.0091, one sample t-test). Further, birds raised with a female had longer syllable durations (median syllable duration=191 ms) than birds raised alone (median syllable duration=145 ms, p=0.0029, Wilcoxon rank sum test, Figure 1F). For both groups the syllable duration was different from the tutor song playback (mean syllable duration=170±18 ms, with female: p=0.045, alone: p=0.0117, one sample t-test). In contrast, gap duration was not different between groups (with female: mean gap duration=58±16 ms; alone: mean gap duration=51±18 ms, p=0.1639, Wilcoxon rank sum test, , Figure 1G) but different from the tutor song playback (mean gap duration=121±3 ms, with female: p<0.001, alone: p<0.001, one sample t-test). This result is in line with previous studies, showing that zebra finches have a bias towards species-specific gap durations, which were shorter than the gap duration of the synthesized ABAB tutor song[44,45].

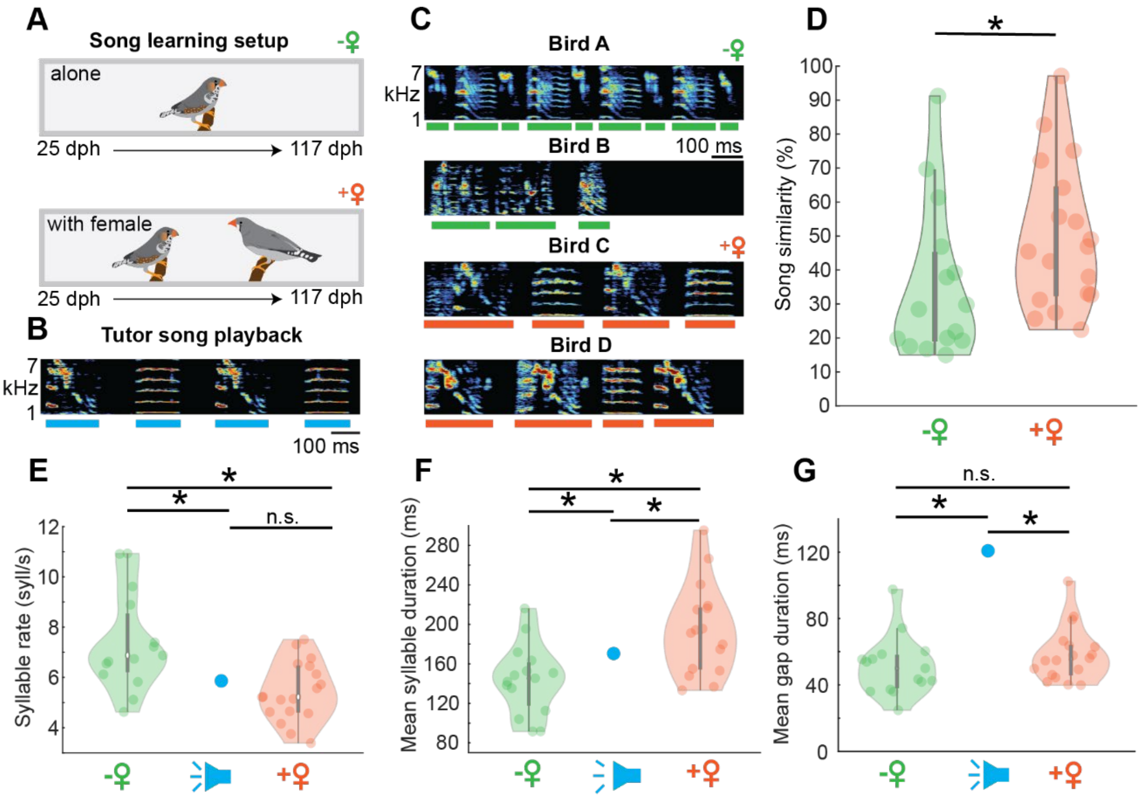


Figure 1: Female presence supports song learning. A) From day 25 to day 117 post hatch (dph) individual juvenile male zebra finches were trained with a tutor song playback in two social conditions – alone (n= 15 juveniles) or with an adult female (n=18 juveniles). B) Sonogram of the tutor song playback that could be elicited by a packing key (n = 40 playbacks/day per bird). Blue lines indicate the syllables. C) Examples of sonograms of learnt song from 4 birds older than 91 dph, green and orange lines indicate syllables. D) Song similarity of birds raised alone (green) or raised with a female (orange). E) Syllable rate of birds raised alone (green), of the tutor song playback (blue) and of birds raised with a female (orange). Asterisk marks significant difference ($p<0.05$); n.s. = not significant. F) Mean syllable durations produced by birds from both groups. Mean syllable duration of tutor song playback in blue. G) Mean gap durations of learnt song from both groups and tutor song playback.

Previous work has demonstrated that the ability to learn specific tutor songs is partly influenced by genetics, and the extent to which the song is inherited depends on how effective the tutoring was[46]. We investigated whether the presence of a female can mitigate genetic biases or influence any father song memory that juveniles may have formed before isolation from their genetic fathers at 30 days post-hatch. To address potential confounding factors arising from similarities between the fathers' songs and the tutor song, we first examined acoustic similarities between these sounds. The analysis revealed that the tutor song exhibited distinct acoustic characteristics different from the fathers' songs (Figure 2A, B). We then compared the learned songs of a subset of juvenile birds with known father identities (n=16 birds) to the songs of their genetic fathers (Figure 2A, n=11 father birds). We observed, that birds raised with a female had a higher song similarity to the tutor song and a better match of the tutor’s song syllable rate compared to the isolated ones (song similarity: p=0.007, syllable rate: p=0.003, linear mixed-effect model, Figure 2A-C) independent of their genetic background or early sensory experience prior to 30 days post hatch.

In summary, our findings suggest that the presence of a female can positively influence song learning in juvenile birds, enhancing similarity to the tutor song and improving the match to the syllable rate of tutor's song, irrespective of genetic background or early sensory experiences. Females might use different sensory modalities to provide behaviorally relevant feedback. It has been previously demonstrated that a specific visual display (fluff ups) by females during juvenile song practice leads to a better singing performance[21] but even blind folded juveniles will produce a more similar copy of the tutor song when a female is present[22]. Additionally, female birds increase their vocalization rate when hearing adult male song[26]. Therefore, we hypothesized that the female’s vocal feedback to the practicing juvenile might also support song learning.

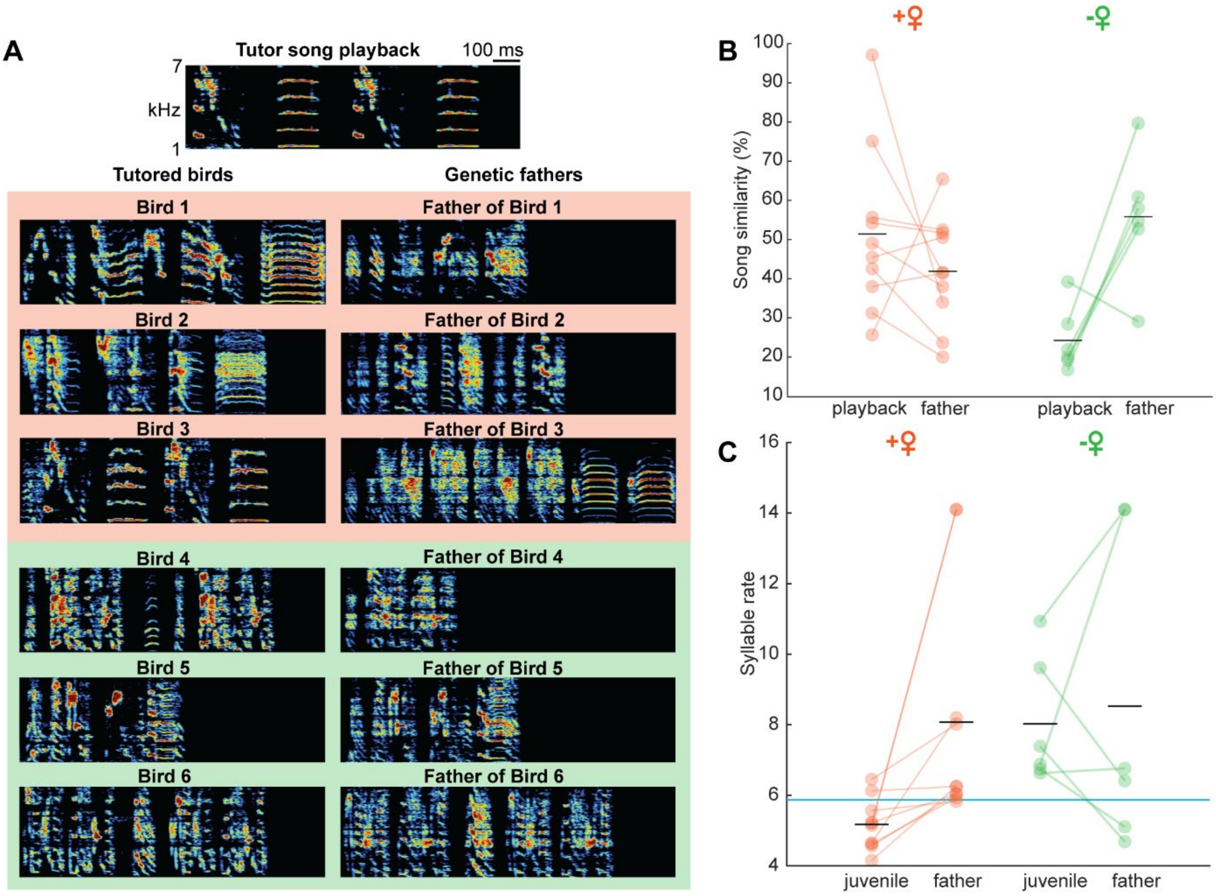


Figure 2: Female presence aids song learning from playback. A) Examples of learnt song from tutored birds (n=16 birds) and their respective fathers (n= 11 birds). *Top*: ABAB tutor song playback. *Below (left)*: Sonograms of example songs produced by Birds 1-3 raised in female presence and birds 4-6 raised alone *Below (right)*: Sonograms of example songs produced by the genetic fathers of bird 1 to 6. B) Song similarity between ABAB playback and the song of the ABAB tutored bird or between the fathers' song and the song of the ABAB tutored bird, from birds raised in female presence (orange) and birds raised alone (green). The tutor song was acoustically distinct from the father's song (similarity(father/tutor) = 31.45 %, similarity (untutored birds/tutor) = 32.02 %, p=0.833, Wilcoxon rank sum test).C) Syllable rate comparison between learnt song and fathers' song from birds raised with a female (orange) and birds raised alone (green). Black dashed line represents an exact match between learnt syllable rate and fathers' syllable rate. Blue line represents the syllable rate of tutor song playback.

Female vocal feedback increases with improved song performance

To explore the vocal behavior of female zebra finches in relation to juvenile song practice we individually raised six juvenile males with four different adult females while also providing tutor song playback. To identify the vocalizations produced by the female and separate them from the juveniles' vocalizations we equipped the females with small telemetric microphones that were attached to their backs, which predominantly captured sounds produced by the females[47]. Simultaneously, we recorded the sounds produced by both animals with a microphone attached to the cage (Figure 3A, B). To detect the juveniles' songs and the female calls, but exclude any other sounds that might have been picked up by both microphones, we trained a neural network[48] with manually assigned vocalizations as a training set (see Methods). We identified 98344 juvenile songs (13572-24712 songs per juvenile) and 18672 female calls (932-8232 female calls per juvenile) which were produced during time windows between the juveniles' song onsets and song offsets. Since the sixth zebra finch did not learn to produce a regular song but only produced harmonic stack calls (Supplementary Fig. 1), this bird was excluded from further analysis.

To understand whether the females changed their calling behavior in relation to song performance, we assessed the number of female calls emitted during song practice. We observed that, for three out of five juveniles, the number of female calls increased with the age of the juvenile bird (juvenile 1-3: $p<0.05$, linear regression model, Figure 3C). Although the female that was initially housed with juvenile 1 later accompanied juvenile 4, we did not observe a correlation between the frequency of female calls and the age of juvenile 4 or juvenile 5, who was raised with a different female (juvenile 4-5, $p>0.05$, linear regression model Figure 3C). Juvenile birds did not produce more songs as the female feedback increased throughout the learning phase (Supplementary. Fig. 2). This observation led us to ask whether the female calls might serve as an indicator for the quality of the final song. Therefore, we assessed the similarity of the learned song to the tutor song playback at the end of learning. We found that in the cases when the female call ratio linearly increased with age, the male birds produced songs with higher similarity to the tutor song (song similarity range: 42.33-73.1%, Figure 3D). Females housed with juvenile 4 and juvenile 5 produced calls that were uncorrelated to juvenile song practice over development (Figure 3C). In these cases, calling did not lead to the production of a similar tutor song (song similarity range: 19.44-29.11% Figure 3D). Although all birds were trained with the same tutor song, only juveniles 1-3 had a higher similarity to tutor song than untutored birds (Figure 3D, similarity of untutored birds ($n = 7$ birds) to ABAB tutor song playback= 32.023 ± 4.44 %). We employed the similarity between untutored birds' songs, along with a 95% confidence interval as a threshold, to categorize birds into good and poor learners based on their ability to replicate the provided tutor song. Scrutinizing this criterion (40.7% similarity threshold) led to the observation that juvenile birds, which received progressively more auditory input from females during their song performance, exhibited better replication of the tutor song compared to juveniles that did not experience an increase in calling by females over time (Figure 3D). When calculating the syllable rate for all five juvenile birds we observed that juveniles 1-3 were closer to the syllable rate of the tutor song playback than juveniles 4 and 5, that received uncorrelated feedback (Figure 3E).

These findings indicate that female birds produce vocalizations in relation to song practice and that this vocal input might lead to a better spectral and temporal copy of the learned song to tutor song.

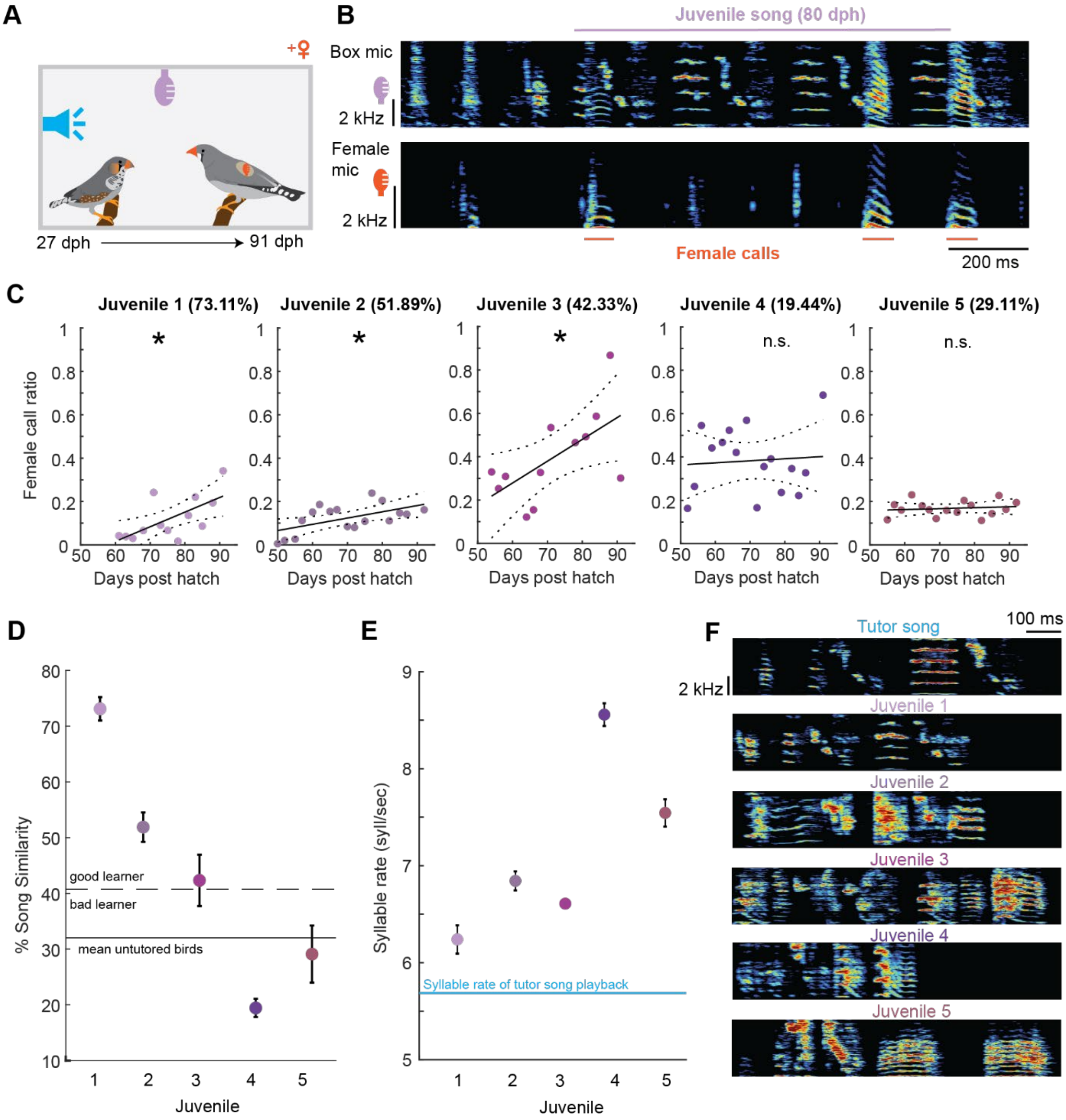


Figure 3: Correlated female calls to juvenile practice improve song learning. A) Experimental setup with two microphones – box microphone (purple) and backpack microphone (orange). B) Example spectrograms of a simultaneous recording from the box microphone and the backpack microphone during song practice. Song (purple) and female calls (orange) were automatically detected using the Deep Audio Segmenter (DAS)[48]. C) Ratio of female calls occurring during juvenile song practice relative to the duration of the song production (see Methods) for each individual juvenile 1 -5 (n= 5 juveniles, n= 4 different females). Purple circles represent call rate during songs. Black line indicates the linear correlation (juvenile 1: R^2=0.435, F=8.46, juvenile 2: R^2=0.325, F=7.23, juvenile 3: R^2=0.38, F=6.12, juvenile 4: R^2=0.005, F=0.008, juvenile 5: R^2=0.02, F=0.29), dotted line shows confidence intervals of calculated correlation. Numbers in brackets indicate the mean song similarity score to tutor song on the last day of practice. D) Song similarity (mean value in purple, error bars represent standard deviation, n=10 song motifs per juvenile) at day 91 (juvenile 1 and 4) or 92 (juvenile 2, 3 and 5) post hatch for all five juveniles. Black line indicates song similarity of untutored birds, dashed lines indicate confidence interval of 1.96 standard error of the mean. E) Syllable rate of the song of the individual juveniles. Syllable rate of the tutor song playback is indicated with the blue line. F) (*Top panel*) Spectrogram of the tutor song. (*Bottom three panels*) Spectrogram of the learnt song of juveniles 1-5.

Female vocalizations evoke neural responses in HVC projection neurons of listening juvenile and adult males

Perception of the tutor song induces changes in the neural circuitry of the premotor nucleus HVC in juvenile birds[36,38]. Since the premotor nucleus HVC receives auditory inputs[49–51] and is involved in song learning and production[37,52], we asked if female vocalizations can similarly evoke neural responses in the premotor circuitry in juvenile male zebra finches.

We performed intracellular recordings in HVC of awake, listening male juveniles while presenting female call playbacks (Figure 4A, B). In total, we recorded four antidromically identified HVC_{RA} neurons (neurons that project to the robust nucleus of the arcopallium (RA) and are directly involved in song production) and 19 unidentified HVC projection neurons. Since no quantifiable difference in neuronal response patterns could be observed we pooled the data and considered the recorded neurons as HVC projection neurons (n=23 HVC projection neurons, n=9 juvenile birds). The cells' firing rate was increased during the playback of a female call (silence: firing rate=3.64±5.02 Hz, female call: firing rate=6.64±8.16 Hz, p=0.0286, linear mixed-effect model, Figure 4C). Next, we analyzed the spiking precision which describes how stereotyped and time-locked spiking events are related to a female call. Due to the variability in the firing properties of each cell, we initially computed spiking precision during periods of silence as a baseline. Subsequently, we compared the baseline spiking precision to the spiking precision observed during female calls. We found that HVC projection neurons exhibit precise spiking events in response to female calls (silence: spiking precision=-0.003±0.62, female call: spiking precision=0.78±1.14, p=0.005, linear mixed-effect model, Figure 4D). This observation led us to ask whether the surrounding population of HVC projection neurons receive time-locked, stereotyped inputs during female call presentation, which might not be reflected when solely analyzing the spiking precision. Therefore, we calculated the subthreshold correlation and found that fluctuations in cell membrane potential were more correlated when a female call was presented compared to a silent period (silence: subthreshold correlation=0.011±0.06, female call: subthreshold correlation=0.16±0.17, p<0.001, linear mixed effect model, Figure 4E). The overall change in subthreshold input during female call presentation was not different from regular occurring subthreshold activity during silence (median Δ subthreshold= 1.5835 mV, p=0.56, Wilcoxon rank sum test) (Figure 4F) indicating that the more correlated activity observed in the subthreshold activity (Figure 4E) is not systematic but either reflects a release from local inhibition or the transmission of excitatory female call-related information from upstream auditory areas.

Next, we asked if the timing of the female calls with respect to the juveniles' song has an impact on the neuronal activity. Therefore, we presented the juveniles with playbacks of the birds own song (BOS) or with BOS interleaved with a female call during song practice while recording 15 of the aforementioned recorded neurons in 6 birds (Figure 4G). The neurons increased their firing rate in relation to BOS + female call compared to BOS alone (BOS: firing rate=2.81±2.94 Hz, BOS + female call: firing rate =5.41±6.31 Hz, p=0.03, linear mixed-effect model, Figure 4H). In contrast, HVC neurons did not exhibit time-locked spiking activity (BOS: spiking precision=0.39±1.58, BOS + female call: spiking precision=0.795±1.38, p=0.205, linear mixed-effect model).

In the BOS + female call condition, nor the subthreshold correlation neither the change in subthreshold activity were significantly different on a population level (Figure 4K,L). These findings suggest that auditory responses to female calls in HVC are reduced during female calls within BOS playback, when compared to auditory responses to female calls alone (Figure 4F).

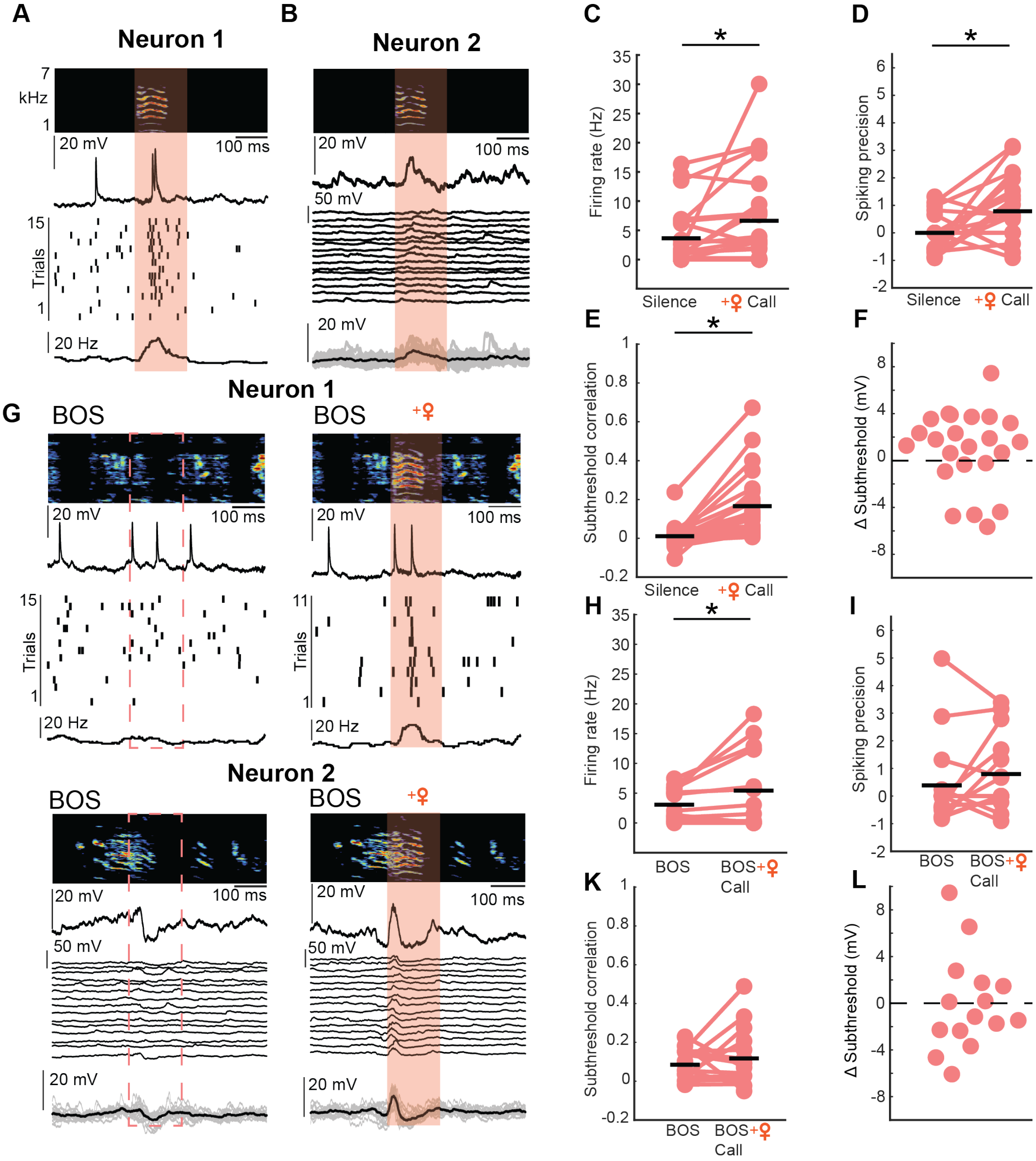


Figure 4: Female vocalizations elicit neural responses in a subset of HVC projection neurons in juvenile birds. A) Example recording of an HVC projection neuron increasing spiking activity during female call playback (call onset to call offset +50 ms). Orange-shaded area represents timing of a female call + 50 ms, top: Spectrogram of female call playback. middle: Example trace of a membrane potential of an HVC projection neuron during female call presentation, below: Spike dot raster of the neuron shown above during 15 representations of a female call, bottom: peri-stimulus time histogram (black line). B) Recording of an HVC projection neuron during 15 representations of a female call. Top: Spectrogram of female call playback, middle: example trace of membrane potential, below: membrane potential traces from 15 trials, bottom: average membrane potential (black) and trial membrane potentials overlaid (grey). C) Firing rate of all recorded cells during silence and call playbacks (n = 23 projection neurons recorded from 9 birds) black lines indicate the mean. D) Spiking precision of 23 recorded cells. E) Subthreshold correlation of all 23 cells. F) Difference (Δ) of subthreshold activity between silence and female call playbacks. G) Example recording of an HVC projection neuron (same cell as in A) with increasing spiking activity locked to female call during bird's own song playbacks. Left: Segment of BOS without female call, right: same segment with female call. H) Firing rate of all 23 cells during the segment of BOS with the overlapping female call and during

the same segment of BOS without a female call. I) Spiking precision comparison as in H. J) HVC projection cell (same cell as in B) with subthreshold responses to female call playback within BOS. K) Subthreshold correlation comparison as in H. BOS: subthreshold correlation=0.08±0.08, BOS + female call: subthreshold correlation=0.12±0.15, p=0. 405, linear mixed-effect model). L) Difference (Δ) in membrane potential (as in F) between BOS playback and BOS playback interleaved with female call (median Δ subthreshold = -1.1492 mV, p=1, Wilcoxon rank sum test).

Next, we asked whether the same response pattern evoked by a female call can be observed in adult birds which cannot undergo further learning[53]. Therefore, we performed multi-channel recordings in HVC of awake listening adult male birds (n = 5 birds) that were presented with female calls. Based on firing rate characteristics[54,55], see Methods) we identified 133 HVC projection neurons. On a population level, HVC projection neurons significantly increased their firing rate in response to female call playback (firing rate during female call=1.93±3.6 Hz, firing rate during silence =0.92±0.53 Hz; p=0.001, linear mixed-effect model) (Figure 5A-C). Additionally, we also found that HVC projection neurons in adults displayed precise spiking responses to the female calls on a population level (spiking precision during playback = 2.49±2.52, spiking precision during silence=0.04±0.89; p<0.001, linear mixed-effect model) (Figure 5A, B, D).

This finding indicates that female calls can induce spiking activity in HVC projection neurons outside the context of singing beyond development. Female calls are a salient and positive feedback signal for adults resulting in dopamine release in other brain areas[56]. Our result further supports the notion that females may provide feedback which evaluates the courtship performance of the males. During adult song production, however, female calls do not change the stereotyped activity pattern of HVC projection neurons[42]. A previous finding that HVC inhibitory interneurons exhibited a variety of responses to call playbacks[57] suggests that HVC inhibitory interneurons might gate off the auditory information in order to shield HVC projection neurons from undergoing plasticity which might affect the stereotyped production of the adult song.

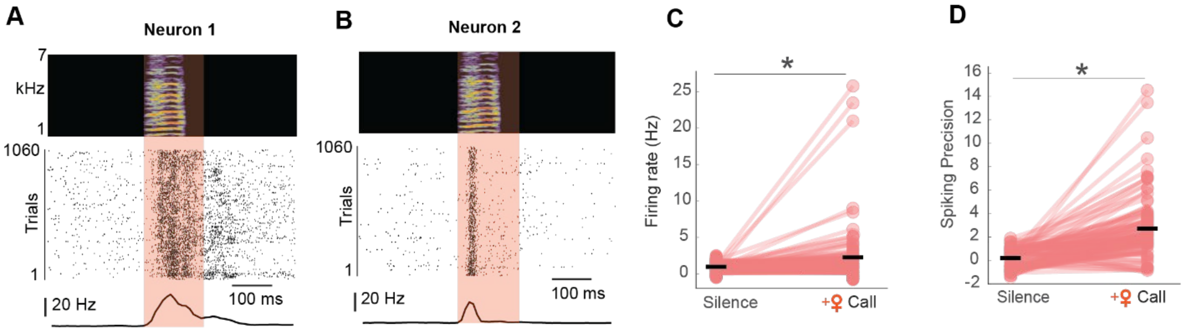


Figure 5: Auditory responses to female calls in HVC projection neurons of adult birds. A) Example HVC projection neuron (neuron 1) recorded during female call playbacks (call onset to call offset + 50 ms[58]). Top: sonogram of female call playback, middle: spike dot raster, bottom: peri-stimulus time histogram. B) Neuron 2, as in A. C) Firing rate of 133 HVC projection neurons recorded during silence and during the female call playback D) Spiking precision of 124 HVC projection neurons recorded during silence and the female call playback.

Female vocalizations induce responses in neural activity of HVC neurons in singing juveniles but not adults

The more stereotyped an adult male song is across renditions, the more attractive it is for a female zebra finch[28]. Female calls heard during courtship song production do not change the activity of HVC projection neurons in adult zebra finches[42]. Thus, female calls are unlikely to affect the stereotypy of ongoing male songs, which would presumably be beneficial in the context of courtship. We have presented evidence that feedback from the female appears to modulate juvenile song development. Therefore, we asked if the female calls have an impact on HVC projection neuron activity while the juvenile birds are practicing.

We performed intracellular recordings of HVC projection neurons in freely-moving, singing juvenile birds (n = 9 neurons in 4 birds, Figure 6A, B). The observed activity profile in juvenile birds older than 76 days post hatch was similarly stereotyped as the pattern observed in HVC projection neurons of singing adults[42] (juveniles: median subthreshold correlation=0.81, adults: median subthreshold correlation= 0.80, Figure 6C). To explore if the neural activity of the projection neurons is affected by the female call, we presented a female call playback during 22 instances of 44 song motifs. On a population level, the neurons exhibited a similar number of spikes during perturbed and unperturbed syllables (mean number of spikes = 0.04±0.40, p=0.917, linear mixed-effect model Figure 6D, F). However, in six out of the 22 cases when a female call occurred, neurons changed the number of spikes compared to baseline firing during song production (Figure 6D-F). This is an observation that is in contrast to data from adult zebra finches. In adults the number of spikes in projection neurons was unaffected when a female called during song production (Δ number of spikes=0; Figure 6F)[42]. To compare subthreshold changes caused by female call perturbations in adult and juvenile HVC projection neurons, we used the maximum delta subthreshold (as shown in Figure 3 of Vallentin and Long 2015) during a female's call while an adult zebra finch was singing as a threshold to determine if any activity changes occurred (Figure 6G). We found that in 5 out of the 22 cases neurons exhibited subthreshold changes in response to female calls that extended beyond the maximally occurring changes observed in adult birds. These findings indicate that female call perturbations during song practice in juveniles had an effect on the ongoing membrane potential activity and this might be a potential mechanism by which female feedback can induce plasticity within HVC projection neurons. However, when testing whether female calls have a direct effect on the ongoing song performance we did not find a significant improvement of the syllable following female call perturbation (Supplementary Figure 3). Similar to proposed endogenous mechanisms[59], exogenously introduced variability within this song production pathway could potentially facilitate exploratory singing behavior over subsequent song motifs, until a more accurate song copy has been achieved.

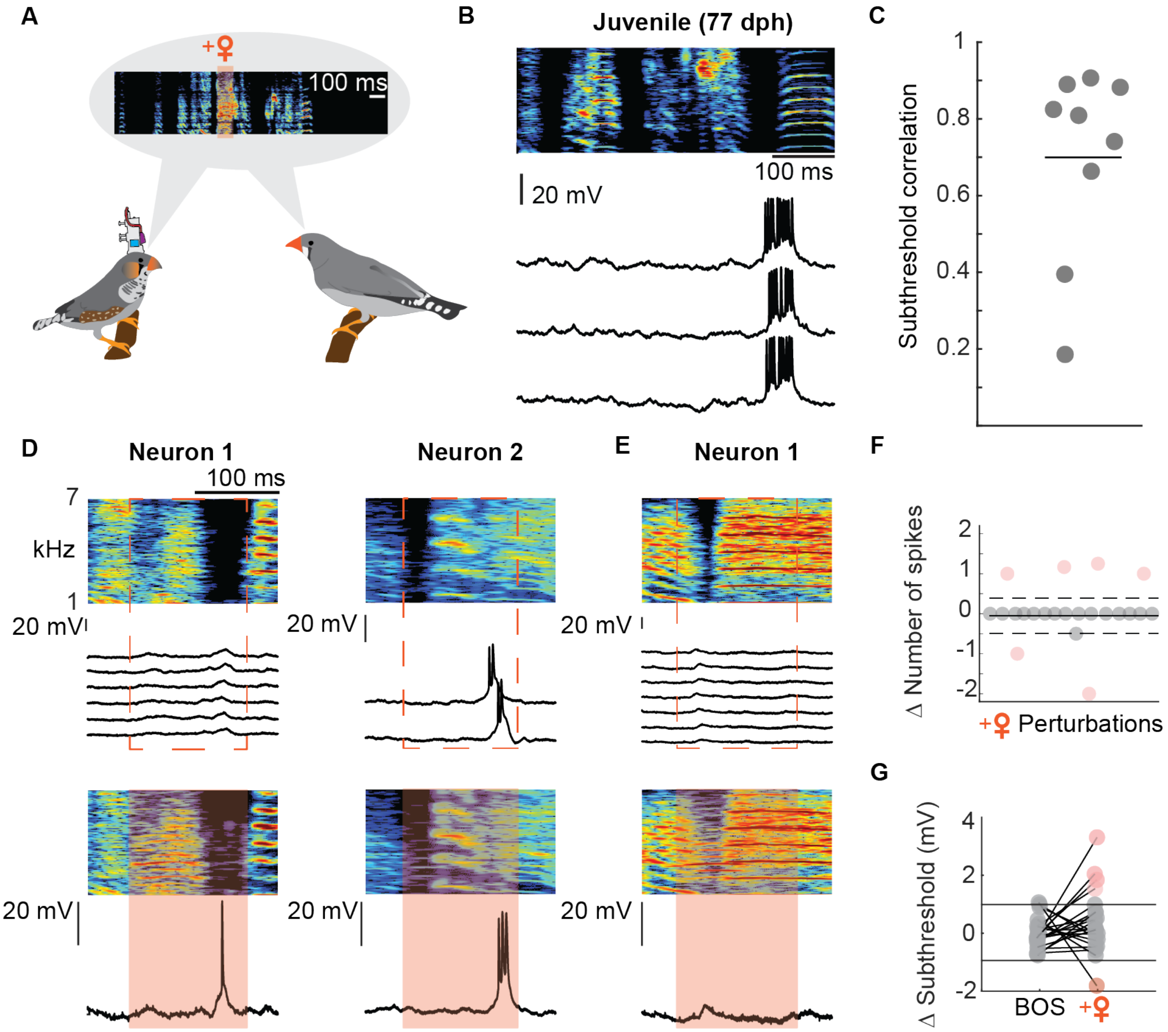


Figure 6: Female vocalizations induce changes in neural activity during song production. A) Experimental setup for intracellular recordings in freely-moving, singing juvenile birds (n=4 juvenile birds). B) HVC projection neurons in juvenile birds older than 76dph showed stereotyped neural activity during song production. Top: sonogram of the produced song motif. Bottom: intracellular membrane potential of the same neuron recorded during three renditions of the song motif. C) Subthreshold correlations across song renditions for nine recorded HVC projection neurons. D-E) Examples of female call perturbations during song practice. Top: Spectrogram of song segments with corresponding membrane potentials for each cell during song production. Dashed line represents when a female call will occur on the next song rendition. Bottom: corresponding song segments to the top, during female call perturbations. Below: Membrane potential during call perturbation. Orange shaded area highlights the time from call onset to call offset +50 ms. F) Delta (Δ) of number of spikes during female call perturbations (n=22 perturbations in nine HVC projection neurons). Black line indicates |mean Δ number of spikes| during BOS in adult zebra finches (data from Vallentin & Long 2015), dashed line indicates confidence interval of 1.96 standard deviations from the mean Δ number of spikes during juvenile's BOS. Orange dots highlight instances when the number of spikes changed beyond the confidence interval. G) Δ subthreshold changes during female call. Black lines indicate maximally observed Δ subthreshold changes in adult birds (data from Vallentin & Long 2015). Orange dots highlight instances when the Δ subthreshold changed beyond changes observed in adults.

Discussion

Social feedback is important for learning[7] but measuring its exact contribution is often difficult. By raising juvenile zebra finches with a tutor song playback accompanied by a female bird or alone, we were able to quantify the impact of female feedback on the song learning performance. We found that the female presence during the learning process of juveniles improved tutor song copying (Figure 1 & 2). Furthermore, in juveniles that copied the tutor song more accurately, the female calling rate increased along with the progression of the song learning phase (Figure 3). Correspondingly, female calls heard during the song development period resulted in activity changes in HVC projection neurons (Figure 4, 5 & 6). This female call-evoked activity persisted during adulthood, outside of the context of song production and might serve as a general mechanism for vocal feedback to signal the quality of song or encourage song production.

Song learning in differentially socialized juvenile zebra finches

The presence of a live tutor during learning enhances the ultimate similarity of the learned song to the tutor song, compared to hearing tutor song playbacks alone[20]. However, in these scenarios it is difficult to identify the social factors aside from the song that influence learning. We therefore paired juveniles with adult non-singing females to add a social component without allowing for tutor-tutee interactions. Since females have preferences for certain features of male song[28] and these preferences largely depend on early exposure[27,29] it is plausible that adult females also apply their learned song preferences to influence and shape juvenile song learning. Our control group was juvenile birds that are raised in social isolation, which could impose stress on these birds and therefore impact how well they learned to copy the tutor song[60]. We were nevertheless able to show, that some of the isolated juveniles learned to copy the song of their genetic father, which indicates that also in the social isolation these birds were capable of vocal learning.

The mere presence of a sibling female has been shown to enhance the copy to tutor song in blind-folded juvenile zebra finches[22]. It is therefore plausible, that even in the absence of visual cues, male juvenile birds are supported throughout the song learning process by females. Additionally, we noted that birds raised in the presence of females were less inclined to learn a song similar to their fathers' song. Early life exposure to fathers' song can already be sufficient for male juveniles to learn the fathers' song[61,62]. However, this early life experience can be overruled, if the juvenile birds are tutored by a male tutor[46]. In our study, we show that the presence of a non-singing female is sufficient to copy the tutor song rather than the fathers' song.

An additional possibility how females might promote song learning is their interactions with the tutor song which might be sufficient to enhance sensory acquisition in juvenile birds. We were not able to rule out that females display preference-related behavior when hearing the tutor song playback. However, social interactions with the pupil play a major role in how juveniles select tutors and ultimately in their song learning success[6]. We therefore hypothesize that direct social interactions with a female bird would have a greater influence on juvenile song learning than juveniles' passive observation of female preference behavior towards a song playback.

Female calls could provide auditory feedback during song practice

Since females do not sing but only produce calls[27], vocal feedback during song practice is a candidate mechanism with which females can direct juveniles to adjust their song performance during the critical period. Our results demonstrate that females produce calls during song practice and increase their call rate as song learning progresses. When female vocalizations were not correlated with the learning phase, the learned song was least similar to the tutor song. Throughout development, juvenile birds produce songs that vary greatly in their spectral and temporal features[63]. Female birds produced more calls as the juvenile song became more stereotyped, raising the possibility that female calls might function as rewarding stimuli, when juvenile birds produce a preferred version of the song. Although it has been shown that the female presence alone does not trigger dopamine release in HVC of juvenile birds[64], female calls are perceived as rewarding stimuli in adult birds[56]. Further studies addressing the involvement of the dopaminergic system are thus important to characterize the significance of vocal feedback in learning juveniles. Since our dataset does not provide a conclusive answer whether female calls alone aid song learning it is an intriguing future idea to investigate song learning in juveniles that are solely exposed to female call playbacks in a contingent or non-contingent manner in relation to song practice. Similar to tutor song playback experiments lacking a social partner[20] we would expect that female call playbacks still improve song learning but are less impactful compared to the presence of a live female calling during song practice. Given the potential of female calls to serve as a rewarding stimulus it should be further explored whether and how female calls also shape sensory learning of the tutor song.

Female calls elicit auditory responses in the premotor nucleus HVC of zebra finches

During the learning phase the premotor nucleus HVC receives auditory input which is later suppressed by synaptic inhibition[36]. Therefore, auditory-evoked activity in HVC projection neurons is absent during song production in adults[36,42]. Our results show auditory-evoked activity in HVC projection neurons of juvenile birds, in response to female calls. Auditory responses within the song network might support the stabilization of synaptic connections within HVC and therefore promote the production of a preferred and stereotyped version of the song. Female calls also elicit auditory-evoked responses in HVC of listening adult zebra finches outside of the context of song production. In line with this finding, auditory responses to female calls in HVC of adult males have previously been recorded during vocal call interactions[24]. Taken together, these observations support the notion that female vocalizations might play a role in directing HVC circuit dynamics during song production.

Projection neurons in HVC show responses to female calls during song production

Although female calls activate dopaminergic neurons in adult birds[56], auditory responses to female calls in HVC projection neurons appear to be absent during song production in adults[42]. Our observed auditory-evoked responses to heard calls in listening juvenile birds led us to explore if, similar to HVC premotor neuron responses to tutor song, responses to female calls during song production are initially present in juveniles. We observed exemplary changes in neural activity during song production, specifically when the song was perturbed by a female call. This finding indicates that female calls can induce responses in projection cells of HVC during song practice. Our results highlight the importance of social feedback from non-singing females during the song learning phase as well as its potential to shape the neural circuitry for song learning in juveniles. We provide novel insights into an exogenous feedback mechanism in learning zebra finches and highlight the contribution of vocal feedback, that might deepen our understanding of reinforcement learning within a complex, socially-embedded behavior.

Acknowledgements

We thank Fabian Heim, Tamir Eliav, Giacomo Costalunga, Jonathan Benichov and Aditi Agarwal for useful comments on a previous version of the manuscript; Ofer Tchernichovski for providing data from untutored birds; Lisa Trost and Petra Müschenborn for technical support; Constance Scharff for advice on the project and constant support, Jeanette Meinecke and Barbara Buhlmann for animal care and husbandry, Philipp Norton for helping to acquire the backpack microphone data.

Funding Sources

DFG - Emmy Noether Project VA 742/2 awarded to D.V.

DFG – SFB 1315 Projektnummer: 327654276 awarded to D.V.

ERC-2017-StG - 757459 MIDNIGHT awarded to D.V.

Competing interests

Authors declare no competing financial interests.

Methods

Animal housing

All procedures described were approved by the Regierungspräsidium Oberbayern (VET 02-21-201, VET 02-21-102) or Landesamt für Gesundheit und Soziales (LAGeSo Berlin) (G 0225/16) at the Freie Universität Berlin. Animals were housed with a light cycle from 7 a.m. to 7 p.m. and were provided with food, water, and grit *ad libitum.*

For the song learning experiments, juvenile male birds (25-35 days post hatch) were placed in sound-attenuated chambers. Each juvenile bird was housed either with an adult female or alone until the end of the song learning phase (max. 117 days post hatch).

For the electrophysiological recording experiments, juvenile birds were raised by their biological parents in the mixed-sex aviary or breeding cages. Adult male birds (at least 100 days post hatch) were housed with female adult birds for at least 72 hours before extracellular recordings.

Song learning

Juvenile birds (n=33 birds) were raised with both parents until nutritional independence (25-35 days post hatch). Afterwards juveniles were housed in sound-attenuated chambers during the critical period of song learning. Of all 33 juvenile birds, 22 birds were not related to each other. The remaining 11 birds came from 4 breeding pairs (pair A to pair D). Siblings from pairs A-C were distributed in both social conditions – raised in female presence or alone. Both siblings from pair D were raised in the alone condition. In the social context (with female), each juvenile bird was accompanied with one adult female throughout the song learning phase (n=18 birds). Adult female birds were acquired from the previously described mixed-sex aviaries. In the socially isolated context (alone) each juvenile bird was housed alone (n=15 birds). Both groups were trained as follows: At the beginning of the training phase (first two weeks) birds received 20 passive tutor song playbacks per day (7 a.m. to 2 p.m.). During the following phase, the birds could actively elicit playbacks of the tutor song themselves via pecking keys (n=40 playbacks per day, 7 a.m. to 7 p.m.). The tutor song playback consisted of two synthetized zebra finch syllables repeated twice – as previously described in[65]. The first syllable of the pair (syllable A, pitch = 2456.6 Hz, wiener entropy = -3.06; continuity of frequency = 360.53 Hz) had a duration of 185 ms and the second one (syllable B, pitch = 1165.4 Hz, wiener entropy = -3.754, continuity of frequency = 241.41 Hz) had a duration of 155 ms. The silent gap between syllables had a duration of 119 ms and 124 ms, respectively (Figure 1B).

Song similarity quantification

The entire song learning phase was monitored via audio recordings of song practice using SAP 2011[43]. At the end of the learning phase, similarity to the tutor song and syllable rate of the learned songs were analyzed. For calculating the similarity to the tutor song, we set SAP 2011 parameters to 'asymmetric' and 'time-courses' to capture if any syllables from the tutor song had been copied irrespective of their sequential order. Song similarity was then determined as the naturally occurring similarity across repetitions of the birds own song when ten different repetitions were compared (see [36]).

Syllable rate was determined by using sound-envelope based amplitude thresholding to segment syllables (200-250 syllables per bird recorded at 91-117 dph, n=33 birds). After segmentation, the number of syllables detected were added up and divided by the sum of syllable durations of all syllables detected:

$$Syllable\ rate = \frac{number\ of\ syllables}{sum(syllable\ duration)}$$

Bird's own song and female call recordings used in electrophysiological experiments

Songs of male birds were recorded with SAP 2011 at 44.1 kHz. Male birds were placed individually in sound-attenuated chambers, accompanied by a female bird. After a brief habituation period, male juveniles produced songs that were then used as BOS playbacks for intracellular recordings. The female call was extracted from a vocal interaction period between an adult female and a juvenile male bird. The same female call was used for all intracellular recordings. For the experiments with adult birds, female calls were recorded during natural interaction between the male and the female. In total, a stack call from two different females were used as playbacks. Each adult male bird was only presented with the call playback from a female it was housed with. The sound recordings were normalized, band pass filtered (300-14000Hz), and a 20 kHz pure tone (outside zebra finch hearing range)[66] was embedded for the duration of the call to detect call onsets and offsets.

Tracking female vocalizations during song learning

Juvenile birds (n=6 birds) were paired with an adult female (n=4 birds) in a song learning setup as described above. During the song learning phase, juvenile birds were tutored with a song playback of the same adult bird. Female zebra finches were equipped with a microphone attached to their back with soft elastic leg straps[47]. The box microphone recording all sounds and the backpack microphone recording only the female sounds were then synchronized using a custom Matlab program. Recordings were made with the same Matlab program at 40 kHz sampling rate.

Analysis of female vocalizations during song practice

We trained a neural network[48] to detect juvenile song throughout the song learning phase and a separate neural network was trained to recognize female vocalizations. We used the default setting in DAS and used 4 TCN blocks for training. As a quality measurement we ensured that the precision of the network was at least 80%, recall was set to at least 68%. If the precision score of the neural network per bird was below 80%, we added additional training data to reach the target.

We used one neural network for each bird to detect syllables from the box microphone recordings and female calls from the backpack microphone recordings on every 2-3 days of the learning phase. Once the juvenile male song syllables and adult female calls were detected, we analyzed how many female calls occurred during song practice.

Song was considered to be an at least 400 ms long event of consecutive syllables. The end of each song was defined by at least 350 ms of silence between syllables. To detect how many calls occurred during song practice, a time window of interest was specified as onset of song to offset of song. A call was considered to be song practice - related, if the onset of the call was detected within this window.

To evaluate, if female birds call during song practice, a call ratio was calculated to measure the number of calls throughout the duration of song:

$$Call\ ratio(nr/sec) = \frac{Number\ of\ call\ onsets\ detected\ within\ the\ window\ of\ interest}{Duration\ of\ the\ window\ of\ interest\ (sec)}$$

Analysis of syllable similarity to tutor song after female call perturbation

We used the dataset of monitored female feedback during the song learning phase (Figure 3) and detected which juvenile songs received female calls during song during every 7-12 days. We then extracted 300 ms long snippets of song directly after the female call. These snippets included the next syllable after female call. For control, we also extracted 300 ms long snippets of unperturbed songs. To ensure higher comparability, we extracted song snippets from unperturbed songs at the same duration from song onset as the perturbed snippets. We analyzed 264 song snippets of four birds, bird 5 was not analyzed due to lack of female calls during song. Similarity analysis was done as described above.

Surgery

For head-fixed awake electrophysiological recordings, male zebra finches (n=6 birds) were implanted with a headplate for fixation. First, male birds were anesthetized with isoflurane (concentration: 1-3% isoflurane, 97-99% oxygen). After careful incision, the skull was exposed and the trabecular bone structure above the area of interest was removed with a dental drill (carbide bur, FG ¼, Johnson-Promident). Second, the robust nucleus of the arcopallium (RA) was targeted according to coordinates (head angle: 65°, bifurcation of the midsagittal sinus was used as a reference point; RA: posterior 1.85 mm, lateral 2.25 mm, ventral 1.8 mm; HVC: anterior 0.2 mm, lateral 2.3 mm, ventral 0.2 mm, for extracellular recordings head angle was 45°, HVC coordinates: 0.3 mm anterior, 2.35 mm lateral). Third, a carbon fiber electrode (Kation Scientific, LLC) was lowered and RA was identified based on its characteristic firing pattern[67,68]. The location of HVC was then confirmed via antidromic stimulation from RA[40]. Fourth, the RA stimulating electrode and a headplate were implanted using light acrylic and dental cement (Paladur, Heraeus). For signal reference, a small craniotomy above the cerebellum was made and a groundwire (0.05 mm bare, silver) was placed between the skull and the dura mater covering the cerebellum. Lastly, all craniotomies were sealed with a silicone elastomer (Kwik-Cast) to prevent desiccation.

Electrophysiological recordings

Intracellular recordings were performed after one recovery day after surgery. Animals were placed in a small foam-lined container to restrict movement. The head of the animal was fixed to allow stable access to the craniotomy.

- *Intracellular recordings in awake head-fixed juvenile birds*

Sharp intracellular electrodes (borosilicate glass with filament, 0.1 mm diameter) were pulled using a micropipette puller (Model P-97, Sutter Instrument) and backfilled with potassium acetate (concentration: 3M). The identified location of HVC was secured by building a well with silicone elastomer around the craniotomy, that was then filled with phosphate buffered saline (PBS). Dura was carefully removed using a dura-pick to gain access to HVC. Intracellular recordings were performed by lowering the glass pipette into HVC with a micromanipulator (model MP-285A, Sutter Instrument), no more than 10 micrometers at a time. Current depth of the electrode was assessed by the micromanipulator. Every time the electrode was lowered, a brief buzzing pulse (10-20 ms) was elicited to enter the cell membrane. Cells selected for further analysis had at least 30 mV action potentials with a resting membrane potential below -50 mV and the recording lasted for at least 3 minutes. To identify the neuron type, antidromic stimulation was performed and only cells with low jitter antidromic spikes were considered to be HVC_{RA} neurons. Based on the low firing rate[41] of the recorded neurons and the characteristic spike waveform we concluded that none of the cells recorded were inhibitory interneurons.

- *Intracellular recordings in freely moving juvenile birds*

We used juvenile birds that were normally reared within our breeding colony up to 60 days post hatch. After 60 days post hatch, birds were placed into an adjacent aviary where several adult female birds were already housed. All juvenile birds maintained a visual and auditory contact with their biological parents. To record intracellularly from a freely moving bird, we assembled and implanted a motorized microdrive[41]. After the surgery (as described above), we allowed the juvenile bird to recover for the next 2-3 days until the bird was singing with the implanted microdrive. Intracellular recordings were performed as described above. Upon successful recording of a cell, the juvenile bird was presented with one of the adult female birds from the aviary to motivate singing behavior. The first 2-3 motifs of directed singing were left unperturbed. For subsequent motifs we triggered a playback of a female call manually to sample neural activity of female call perturbations during song production.

- *Extracellular recordings in awake head-fixed adult birds*

Neuropixels probes[69] were used to record in awake, head-fixed adult birds. Only recording sites that covered the ventral length of HVC were considered. A New Scale micromanipulator was used to insert a single probe at an angle of 10°. Once it reached the desired depth, it was stabilized for 20 minutes. The recording was carried out using an external reference and shared ground-reference configuration. Then, playbacks of a female call were presented every 1.5 seconds for a minimum of 60 trials. Audio and neural signals were both acquired simultaneously and aligned offline.

Data analysis

All statistical test used are specified in the main text. If statistical test for normally distributed data were applied, the dataset was first confirmed to be normally distributed by an Anderson-Darling test. All medians, means ± standard deviations are reported in the text.

Firing rate: To determine whether a cell responded in relation to a playback of a female call, we analyzed the change in firing rate, during call playback (call duration +50 ms[58]) and compared it to the firing rate of a silent period of the same duration that occurred before the call (offset of silent period was aligned to call onset -50 ms to avoid any overlap). Firing rate was calculated for every trial as number of spikes per given time window. For every cell, the average firing rate was calculated across all trials of the same auditory stimulus or silent period, respectively. For intracellular recordings, spikes were detected at 15 mV threshold above the baseline membrane potential. For extracellular recordings, data was acquired using SpikeGLX, spikes were detected and single units sorted using Kilosort 2.5. Doubled counted spikes were removed using ecephys spike sorting repository. Sorted clusters were manually curated using phy, and single units were obtained after calculating quality metrics considering an interspike interval violation index (isi)>0.25, amplitude cutoff distribution<0.3, and presence ratio>90%[70]. To differentiate projection neurons from interneurons, we used previously described spontaneous firing rates typical for both types of neurons. We considered neurons with a spontaneous firing rate <= 2 Hz as HVC projection cells, and correspondingly, cells with a firing rate > 2 Hz as HVC interneurons.

Spiking precision: To assess how time-locked the cell is firing, a precision score was calculated as previously described in[36]: Differences in latencies between all spikes, across all trials of the same auditory stimulus were calculated and the mean latency difference was extracted. To test for significance, a permutation test was used: Spikes were shuffled across trials 1000 times preserving the statistical distribution of their occurrence and then the mean latency from the shuffled data was extracted. Responsive cells were defined as having a precision score outside of 95% of the shuffled mean latency distribution.

Subthreshold correlation: To specifically determine changes in subthreshold activity in relation to playbacks, we cut off the spikes detected at 15 mV threshold above baseline membrane potential and linearly interpolated the resulting membrane potential traces. The subthreshold activity during playback and the aforementioned silent period were compared. To analyze the stereotypy of the subthreshold activity during auditory stimuli, the correlation of subthreshold traces at a 0 lag across all trials of the same stimulus was calculated. An average value per cell across all trials of the same stimulus was calculated. Confidence interval for population analysis was defined as ± 1.96 standard error of the mean.

Responsiveness to female call perturbation: To determine whether a cell was responsive to a female call perturbation, we first detected the number of spikes during the female call perturbation +50 ms to allow for delayed auditory responses[58]. We then compared the number of spikes during the female call perturbation to the average number of spikes during the same epoch during song without the female call. For control, we calculated the difference in number of spikes during BOS renditions only. We then calculated a confidence interval of 1.96 standard deviation from the mean of the delta number of spikes during BOS. We marked any delta number of spikes during the female call perturbation In Figure 6F, if it exceeded the confidence interval boundaries.

Statistical analysis:

- Neuronal data: For metrics that could be measured in individual trials (firing rate), we analyzed the data for each trial, controlling for neuron identity in order to account for any variability in the number of trials for each neuron. In metrics that measured properties across trials (precision), we analyzed the summary parameters for each neuron, controlling for bird identity. In order to account for the nested structure (neurons within birds) in continuous data, we fitted linear mixed model by maximum likelihood estimation (changes in neuron firing rate/ spiking precision depending on condition (baseline or female call). We used the following model specifications:

neural activity ~ 1 + condition + (1 | bird ID) + (1 | bird ID:cell ID)

- Behavioral data: To test for song similarity/similarity in syllable rate differences and account for the genetic background or early sensory experience of individual birds, we fitted a linear mixed model with the following model specifications:

Similarity (song or syllable rate) ~ 1 + condition + (1 | father ID) + (1 | father ID:bird ID).

Detailed results for individual statistical analysis can be found in the Supplementary Table 1.

Data availability

The data used in this study is available from the corresponding author upon request.

Code availability

The custom-made Matlab code will be made available from the corresponding author upon request.

References

1. Bolhuis, J. J., Okanoya, K. & Scharff, C. Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience* **11**, 747–759 (2010).
2. Jarvis, E. D. Evolution of vocal learning and spoken language. *Science* **366**, 50–54 (2019).
3. Böhner, J. Song learning in the zebra finch (taeniopygia guttata): Selectivity in the choice of a tutor and accuracy of song copies. *Animal Behaviour* **31**, 231–237 (1983).
4. Tchernichovski, O. & Nottebohm, F. Social inhibition of song imitation among sibling male zebra finches. *Proceedings of the National Academy of Sciences* **95**, 8951–8956 (1998).
5. Baptista, L. F. & Petrinovich, L. Social interaction, sensitive phases and the song template hypothesis in the white-crowned sparrow. *Animal Behaviour* **32**, (1984).
6. Clayton, N. S. Song tutor choice in zebra finches. *Animal Behaviour* **35**, 714–721 (1987).
7. Goldstein, M. H., King, A. P. & West, M. J. Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 8030–8035 (2003).
8. Gultekin, Y. B. & Hage, S. R. Limiting parental feedback disrupts vocal development in marmoset monkeys. *Nature Communications 2017 8:1* **8**, 1–9 (2017).
9. Allen, R. E. & Oliver, J. M. *The Effects of Child Maltreatment on Language Development*. 299–305 (1982).
10. Chen, Y., Matheson, L. E. & Sakata, J. T. Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 6641–6646 (2016).
11. Immelmann, K. Prägungserscheinungen in der Gesangsentwicklung junger Zebrafinken. *Naturwissenschaften 1965 52:7* **52**, 169–170 (1965).
12. Eales, L. A. The influences of visual and vocal interaction on song learning in Zebra finches. *Animal Behaviour* **37**, 507–508 (1989).
13. Böhner, J. Early acquisition of song in the zebra finch, Taeniopygia guttata. *Animal Behaviour* **39**, 369–374 (1990).

14. Williams, H. *Models for Song Learning in the Zebra Finch: Fathers or Others? Anita. Behav* vol. 39 745–757 (1990).

15. Zann, R. Song and call learning in wild zebra finches in south-east Australia. *Animal Behaviour* **40**, 811–828 (1990).

16. Mann, N. I. & Slater, P. J. B. Song tutor choice by zebra finches in aviaries. *Animal Behaviour* **49**, 811–820 (1995).

17. Zann, R. A. *The Zebra Finch - A Synthesis of Field and Laboratory Studies*. 197–247 (1996).

18. Immelmann K. Song development in the zebra finch and other estrildid finches | Semantic Scholar. *Bird Vocalizations* (1969).

19. Price, P. H. Developmental determinants of structure in zebra finch song. *Journal of Comparative and Physiological Psychology* **93**, 260–277 (1979).

20. Derégnaucourt, S., Poirier, C., Kant, A. V. der, Linden, A. V. der & Gahr, M. Comparisons of different methods to train a young zebra finch (Taeniopygia guttata) to learn a song. *Journal of Physiology Paris* **107**, 210–218 (2013).

21. Carouso-Peck, S. & Goldstein, M. H. Female Social Feedback Reveals Non-imitative Mechanisms of Vocal Learning in Zebra Finches. *Current Biology* **29**, 631-636.e3 (2019).

22. Adret, P. Vocal imitation in blindfolded zebra finches (Taeniopygia guttata) is facilitated in the presence of a non-singing conspecific female. *Journal of Ethology* **22**, 29–35 (2004).

23. Elie, J. E. & Theunissen, F. E. Zebra finches identify individuals using vocal signatures unique to each call type. (1234) doi:10.1038/s41467-018-06394-9.

24. Ma, S., Maat, A. ter & Gahr, M. Neurotelemetry Reveals Putative Predictive Activity in HVC during Call-Based Vocal Communications in Zebra Finches. *J. Neurosci.* **40**, 6219–6227 (2020).

25. Benichov, J. I. *et al.* The Forebrain Song System Mediates Predictive Call Timing in Female and Male Zebra Finches. *Current biology : CB* **26**, 309 (2016).

26. Williams, H. Birdsong and singing behavior. in *Annals of the New York Academy of Sciences* vol. 1016 1–30 (New York Academy of Sciences, 2004).

27. Chen, Y., Clark, O. & Woolley, S. C. Courtship song preferences in female zebra finches are shaped by developmental auditory experience. *Proceedings of the Royal Society B: Biological Sciences* **284**, (2017).

28. Woolley, S. C. & Doupe, A. J. Social context-induced song variation affects female behavior and gene expression. *PLoS Biology* **6**, 0525–0537 (2008).

29. Riebel, K. Early exposure leads to repeatable preferences for male song in female zebra finches. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **267**, 2553–2558 (2000).

30. Clayton, N. S. Song discrimination learning in zebra finches. *Animal Behaviour* **36**, 1016–1024 (1988).

31. Miller, D. B. Long-term recognition of father's song by female zebra finches. *Nature* **280**, 390–391 (1979).

32. Riebel, K., Smallegange, I. M. & Terpstra, N. J. Sexual Equality in Zebra Finch Song Preference: Evidence for a Dissociation between Song Recognition and Production Learning. **269**, 729–733 (2002).

33. Wei, J., Liu, Q. & Riebel, K. Generalisation of early learned tutor song preferences in female zebra finches (Taeniopygia guttata). *Behavioural Processes* 104731 (2022) doi:10.1016/J.BEPROC.2022.104731.

34. Lauay, C., Gerlach, N. M., Adkins-Regan, E. & Devoogd, T. J. Female zebra finches require early song exposure to prefer high-quality song as adults. *Animal Behaviour* **68**, 1249–1255 (2004).

35. Okubo, T. S., Mackevicius, E. L., Payne, H. L., Lynch, G. F. & Fee, M. S. Growth and splitting of neural sequences in songbird vocal development. *Nature* **528**, 352–357 (2015).

36. Vallentin, D., Kosche, G., Lipkind, D. & Long, M. A. Neural circuits: Inhibition protects acquired song segments during vocal learning in zebra finches. *Science* **351**, 267–271 (2016).

37. Roberts, T. F. *et al.* Motor circuits are required to encode a sensory model for imitative learning HHS Public Access Author manuscript. *Nat Neurosci* **15**, 1454–1459 (2012).

38. Roberts, T. F., Tschida, K. A., Klein, M. E. & Mooney, R. Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning. *Nature* **463**, 948–952 (2010).

39. Nick, T. A. & Konishi, M. Neural song preference during vocal learning in the zebra finch depends on age and state. *Journal of Neurobiology* **62**, 231–242 (2005).

40. Hahnloser, R. H. R., Kozhevnikov, A. A. & Fee, M. S. An ultra-sparse code underliesthe generation of neural sequences in a songbird. *Nature 2002 419:6902* **419**, 65–70 (2002).

41. Long, M. A., Jin, D. Z. & Fee, M. S. Support for a synaptic chain model of neuronal sequence generation. *Nature* **468**, 394–399 (2010).

42. Vallentin, D. & Long, M. A. Motor origin of precise synaptic inputs onto forebrain neurons driving a skilled behavior. *Journal of Neuroscience* **35**, 299–307 (2015).

43. Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B. & Mitra, P. P. A procedure for an automated measurement of song similarity. *Animal Behaviour* **59**, 1167–1176 (2000).

44. Araki, M., Bandi, M. M. & Yazaki-Sugiyama, Y. Mind the gap: Neural coding of species identity in birdsong prosody. *Science* **354**, 1282–1287 (2016).

45. James, L. S., Wang, A. S., Bertolo, M. & Sakata, J. T. Learning to pause: Fidelity of and biases in the developmental acquisition of gaps in the communicative signals of a songbird. *Developmental Science* **26**, (2023).

46. Mets, D. G. & Brainard, M. S. Genetic variation interacts with experience to determine interindividual differences in learned song. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 421–426 (2017).

47. Gill, L. F. *et al.* A minimum-impact, flexible tool to study vocal communication of small animals with precise individual-level resolution. *Methods in Ecology and Evolution* **7**, 1349–1358 (2016).

48. Steinfath, E., Palacios-Muñoz, A., Rottschäfer, J. R., Yuezak, D. & Clemens, J. Fast and accurate annotation of acoustic signals with deep neural networks. *eLife* **10**, (2021).

49. Nottebohm, F., Paton, J. A. & Kelley, D. B. Connections of vocal control nuclei in the canary telencephalon. *Journal of Comparative Neurology* **207**, 344–357 (1982).

50. Bauer, E. E. *et al.* A Synaptic Basis for Auditory–Vocal Integration in the Songbird. *Journal of Neuroscience* **28**, 1509–1522 (2008).

51. Akutagawa, E. & Konishi, M. New brain pathways found in the vocal control system of a songbird. *The Journal of comparative neurology* **518**, 3086–3100 (2010).

52. Long, M. A. & Fee, M. S. Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* **456**, 189–194 (2008).

53. Funabiki, Y. & Funabiki, K. Factors Limiting Song Acquisition in Adult Zebra Finches. *Inc. Develop Neurobiol* **69**, 752–759 (2009).

54. Egger, R. *et al.* Local Axonal Conduction Shapes the Spatiotemporal Properties of Neural Sequences. *Cell* **183**, 537-548.e12 (2020).

55. Kozhevnikov, A. A. & Fee, M. S. Singing-related activity of identified HVC neurons in the zebra finch. *Journal of Neurophysiology* **97**, 4271–4283 (2007).

56. Roeser, A. *et al.* Dopaminergic error signals retune to social feedback during courtship. *Nature* (2023) doi:10.1038/s41586-023-06580-w.

57. Norton, P., Benichov, J. I., Pexirra, M., Schreiber, S. & Vallentin, D. A feedforward inhibitory premotor circuit for auditory–vocal interactions in zebra finches. *Proceedings of the National Academy of Sciences of the United States of America* **119**, (2022).

58. Margoliash, D. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *Journal of Neuroscience* **3**, 1039–1057 (1983).

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

60. Perez, E. C. *et al.* The acoustic expression of stress in a songbird: does corticosterone drive isolation-induced modifications of zebra finch calls? *Horm Behav* **61**, 573–581 (2012).

61. Funabiki, Y. & Konishi, M. Long Memory in Song Learning by Zebra Finches. *Journal of Neuroscience* **23**, 6928–6935 (2003).

62. Roper, A. & Zann, R. The Onset of Song Learning and Song Tutor Selection in Fledgling Zebra Finches. doi:10.1111/j.1439-0310.2005.01169.x.

63. Tchernichovski, O., Mitra, P. P., Lints, T. & Nottebohm, F. Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science* **291**, 2564–2569 (2001).

64. Tanaka, M., Sun, F., Li, Y. & Mooney, R. A mesocortical dopamine circuit enables the cultural transmission of vocal behaviour. *Nature* **563**, 117–120 (2018).

65. Lipkind, D. & Tchernichovski, O. Quantification of developmental birdsong learning from the subsyllabic scale to cultural evolution. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 15572–15579 (2011).

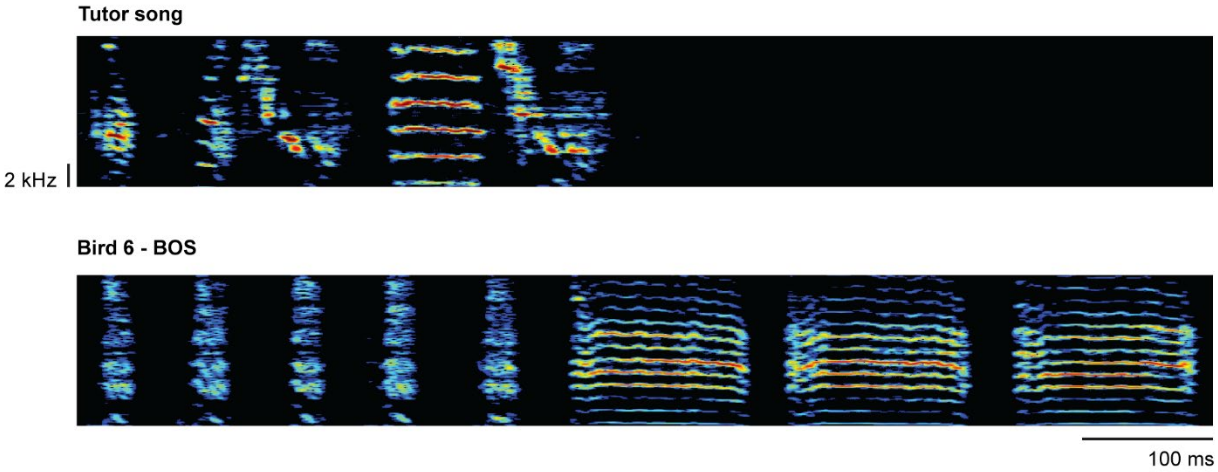
66. Okanoya, K. & Dooling, R. J. Hearing in passerine and psittacine birds: A comparative study of absolute and masked auditory thresholds. *Journal of Comparative Psychology* 7–15 (1987).

67. Yu, A. C. & Margoliash, D. Temporal Hierarchical Control of Singing in Birds. *New Series* **273**, 1871–1875 (1996).

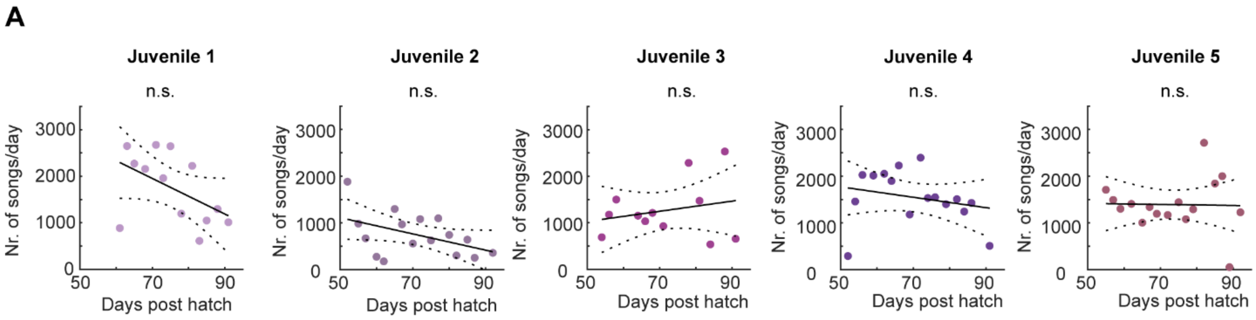
68. Spiro, J. E., Dalva, M. B. & Mooney, R. Long-range inhibition within the zebra finch song nucleus RA can coordinate the firing of multiple projection neurons. *Journal of Neurophysiology* **81**, 3007–3020 (1999).

69. Jun, J. J. *et al.* Fully integrated silicon probes for high-density recording of neural activity. *Nature 2017 551:7679* **551**, 232–236 (2017).

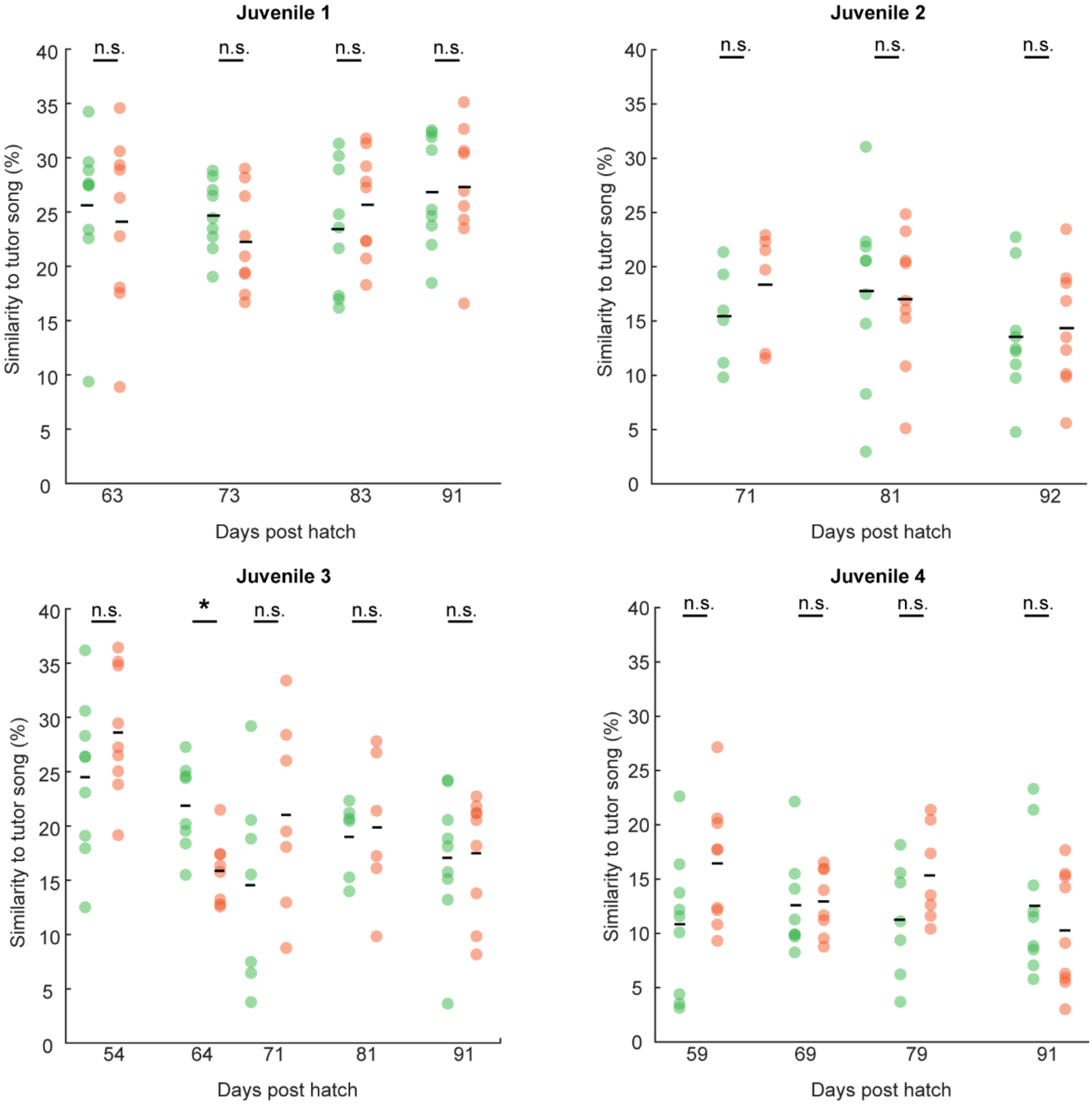
70. Hill, D. N., Mehta, S. B. & Kleinfeld, D. Quality Metrics to Accompany Spike Sorting of Extracellular Signals. (2011) doi:10.1523/JNEUROSCI.0971-11.2011.



Supplementary Figure 1: Bird 6 learnt song that consisted of unmodulated distance calls. Top: Tutor song playback that bird 6 was trained with. Bottom: Birds own song from bird 6 at the last day of training (90 days post hatch). First five song elements are introductory notes.



Supplementary Figure 2: Number of practice songs per day. A) Number of practice songs per day from 5 juveniles. Number of songs per day did not increase during the song learning phase.



Supplementary Figure 3: Similarity of perturbed and unperturbed song to tutor song. Similarity to tutor song playback of song syllables from up to 10 songs per day that were either perturbed (orange circles) or unperturbed (green circles). Juveniles 1-4 are the same birds as juveniles 1-4 in Figure 3. Reported statistic is a comparison of similarity of syllables that followed female call interruptions versus similarity of syllables that were sang without previous interruption (p>0.05 Wilcoxon rank sum test for all except for juvenile 3 on day 64 dph).

Supplementary Table 1

Genetic background or early sensory experience on song similarity after tutoring (Fig. 2)

Number of observations	Fixed effects coefficients	Random effects coef.	Covariance parameters
16	2	27	3

Formula: Song similarity ~ 1 + social condition + (1 | father ID) + (1 | father ID:bird ID)

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
Intercept	-2.58815	14.508	-0.19861	14	0.84542	-33998	28.235
‘background’	27.241	8.5564	3.172	14	0.0067886	8.789	45.492

Genetic background or early sensory experience on syllable rate after tutoring (Fig. 2)

Number of observations	Fixed effects coefficients	Random effects coef.	Covariance parameters
16	2	27	3

Formula: Syllable rate ~ 1 + social condition + (1 | father ID) + (1 | father ID:bird ID)

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
Intercept	2.3178	0.8617	2.6898	14	0.017605	0.46964	4.166
‘background’	2.856	0.59112	4.8315	14	0.00026635	1.5882	4.1238

Female calls on firing rate of juvenile HVC projection neurons (Fig. 4C)

Number of observations	Fixed effects coefficients	Random effects coef.	Covariance parameters
46	2	32	3

Formula: Firing rate ~ 1 + condition + (1 | bird ID) + (1 | bird ID:cell ID)

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
Intercept	0.91212	2.4868	0.366786	44	0.715544	-4.0997	5.924
‘condition’	3.0066	1.3282	2.2638	44	0.028581	0.32974	5.6835

Female calls on spiking precision of juvenile HVC projection neurons (Fig. 4D)

Number of observations	Fixed effects coefficients	Random effects coef.	Covariance parameters
46	2	32	3

Formula: Spiking precision ~ 1 + condition + (1 | bird ID) + (1 | bird ID:cell ID)

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
Intercept	-0.78957	0.41848	-1.8867	44	0.065805	-1.633	0.05383
‘condition’	0.78671	0.26467	2.9724	44	0.00477761	0.25329	1.3201

Female calls on subthreshold cross correlation of juvenile HVC projection neurons (Fig. 4E)

Number of observations	Fixed effects coefficients	Random effects coef.	Covariance parameters
46	2	32	3

Formula: Subthreshold cross correlation ~ 1 + condition + (1 | bird ID) + (1 | bird ID:cell ID)

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
Intercept	-0.14371	0.050165	-2.8646	44	0.006376	-0.24481	-0.042604
‘condition’	0.1545	0.030224	5.1116	44	6.6629e-06	0.093583	0.21541

Female calls during BOS playback on firing rate of juvenile HVC projection neurons (Fig. 4H)

Number of observations	Fixed effects coefficients	Random effects coef.	Covariance parameters
30	2	21	3

Formula: Firing rate ~ 1 + condition + (1 | bird ID) + (1 | bird ID:cell ID)

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
Intercept	0.70166	1.9121	0.36695	28	0.71642	-3.2152	4.6185
‘condition’	2.3585	1.0507	2.2447	28	0.032873	0.20611	4.5107

Female calls during BOS playback on spiking precision of juvenile HVC projection neurons (Fig. 4I)

Number of observations	Fixed effects coefficients	Random effects coef.	Covariance parameters
30	2	21	3

Formula: Spiking precision ~ 1 + condition + (1 | bird ID) + (1 | bird ID:cell ID)

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
Intercept	-0.022392	0.57952	-0.03864	28	0.96945	-1.2095	1.1647
‘condition’	0.4088	0.31538	1.2962	28	0.20548	-0.23722	1.0548

Female calls during BOS playback on subthreshold cross correlation of juvenile HVC projection neurons (Fig. 4K)

Number of observations	Fixed effects coefficients	Random effects coef.	Covariance parameters
30	2	21	3

Formula: Subthreshold cross correlation ~ 1 + condition + (1 | bird ID) + (1 | bird ID:cell ID)

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
Intercept	0.052106	0.062287	0.83644	28	0.40993	-0.075482	0.17969
‘condition’	0.032602	0.038556	0.84557	28	0.40496	-0.046377	0.11158

Female calls on firing rate of adult HVC projection neurons (Fig. 5)

Number of observations	Fixed effects coefficients	Random effects coef.	Covariance parameters
266	2	138	3

Formula: Firing rate ~ 1 + condition + (1 | bird ID) + (1 | bird ID:cell ID)

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
Intercept	0.11771	0.58634	0.20076	264	0.84104	-1.0368	1.2722
‘condition’	1.0067	0.30861	3.262	264	0.001252	0.39903	1.6143

Female calls on spiking precision of adult HVC projection neurons (Fig. 5)

Number of observations	Fixed effects coefficients	Random effects coef.	Covariance parameters
248	2	129	3

Formula: Spiking precision ~ 1 + condition + (1 | bird ID) + (1 | bird ID:cell ID)

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
Intercept	-2.5711	0.45859	-5.6065	246	5.5301e-08	-3.4744	-1.6678
‘condition’	2.4515	0.22312	10.988	246	4.1148e-23	2.0121	2.891

Female calls on spike count of juvenile HVC projection neurons (Figure 6)

Number of observations	Fixed effects coefficients	Random effects coef.	Covariance parameters
44	2	13	3

Formula: Spike count ~ 1 + condition + (1 | bird ID) + (1 | bird ID:cell ID)

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
Intercept	0.83731	0.67152	1.2469	42	0.21934	-0.51786	2.1925
‘condition’	0.041667	0.39768	0.10477	42	0.91705	-0.76088	0.84421

3 TEMPORAL DYNAMICS OF SONG CRYSTALLIZATION IN ZEBRA FINCHES

Temporal dynamics of song crystallization in zebra finches

L. Bistere[1,2], S. Wilczek[1] & D. Vallentin[1]

[1] Max Planck Institute for Biological Intelligence, Seewiesen, Starnberg, Germany

[2] Graduate School for Systemic Neurosciences, Munich, Germany

Abstract

Song learning in zebra finches is a well-established model for vocal motor skill acquisition. Young birds undergo a gradual transformation of their song, starting with high variability in juvenile stages and culminating in a stereotyped adult song. We investigated the neural correlates of song crystallization - a critical phase of song development where established syllables undergo refinement. This was achieved by performing intracellular recordings in HVC, a premotor area known for its role in song learning and production. Compared to the highly regular bursting activity observed in adult HVC projection neurons during singing, juveniles during song crystallization exhibited a variable occurrence of bursts with a lower firing rate. These slower and more variable temporal dynamics are also reflected in the subthreshold activity by longer duration and larger amplitude of excitatory post-synaptic potentials in juveniles. These distinct temporal dynamics within the HVC during song crystallization likely underlie the fine-tuning processes responsible for shaping the precise local spectral and temporal features of the zebra finch song.

Introduction

The process of acquiring a motor skill encompasses the refinement of a once highly variable behavior into a finely-tuned and precise execution. For example, in the initial stages of vocal learning, humans often produce vocalizations that are inherently variable and uncontrolled such as cries and babbling sounds[1]. After months of practicing these utterances shape into refined words with articulated pronunciation[2]. Male zebra finches undergo a comparable learning process for singing[3]. Within a critical period, their initially variable subsong transitions into a plastic song, ultimately culminating in the acquisition of a final, stereotyped song closely resembling that of their tutor[4].

Using bioacoustics, the song learning progression of developing zebra finches can be described by assessing changes in spectral features or temporal characteristics of song[4–6]. Subsong (~25-50 days post hatch (dph)) consists of poorly structured sounds with high variability of spectral features (wiener entropy, spectral continuity, pitch and frequency modulation)[7]. Subsong is followed by a plastic song (~50-80 dph) consisting of structured syllable production that gradually develops distinct acoustic features until a stereotyped, crystalized song is achieved which is characterized by precise, reliably repeated features[7]. In adults the spectral features of song are highly stereotyped and acoustic parameters have a low variability.

Throughout the developmental phase, neural dynamics in the vocal production pathway undergo various modifications[8–13], ultimately leading to the emergence of a stable neural pattern during song production[14–19]. Initially the lateral magnocellular nucleus of the anterior neostriatum (LMAN) is necessary for the production and timing of subsong vocal patterns[8]. LMAN projects to the robust nucleus of the arcopallium (RA) which in turn projects to the brainstem vocal and respiratory nuclei[20], leading to audible song production. During the song learning phase, the main input to RA switches from LMAN to HVC, which on a behavioral level is represented by increased stereotypy in timing and spectral features of song [21]. It has been shown to be a gradual and overlapping transition since HVC or LMAN inactivation during the plastic song phase either reverts the juveniles singing into subsong or into the production of adult like stereotyped song sequences[10,22].

We aimed to explore the neural underpinnings that might fine-tune temporal dynamics of plastic song to eventually become a stereotyped song. Between 70 and 90 days post hatch the temporal structure of the song changes i.e. the silent gap duration decreases and the overall timing variability declines[23]. While RA is necessary for song production throughout all stages of development its basic spiking characteristics develop concomitantly with spectral features and it does not generate timing[24]. LMAN is inducing temporal variability but only during the subsong phase, since after 60 dph, the RA motor program is predominantly driven from HVC[8]. Although HVC is not necessary for subsong it is necessary for all later stages of song production and it has been shown to generate the timing of the song in adults[24] but also juveniles in which HVC lesions result in the abolishing adult like gap and syllable durations and HVC cooling results in an elongation of syllables and gaps[13].

During the early plastic song phase, HVC projection neurons produce rhythmic bursts during several syllables[11]. As the song matures, the bursting activity becomes tied to one particular timepoint during song[11] but how these neural changes relate to alterations in the temporal features of singing behavior remains unclear. Here we wanted to directly investigate the neural activity changes at a single cell level that underly the transition from plastic song to crystalized song. We quantified the neural dynamics including changes in spike characteristics and the corresponding intrinsic properties as well as the activity of surrounding population by analyzing subthreshold membrane potential in singing zebra finches. Ultimately the crystallized song becomes the stereotyped mating song of the male which is more attractive to the female the more temporally precise it is sang.

Results

HVC projection neurons in juveniles exhibit variable bursting patterns

In adult zebra finches the motor program underlying song production is made up of highly stereotyped sequences of bursts in HVC that are precisely time-locked to a particular timepoint in song. To investigate network activity during late-stage song learning, we performed intracellular recordings from HVC projection neurons in freely moving, singing male juvenile zebra finches between the ages of 74 to 94 days post hatch (Table 1). These birds displayed a relatively stereotyped singing behavior. When comparing the temporal features of song, we observed that juveniles tended to produce shorter syllables than adults (syllable durations (median±std) juveniles=108.2±67.8 ms, adults=166.2±95.4 ms, linear mixed-effects model, p=0.04). In contrast, the gap durations of juvenile birds were comparable to those of adults (gap durations (median±std) juveniles=35.8±24.9 ms, adults=31.3±16.9 ms, linear mixed-effects model, p=0.21, Supplementary Figure 1). Since HVC controls the duration of both syllables and gaps during plastic song phase[13,24], we next analyzed the neural activity of HVC projection neurons in singing juveniles and adults. When linearly warping the motifs to the same duration we found that HVC projection neurons elicited sparse bursts of action potentials during a specific time point during song production. The distribution of bursts in time from all recorded neurons covered the entirety of the motif in juveniles and adults[15,16] (Figure 1A, 1B). This finding supports the notion that during the late phase of song learning neural sequences are already formed and the song is produced in a stereotyped manner. However, it cannot explain observed temporal variability in juvenile and adult songs.

In adult zebra finches HVC projection neurons display a restricted number of bursts during song production[14,25]. To explore whether our recorded neurons produced ultra-sparse bursting patterns, we calculated the maximum number of bursts per trial for each HVC projection neuron. In juveniles we observed an increased number of bursts during motif per projection neuron when compared to adults (juveniles: 0-4 bursts per motif, median=2, n=10 neurons in n=4 juveniles; adults: 0-4 bursts per motif, median=1, and n=54 neurons in n=10 adults p = 0.03, Wilcoxon rank sum test, Figure 1C). Next, we measured how stereotyped bursting activity was across song motifs. We identified bursts as reoccurring bursts across song motifs, if their onset time was within ± 20ms across trials. In adults we observed 85.94% (n=55/64 bursts) of all recorded bursts to be reoccurring, in juveniles we found 60%% (n=12/20 bursts) of reoccurring bursts (p= 0.0224, Fisher Exact Test). These observations are in line with previous work[11] and suggest, that the structural connectivity within HVC exhibits greater variability in late-stage juveniles compared to adults.

In addition to bursting activity some neurons also exhibited single action potentials during song production which reduces the sparseness of the neural code and might induce less reliable behavioral outcomes. To test whether single action potentials might contribute to a less temporally stereotyped song we explored the spiking activity of the HVC neurons (spikes within a burst and single spikes). We observed, that the median number of spikes per trial was higher in juvenile birds than in adults (juveniles: mean=7.1±4.72, median=6.5, adults: mean=4.45±4.28, median=4, Wilcoxon rank sum test, p=0.048). Higher number of spikes per trial could also be accounted to single spikes alone outside of bursts. We therefore considered median single spikes across song renditions and found that HVC neurons in juveniles elicited comparable number of single action potentials to adults (action potential (juveniles)= 1.5±2.56, median=0.25, action potential (adults)= 0.48± 1, median=0, p= 0.24, Wilcoxon rank sum test, Figure 1D), suggesting that the higher number of spikes per trial may be attributed to internal HVC projection neuron burst dynamics in juveniles. These results show that in late-stage development the HVC projection neurons in juveniles exhibit detectable differences in bursting activity when compared to adults. However, this activity is not reflected in spiking activity outside of bursts. We next explored, whether the late-stage development can be accounted to the temporal dynamics within bursts in HVC projection neurons.

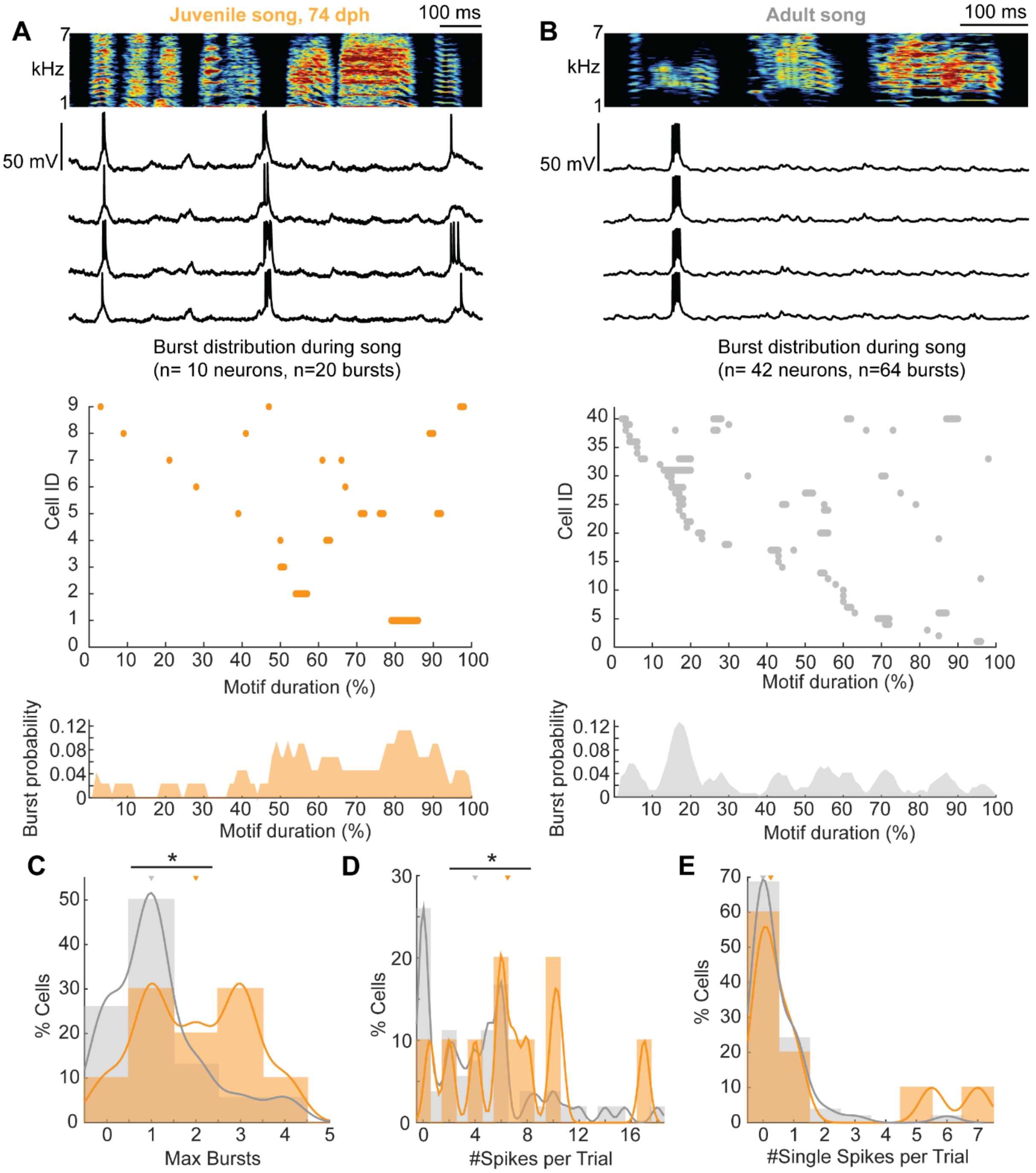


Figure 1: HVC projection neurons in juvenile birds burst frequently during song production. A) Top: example recording of an HVC projection neuron (top: spectrogram of a song motif, below: intracellular traces of an HVC projection neuron during four repetitions of the song motif), Middle: Raster-plot of bursting activity during singing a motif (0-100% of duration) in all birds, Bottom: Probability distribution of burst occurrence during motif. B) Same as in A bur for adult birds. C) Maximum number of bursts during song production between juveniles (orange) and adults (grey). D) Number of spikes per trial in juveniles (orange) and adults (grey). E) Distribution of single spikes per trial across all recorded projection neurons in juveniles and adults.

Temporal dynamics of intrinsic bursting activity are delayed in juveniles

We explored the intrinsic dynamics of individual bursts by analyzing their individual number of spikes and temporal characteristics of these spikes during song production (Figure 2A, 2B). First, we quantified the median number of spikes occurring per burst. In both juveniles and adults, we observed a similar number of spikes occurring per burst (median spikes per burst ± std, juveniles=3±1.95, adults=3±1.94, Wilcoxon rank sum test, p=0.2653, Figure 2C). Next, we assessed the variability of the number of spikes occurring in each reoccurring burst. Juvenile birds had a comparable distribution of Δ number of spikes per burst (range: Δ number of spikes per burst 0-2; mean Δ number of spikes per burst =0.32±0.61, median=0) to adults (0 to 1.2 spikes jitter, mean=0.37±0.29, median=0.33, Wilcoxon rank sum test, p=0.09 Figure 2D), which indicates stereotyped reoccurring bursts during the late stage development. Since we observed temporal differences in song production between juveniles and adults, we next investigated whether these differences could be attributed to the differences in the temporal structure of bursts in juveniles and adults by quantifying their duration. On a population level we did not observe differences in mean duration of bursts between juveniles and adults (mean ± std, juveniles=10.63±8.43 ms, median=9.64 ms, adults=8.54±6.96 ms, median=6.83 ms, Wilcoxon rank sum test, p=0.15, Figure 2E). However, when assessing the instantaneous firing rate within bursts, we found that it was lower in juveniles than adults (mean ± std juveniles=229.77±72.5, median=226.09, adults=380.37±165.49, median=387.68, Wilcoxon rank sum test, p<0.001, Figure 2F) indicating that premotor signals are timed and integrated on a slower timescale compared to in the adult brain. To quantify whether the lower instantaneous firing rate also produced a distinct temporal pattern of spiking progression within bursts, we next compared the normalized pattern of spiking progression in juveniles and adults (Figure 2G). Once we accounted for the overall higher instantaneous firing rate in adults by demeaning the progression pattern in juveniles and adults, the relative progression in juveniles was not distinguishable from that of the adult HVC projection neurons (Wilcoxon rank sum test, p=0.84, Figure 2H). To verify, whether the surrounding population activity is providing a stereotyped excitatory input preceding the bursts, we analyzed the membrane potential rise during 15 ms window preceding all recorded bursts. Excitatory input accounting to the membrane potential rise preceding the bursts was as stereotyped in singing juveniles as in quiet juveniles where we elicited bursts using current injection (Wilcoxon rank sum test, p=0.2105) or in singing adults (Wilcoxon rank sum test, p=0.1145, Supplementary Figure 2). This finding indicates, that the elicitation of bursts in singing juveniles occurs in a stereotyped way that is comparable to adults. The overall stereotyped characteristics of bursts (i.e. number of spikes, burst duration and stereotyped excitatory input) and the lower instantaneous firing rate indicate precise yet slower temporal dynamics in juvenile HVC projection neurons than in adults. We hypothesized, that these temporal dynamics might also be exhibited in a focal microcircuit within HVC. Here we leveraged the ability to quantify subthreshold activity in individual HVC projection neurons. This metric reflects the integrated input received by a focal neuron from its presynaptic network.

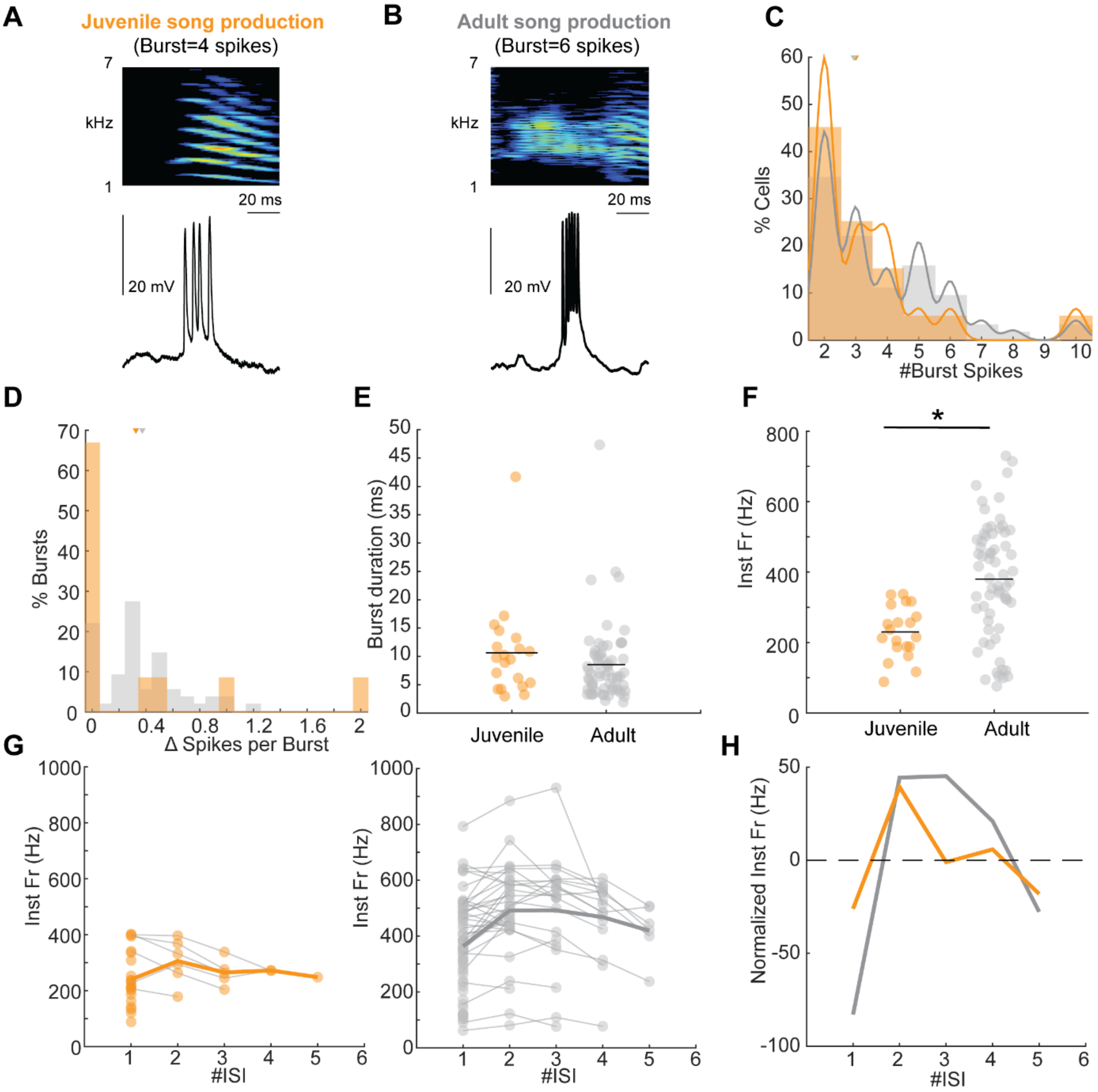


Figure 2: Temporal dynamics within HVC projection neuron bursts. A) Example of a burst in a juvenile during song production (top: spectrogram with a syllable, below: intracellular trace of a single burst), B) same as in A but for an adult bird. C) Number of spikes within a burst in juveniles and adults. D) Distribution of Δ spikes per burst in reoccurring bursts in juveniles and adults. E) Burst duration of all recorded bursts in juveniles and adults. F) Instantaneous firing rate of all bursts in juveniles and adults. G) Progression of instantaneous firing rate in juveniles (left) and adults (right), bold line: mean instantaneous firing rate per number of ISI in juveniles (orange) and adults (grey). H) Demeaned progression of instantaneous firing rate from E and F across ISI.

HVC projection neurons receive quantitatively and temporally distinct subthreshold inputs in juveniles

To quantify the temporal dynamics of inputs that HVC projection neurons are receiving, we assessed the stereotypy of the surrounding population dynamics by correlating subthreshold activity across motif renditions for each recorded neuron (Figure 3A). In juveniles, the subthreshold activity was as stereotyped as in adults (Wilcoxon rank sum test, mean±std, juveniles=0.73±0.23, median=0.82, adults=0.79±0.13, median=0.83, p=0.59, Figure 3B), suggesting a stable neural representation of song production. Since we previously reported temporal differences in syllable structure and bursting activity, we hypothesized, that these differences could also be reflected in the summation of excitatory and inhibitory postsynaptic potentials (PSPs) received by a focal neuron. To address the temporal dynamics, we quantified the duration of the PSPs and observed, that PSP events in juveniles were of longer duration than in adults (mean±std, adults=18.09±3.54, median=17.79, juveniles=21.11±2.94, median=21.04, Wilcoxon rank sum test, p=0.01 Figure 3C), which is in line with our previously observed slower temporal dynamic within bursts in juveniles. Further we quantified the membrane potential amplitude, and found that PSPs in juveniles had a higher membrane potential amplitude than in adults (mean±std, juveniles=6.37±1.2, median=6.1, adults=5.64±1.43, median=5.48, Wilcoxon rank sum test, p=0.004, Figure 3D), which could potentially be accounted to a higher resting membrane potential in juveniles[26]. In juveniles, a higher resting membrane potential might be associated with altered neuronal excitability. This could potentially manifest as changes in the frequency or amplitude of postsynaptic potential events. The frequency of the PSP element occurrence was lower in juveniles than adults (mean±std, juveniles=12.85±4.99, median=12.82, adults=16.03±4.91, median=16.12, Wilcoxon rank sum test, p=0.045, Figure 3E), suggesting a slower-paced, more sparse input from surrounding population on HVC projection neurons. The slower and more sparse membrane potential dynamics during singing in juveniles might explain the temporal differences and the increased spectral variability during song performance[4,23].

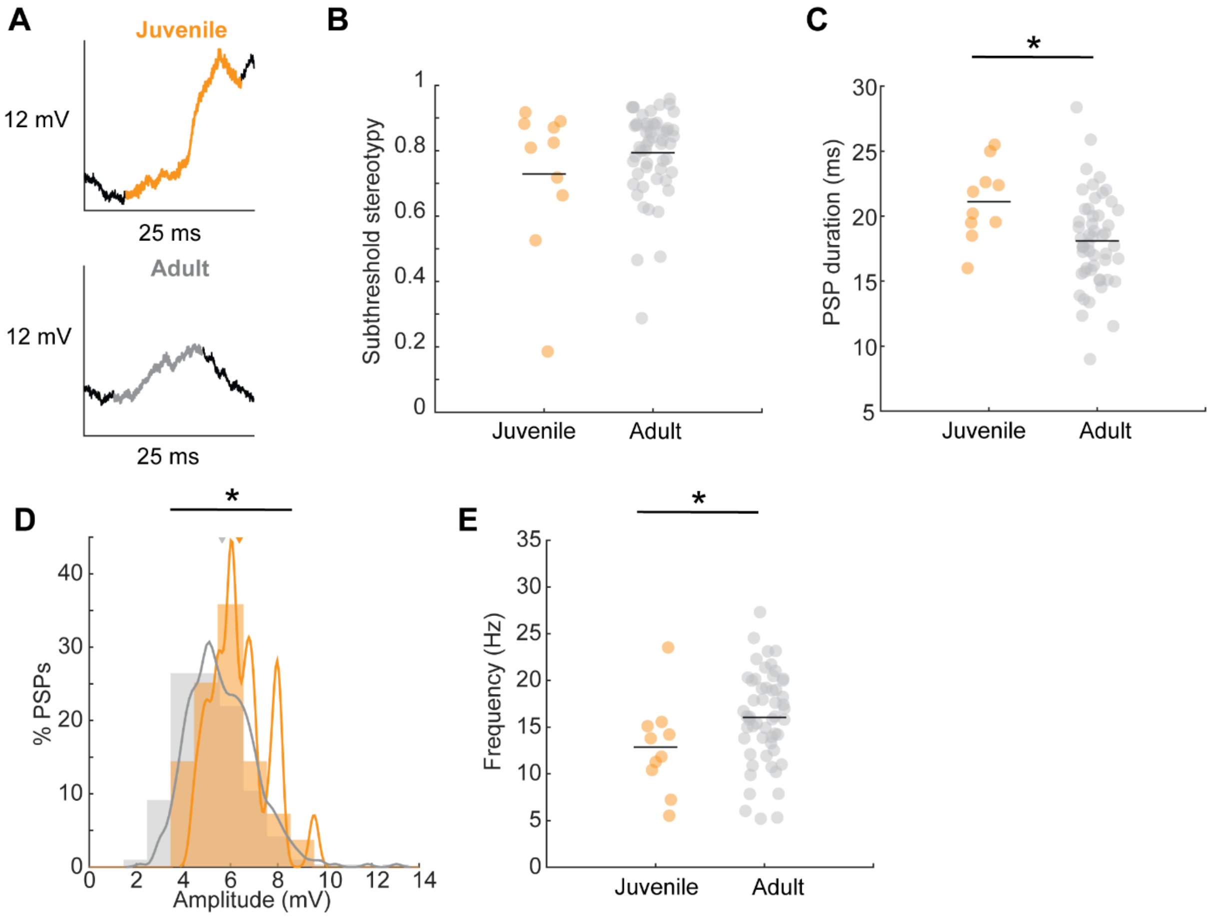


Figure 3: Temporal features of the subthreshold activity. A) Example of excitatory post-synaptic event detection in juveniles (orange) and adults (grey). B) Subthreshold activity exhibited high stereotypy in juveniles and adults. C) Mean PSP duration per HVC projection neuron. D) Distribution of amplitude of all excitatory PSP events recorded. E) Mean frequency of all PSP events per HVC projection neuron.

Discussion

In this study, we investigated the developmental shift in the temporal characteristics of song production and report a temporally distinct song from late-stage juvenile zebra finches compared to adults. Further we elucidate the neural circuitry within the HVC mediating this transformation. The song crystallization phase in late-stage juveniles is characterized by a higher number of bursts of HVC projection neurons during singing, that were more variable in their occurrence than bursts of HVC projection neurons in adult birds. Despite receiving stereotyped excitatory input, bursts in juveniles exhibited a lower firing rate than those in adults. This finding suggests that developmental changes in intrinsic neuronal properties may play a role in how temporal information is encoded within HVC. This finding was further supported by exploring the microcircuit connectivity to individual HVC projection neurons. The stereotyped long post-synaptic events occurring at a lower frequency in juveniles may indicate weaker or lesssynchronous excitatory drive, potentially explaining the lower firing rates observed in their bursts.

In adult birds, HVC_{RA} projection neurons burst sparsely during singing[14]. These bursts are triggered by presynaptic inputs with a range of latencies to form smooth neural sequences during song production[19]. In juvenile birds, the connectivity in HVC develops over time by bursting events gradually becoming more timepoint-specific during singing[11]. In this study, we report that the bursting activity in the late-stage juveniles occurred more frequently with less stereotypy across motifs than in adults, which could suggest an overall premature state of connectivity. Presumably, lack of connectivity in presynaptic input could affect the elicitation of bursts. However, we report a stereotyped excitatory input immediately preceding the bursts in singing juveniles. Although this finding does not quantify the connectivity within HVC, it highlights that the neural dynamics are stable during production of a female-directed song in late-stage juveniles. A similar observation has been reported in juveniles during the plastic song phase, where singing to a female resulted in a highly stereotyped repetitions of song, that was previously variable when the female was absent[27].

Alternatively, neural dynamics may also have distinct temporal signatures due to different membrane properties of the HVC projection neurons in juveniles and adults. During development, the membrane properties of HVC projection neurons exhibit a higher resting membrane potential and an increased spiking amplitude[26], which could lead to a higher general excitability of the HVC projection neurons. In line with this hypothesis, we report an increased amplitude of the excitatory post-synaptic events in juveniles. These excitatory post-synaptic events occurred at a lower frequency during singing in juveniles than in adults. The slower temporal dynamics could be attributed to incomplete myelination of axonal connections and changing membrane properties during development[26], that ultimately impact the signal propagation within HVC during the juvenile stage. In line with this hypothesis, studies in adults demonstrate relatively slow HVC projection neuron conduction velocities[19], suggesting a developmental trajectory where conduction speed increases over time. Describing the anatomical structure of HVC projection neurons and measuring their conductance in juveniles would provide an insight, of whether the speed of conductance changes during development.

Our findings emphasize that maturation processes within the HVC extend beyond established patterns of bursting activity. These processes are additionally reflected in the temporal dynamics of both bursts and subthreshold activity. Interestingly, these distinct temporal dynamics do not appear to disrupt the overall stereotypy of the HVC network. While the precise mechanisms underlying these temporal differences remain to be elucidated, our results raise intriguing questions about the potential role of these dynamics in the complex process of song crystallization.

Materials and Methods

Animal housing

All procedures were approved by the Regierungspräsidium Oberbayern (VET 02-21-201), Landesamt für Gesundheit und Soziales (LAGeSo Berlin) (G 0225/16) at the Freie Universität Berlin or according to Institutional Animal Care and Use Committee at New York University Langone Medical Center guidelines. Juvenile zebra finches (n=4, older than 73 dph) were housed in an aviary or a breeding cage with their genetic parents up to 60 dph, at which they were moved to an adjacent aviary where they had visual and auditory contact with their parents. Adult male birds (n=10, older than 100 dph) were acquired from a breeder.

Surgery

Male zebra finches were anesthetized with isoflurane (concentration: 1-3 % isoflurane, 97-99% oxygen) at a 65° head angle. An incision was made to expose the skull and a square shaped area of trabecular bone was removed above HVC, RA and cerebellum using a dental drill (carbide bur, FG ¼, Johnson-Promident). Nucleus RA was targeted according to coordinates (0 point at the bifurcation of the midsaggital sinus, RA coordinates: posterior 1.85 mm, lateral 2.25 mm, ventral 1.8 mm). A carbon fiber electrode (Kation Scientific, LLC) was used to identify RA based on the firing pattern [18,28]. Next, we targeted HVC using coordinates (0 point at the bifurcation of the midsaggital sinus, HVC coordinates: anterior 0.2 mm, lateral 2.3 mm, ventral 0.2 mm) and confirmed the location with antidromic stimulation from RA [14]. A previously assembled microdrive for intracellular recordings[25] was then implanted at 25° angle above HVC for intracellular recordings. Birds were let to recover 1-3 days before the intracellular recordings.

Intracellular recordings in freely behaving animals

To record intracellularly we used sharp intracellular electrodes (borosilicate glass with filament, 0.1 mm diameter), that were previously pulled using a micropipette puller (Model P-97, Sutter Instrument). Before use, each electrode was backfilled with potassium acetate (concentration: 3M). Using a silicone elastomer, we built a well around the craniotomy in HVC and filled it with phosphate buffered saline (PBS). Next, we removed dura using a dura pick. While birds were in a head-fixed setting, we lowered down the glass electrode in HVC until we could identify surrounding cells using oscilloscope and an audio monitor. We then transferred the bird into the cage of the recording setup and gradually lowered the electrode within HVC, accompanied by a brief buying pulse (10-20 ms), until a successful penetration of a neuron. Next, we presented a female bird to the male bird to motivate singing behavior. For this study, we only selected neurons that had at least 30 mV action potentials, the recording lasted at least 3 minutes and the membrane potential was below -50 mV. We could identify HVC projection neurons based on their low firing rate[25] and characteristic waveform.

Song acquisition

We briefly housed the juvenile and adult birds in sound attenuated boxes together with a female to motivate singing. Once song was recorded, we took one motif from each bird and detected syllables and gaps using threshold-based detection from the sound envelope. Durations of each segment were calculated by subtracting onset timing from the offset timing respectively.

Spike and burst detection

Spikes in each recorded trace were detected as events exceeding the 15 mV threshold from the demeaned trace. We defined bursts as single events where the firing rate of adjacent spikes exceeded 100 Hz and the spikes occurred on the same membrane depolarization event. To compare reoccurring bursts, we defined a burst as reoccurring, if the onset time of the burst (first spike) in next trial occurred ±20 ms from the onset of burst in the previous trial. Delta of spikes per burst was calculated for each identified, reoccurring burst. First, the mode of spikes per burst was calculated for each reoccurring burst. Then, the number of spikes of each burst of the same identity was subtracted from mode and an absolute mean delta of number of spikes was calculated per identified burst.

Instantaneous firing rate

Instantaneous firing rate was calculated as the mean firing rate per reoccurring burst and was defined as number of inter spike interval (ISI) divided by the sum of ISIs per burst. To make a fair comparison of the temporal progression of spiking within bursts, we only considered bursts with no more than 6 spikes per burst for this analysis.

Excitatory post-synaptic potential detection

For detection of the PSPs, we first detected local peaks and events of membrane depolarization exceeding 2 mV amplitude, then measured their amplitude and calculated the frequency by dividing the number of events with the duration of the song motif. For the analysis of PSP duration, we only included PSPs that were shorter than 35 ms to avoid false PSP detection.

Subthreshold stereotypy

Subthreshold stereotypy was defined as a cross correlation at 0 lag of all trials within a cell. First, we cut the spikes of our recorded traces (spike detection as described above). We then demeaned all subthreshold traces for a fair comparison. Each trace was then correlated with every other trace from the same HVC projection neuron.

References

1. Oller, D. K. Chapter 6 - THE EMERGENCE OF THE SOUNDS OF SPEECH IN INFANCY. in *Child Phonology* (eds. Yeni-komshian, G. H., Kavanagh, J. F. & Ferguson, C. A.) 93–112 (Academic Press, 1980). doi:10.1016/B978-0-12-770601-6.50011-5.
2. Lightfoot, D. Language acquisition and language change. *WIREs Cognitive Science* **1**, 677–684 (2010).
3. Doupe, A. J. & Kuhl, P. K. *BIRDSONG AND HUMAN SPEECH: Common Themes and Mechanisms*. www.annualreviews.org (1998).
4. Tchernichovski, O., Mitra, P. P., Lints, T. & Nottebohm, F. Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science* **291**, 2564–2569 (2001).
5. Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B. & Mitra, P. P. A procedure for an automated measurement of song similarity. *Animal Behaviour* **59**, 1167–1176 (2000).
6. Derégnaucourt, S., Poirier, C., Kant, A. V. der, Linden, A. V. der & Gahr, M. Comparisons of different methods to train a young zebra finch (Taeniopygia guttata) to learn a song. *Journal of Physiology Paris* **107**, 210–218 (2013).
7. Zann, R. A. *The Zebra Finch - A Synthesis of Field and Laboratory Studies*. 197–247 (1996).
8. Ölveczky, B. P., Otchy, T. M., Goldberg, J. H., Aronov, D. & Fee, M. S. Changes in the neural control of a complex motor sequence during learning. *J Neurophysiol* **106**, 386–397 (2011).
9. Shank, S. S. & Margoliash, D. Sleep and sensorimotor integration during early vocal learning in a songbird. *Nature* **458**, 73–77 (2009).
10. Aronov, D., Andalman, A. S. & Fee, M. S. A Specialized Forebrain Circuit for Vocal Babbling in the Juvenile Songbird. *Science* **320**, 630–634 (2008).
11. Okubo, T. S., Mackevicius, E. L., Payne, H. L., Lynch, G. F. & Fee, M. S. Growth and splitting of neural sequences in songbird vocal development. *Nature* **528**, 352–357 (2015).
12. Mackevicius, E. L., Gu, S., Denisenko, N. I. & Fee, M. S. Self-organization of songbird neural sequences 1 during social isolation 2. doi:10.1101/2022.02.18.480996.

13. Aronov, D., Veit, L., Goldberg, J. H. & Fee, M. S. Two Distinct Modes of Forebrain Circuit Dynamics Underlie Temporal Patterning in the Vocalizations of Young Songbirds. *J. Neurosci.* **31**, 16353–16368 (2011).

14. Hahnloser, R. H. R., Kozhevnikov, A. A. & Fee, M. S. An ultra-sparse code underliesthe generation of neural sequences in a songbird. *Nature 2002 419:6902* **419**, 65–70 (2002).

15. Picardo, M. A. *et al.* Population-Level Representation of a Temporal Sequence Underlying Song Production in the Zebra Finch. *Neuron* **90**, 866–876 (2016).

16. Lynch, G. F., Okubo, T. S., Hanuschkin, A., Hahnloser, R. H. R. & Fee, M. S. Rhythmic Continuous-Time Coding in the Songbird Analog of Vocal Motor Cortex. *Neuron* **90**, 877–892 (2016).

17. Leonardo, A. & Fee, M. S. Ensemble Coding of Vocal Control in Birdsong. *J. Neurosci.* **25**, 652–661 (2005).

18. Yu, A. C. & Margoliash, D. Temporal Hierarchical Control of Singing in Birds. *New Series* **273**, 1871–1875 (1996).

19. Egger, R. *et al.* Local Axonal Conduction Shapes the Spatiotemporal Properties of Neural Sequences. *Cell* **183**, 537-548.e12 (2020).

20. Wild, J. M. Functional Neuroanatomy of the Sensorimotor Control of Singing. *Annals of the New York Academy of Sciences* **1016**, 438–462 (2004).

21. Mehaffey, W. H. & Doupe, A. J. Naturalistic stimulation drives opposing heterosynaptic plasticity at two inputs to songbird cortex. *Nat Neurosci* **18**, 1272–1280 (2015).

22. Ölveczky, B. P., Andalman, A. S. & Fee, M. S. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS biology* **3**, 0902–0909 (2005).

23. Glaze, C. M. & Troyer, T. W. Development of temporal structure in zebra finch song. *J Neurophysiol* **109**, 1025–1035 (2013).

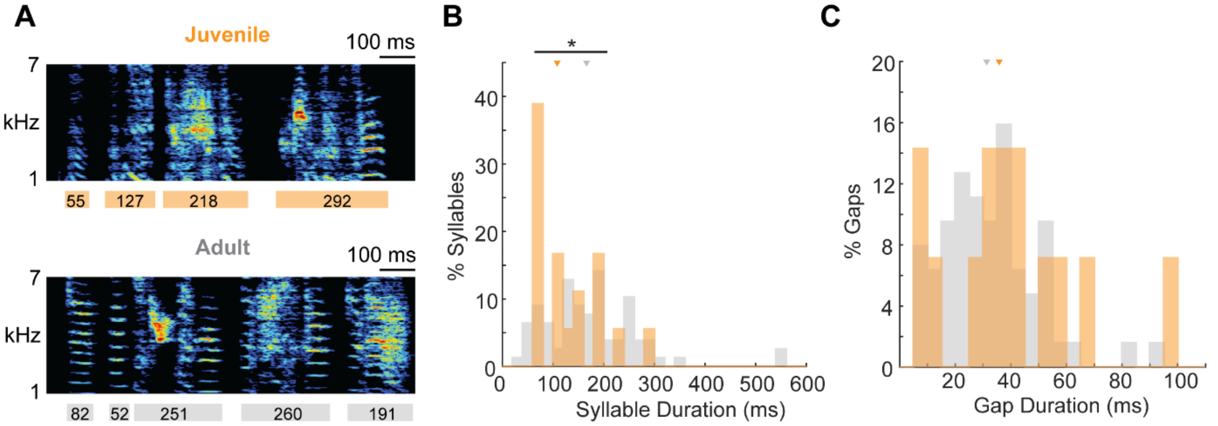
24. Long, M. A. & Fee, M. S. Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* **456**, 189–194 (2008).

25. Long, M. A., Jin, D. Z. & Fee, M. S. Support for a synaptic chain model of neuronal sequence generation. *Nature* **468**, 394–399 (2010).

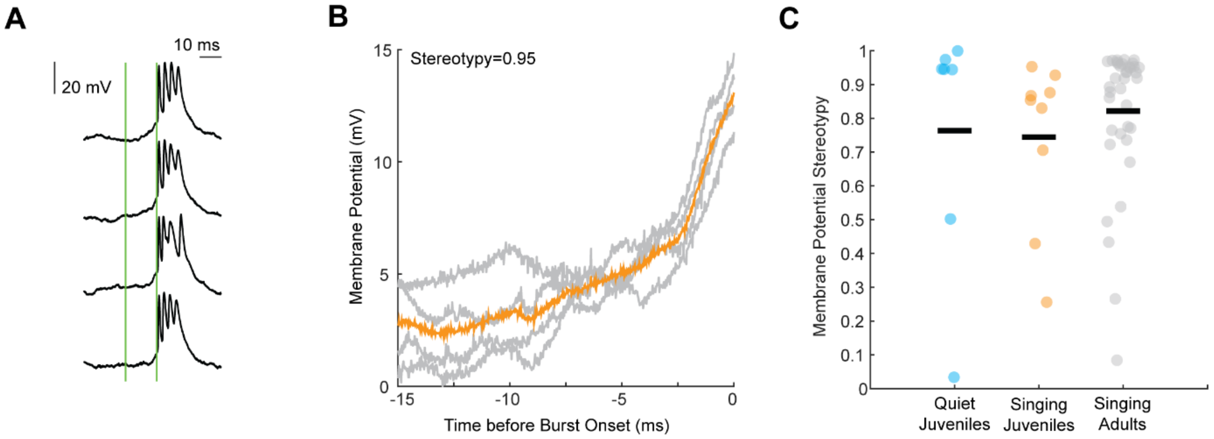
26. Ross, M. T., Flores, D., Bertram, R., Johnson, F. & Hyson, R. L. Neuronal Intrinsic Physiology Changes During Development of a Learned Behavior. *eNeuro* **4**, ENEURO.0297-17.2017 (2017).

27. Kojima, S. & Doupe, A. J. Social performance reveals unexpected vocal competency in young songbirds. *Proceedings of the National Academy of Sciences* **108**, 1687–1692 (2011).

28. Spiro, J. E., Dalva, M. B. & Mooney, R. Long-range inhibition within the zebra finch song nucleus RA can coordinate the firing of multiple projection neurons. *Journal of Neurophysiology* **81**, 3007–3020 (1999).



Supplementary Figure 1: Temporal features of juvenile and adult song. A) Example spectrograms of juvenile and adult song. Colorful bars below indicate detected syllables, numbers within the bars show syllable duration in milliseconds. B) Syllable duration of juvenile (orange) and adult (grey) birds. C) Gap distribution of adults and juveniles. Asterisks indicate mean values.



Supplementary Figure 2: Subthreshold activity preceding bursts. A) Example recording of an HVC projection neuron in a singing juvenile bird (78 days post hatch), aligned to burst onsets across four song motifs. Green lines indicate the selected interval of analysis 15 ms before burst onset. B) Subthreshold activity from the four traces (grey) of the example in A 15 ms window before burst onset. In orange – average subthreshold activity of the 15 ms interval. C) Mean subthreshold stereotypy of all pre-burst intervals per recorded neuron (quiet juveniles: mean=0.76±0.36, median=0.94; singing juveniles: mean=0.74±0.24, median=0.85; singing adults: mean=0.82±0.21, median=0.92), black lines indicate mean values of each group.

Bird ID	Days post hatch	Neuron ID
Bird 1	88	Neuron 1
Bird 1	94	Neuron 2
Bird 2	77	Neuron 3
Bird 2	78	Neuron 4
Bird 3	74	Neuron 5
Bird 3	74	Neuron 6
Bird 3	74	Neuron 7
Bird 4	87	Neuron 8
Bird 4	89	Neuron 9
Bird 4	90	Neuron 10

Table 1: Overview of intracellular recording dataset from juvenile birds. In total, intracellular recordings were obtained from 4 birds and 10 neurons. The age of the birds was between 74 and 94 days post hatch.

4 DISCUSSION

4.1 SUMMARY

In the first chapter, I investigated, whether adult female birds, that are not capable of singing, influence the song learning of juvenile male birds. First, I quantified differences in learned song of birds that were either raised with a female or alone. Both groups were tutored with the same tutor song playback, that could be elicited by pecking a key. By analyzing the spectral features and timing of tutor song and the learned song, I observed that juveniles raised with a female learned to copy the tutor song more accurately than juveniles, that were raised alone. To investigate, whether females are providing vocal feedback to juveniles during song practice, I tracked female vocal behavior during song learning phase using a small audio transmitter attached to the back of the female. In three out of five female-juvenile pairs, female birds increased their call rate during song practice across the song learning phase. In these pairs, the juveniles learned a more accurate copy of the tutor song than in the other two pairs, where female calls during song practice were not correlated with the progression of the learning phase. To explore whether female vocal feedback would also be potent to induce changes in the neural circuitry for song learning in juveniles, I performed intracellular recordings in the premotor area HVC in awake, listening juvenile birds. Projection neurons in HVC modulated their activity during the presentation of female calls alone, and also if female calls were mixed with a playback of birds own song. This finding indicated, that HVC in juvenile birds shows auditory responses not only to tutor song[112], but also to female calls. Lastly, I quantified, whether neural responses to female vocalizations could also be observed in real time during song practice, by recording intracellularly from HVC projection neurons of singing juvenile birds and perturbing their songs with a female call playback. In six out of 22 call perturbations, the HVC projection neurons changed the number of spikes during that particular timepoint in song, when compared to the same timepoint of unperturbed song. Although not all recorded HVC projection neurons were responsive to female call perturbations, the observed responses suggest a possible mechanism of female calls influencing the ongoing activity of HVC projection neurons during song production – a phenomenon, that is absent in adult birds[65]. These findings suggest, that females may guide juvenile song learning process by producing calls during song practice. Female calls in turn are a potent auditory signal to influence the neural activity of song learning circuitry in juveniles.

In the second chapter, I investigated the temporal dynamics of neural activity underlying song crystallization. I performed intracellular recordings in HVC in late-stage juveniles and adults during singing. Neural activity during song crystallization phase was characterized by a more variable occurrence of bursts and a lower instantaneous firing rate of bursts. Analysis of inter spike interval progression during bursts and the subthreshold activity immediately preceding the bursts revealed, that these aspects of bursting activity in juveniles were comparable to adult birds. Using subthreshold activity of HVC projection neurons during singing as a read-out for connected population activity revealed a high subthreshold stereotypy, but longer duration of excitatory post-synaptic events at a higher amplitude than in adult birds. These observations suggest, that neural dynamics during song crystallization phase are slower than in adults and likely serve to fine-tune the temporal features of song until a final courtship song is learned.

4.2 INFLUENCE OF EARLY LIFE EXPERIENCE ON SONG LEARNING

Although tutoring juvenile birds from 30 days post hatch with a tutor song playback is common in the birdsong neuroscience research field, the song learning process might be influenced by the early auditory experiences the juvenile birds had with their genetic fathers or other adult tutor birds[11]. A previous study found, that juveniles, that are exposed to a tutor song for as little as the first 35 days of their life were able to learn a good copy of the song[11]. In the first chapter, the juveniles were raised by both parents before the onset of the song training. To control for the influence the early life experience might have had on the song learning outcome, I compared the learned songs of the juveniles with songs of their genetic fathers. Birds that were raised alone tended to copy their father's song (5 out of 6 birds) rather than the tutor song. In contrast, birds that were raised together with a female copied the tutor song irrespective of early exposure to their genetic father. This finding indicates, that social interactions may play a defining role in song learning. Juvenile birds that copied the song of their genetic fathers likely were motivated by the social interactions with their fathers, whereas juveniles that were raised with a female experienced social interaction before and throughout the song learning phase. The constant exposure to social interactions might have therefore led to a more proportional outcome of the song learning – since juveniles were exposed to a tutor song playback longer (~ 60 days) than to their genetic father (~ 35 days), they were more likely to copy the tutor song playback. These results are in line with previous studies and suggest that social interactions are essential for the choice of the song model[6,8,10,47,124].

4.3 FEMALE ZEBRA FINCHES AS SOCIAL PARTNERS DURING SONG LEARNING

Zebra finch juveniles will learn to copy the tutor song most accurately if they are tutored with a live male tutor[3]. In a highly social species like the zebra finches that naturally live in large flocks, social interactions with other conspecifics are a crucial part in the life of any zebra finch. In line with this idea, zebra finch juveniles will choose to copy the song of an adult male, with whom they have most social interactions with[6,8,10]. Whether social interactions alone can aid juvenile song learning from a song template was so far unknown. Female zebra finches provide the ideal case for studying social influence on song learning, as they do not sing yet they share other vocalizations (‘calls’) with males. In this scenario, the exposure to tutor song can be disentangled from social interactions between the juvenile and the female.

4.3.1 Influence of female song preferences in guiding juvenile song learning

Female zebra finches have preferences for subtle acoustic features of male songs[52], and they show preference towards songs that they were exposed to during early development[30]. On the one hand, this behavior is advantageous to the training paradigm that I used to train the juvenile zebra finches with, since females might show positive feedback to more stereotyped, adult-like songs. On the other hand, the early auditory experience of the company females, that were selected for the experiments, might bias them to reinforce juvenile songs, that may be more similar to their father or mate. The company females for the song learning experiment were randomly picked from the breeding colony and it is difficult to say, which song features in particular did they prefer, as the information about their song preferences, their mate’s song or their father’s song was not available at the time of the experiment.

One method to assess female preference to a song is to quantify the number of calls a female produces as a response to a song playback[53]. Female and male zebra finches produce several different types of calls, that are used in a variety of contexts[35], and can be used for individual recognition[36]. In my first chapter, I did not distinguish between the different call types a female would produce during juvenile song practice. It is possible, that females could signal a preferred version of juvenile song by producing one type of a call, whereas they could use a different call type to indicate a song that they preferred less. One way to investigate, whether female call types are indicative of quality of song rendition would be to assess the juvenile songs, that were followed by a female call less than 1 s after the offset of the last syllable. The female call type should be noted and a similarity score between the last song motif before female call and the final learned song should be quantified. This analysis would additionally provide insight, into whether females use different call types in close relation to juvenile song.

4.3.2 Do females guide juveniles to a specific song template?

As female birds aid the juvenile song learning process, an interesting consideration is what are the females guiding the juvenile song practice towards. Several possibilities should be considered: 1) females could guide juveniles to achieve the most accurate tutor song copy, 2) females may guide juveniles to produce a song that fits the species characteristics of an attractive song and this process might be influenced by female song preferences, and 3) females have more of a ‘’cheer-leading’’ function in song learning and they reinforce song practice whenever the juvenile male sings, independent of the quality of that particular song rendition. Although it cannot be excluded, that females serve as cheerleaders to juvenile song practice, this hypothesis is not supported by the observations from the female-vocalization tracking experiment – in the two cases where female vocalizations occurred randomly and were not correlated with the progression of song learning phase, the juvenile males learned a less accurate copy of the tutor song than in the cases where female vocalizations occurred in a correlated fashion with the song learning phase. To address the first two hypotheses, as mentioned previously, the information about the song features that these females may have preferred was missing. The tutors’ song playback that was used in the song learning experiment was a computerized zebra finch song that did not exist in the breeding colony. The syllable rate of this playback was slower than the syllable rate of most songs that were recorded from the breeding colony. Thus, higher similarity to tutor song in the juveniles that were raised with females might suggest the females guiding the juvenile males to learn a more accurate copy of the tutor song, rather than guiding them towards song features that are already represented in the breeding colony. To analyze this hypothesis more precisely, company females should be raised in a controlled acoustic environment, where they are tutored with a single song playback. Before the tutoring and after reaching adulthood, these females should be tested for preference of the song playback they were raised with and conspecific songs. Among the tutoring experience, other latent features that may influence song preference, i.e. the acoustic properties of song based on males' early experience, should be considered[125]. Females that show preference for the song they were raised with should then be paired with juveniles in song training. The learned song of the juvenile should be compared to the tutor song playback and the song playback that the females were exposed to, to determine which song features did the juvenile implement in his own song.

4.3.3 Social and genetic relation between juveniles and females may affect social feedback during song learning phase

Another factor to consider is the social relation between the females and juveniles. Presumably, the motivation of a genetic mother to help her offspring in song learning is different from a conspecific adult female that is paired with an unfamiliar juvenile during song learning phase. To test this idea, sibling juvenile males could be trained with a tutor song playback in two different social conditions – either with their genetic mother or a conspecific female. If juveniles that were raised with their own mother learn a more accurate copy of the tutor song, it would indicate that the motivation to aid juvenile song learning is also dependent on the social relation between the juvenile and the female.

4.3.4 Female birds may use other social feedback mechanisms to guide juvenile song learning

Apart from vocal feedback, female birds could potentially also use visual feedback to reward preferred versions of juvenile song. In a previous study, exposing juveniles to a female visual display (fluff-up) during song practice led these juveniles to learn a better copy of the tutor song than juveniles that were exposed to the same visual display randomly throughout the day[13]. The sound-isolating boxes in which the juvenile birds with females were raised, were not equipped with cameras, so it is unfortunately not possible to conclude, whether the company females engaged in other social feedback mechanisms than the vocalizations. However, both the increase of female calls during song practice across the song learning phase as well as the neural responses to female calls in HVC of juvenile birds suggest that apart from other possible social reinforcement mechanisms, female calls may be used to aid juvenile song learning process.

4.3.5 How might social interactions between a female and a juvenile form a basis for social feedback during song practice?

An interesting consideration is how might the observed social interactions between a female and the juvenile bird improve the song copy over time. One hypothesis would be, that female birds call in response to a preferred rendition of the juvenile song[53] which in turn acts as a positive reinforcement signal for the juvenile, since female calls are a rewarding signal in adults[64]. To test this hypothesis, two factors could be analyzed: 1) whether females call more to a juvenile song rendition that has a higher similarity to tutor song, and 2) whether female call during a particular syllable leads to a higher song similarity on the next song rendition. In the first chapter, I addressed the second factor by analyzing juvenile songs with female call interruption. In particular, I analyzed 300 ms of a song segment after the female call and compared it to similar song segments without a female call of the same day. I did not observe any differences between these two groups of songs, which could indicate, that female calls may not instantly affect the quality of the next practice song, but rather this effect might be implemented over several days. Another possibility is, that females produce calls to a targeted syllable of the juvenile song. Because juveniles practice the song hundreds of times per day, female birds could assess which syllables of the song may need more practice, and produce calls when the juvenile is singing that particular syllable. The analysis of this hypothesis is challenging, since it is not possible to analyze the similarity of a syllable, when it is simultaneously overlaid with a female call. It remains an interesting question to observe, whether the timing of female call during juvenile song practice is meaningful to improve specific parts of song and whether within the same day females call more to renditions of juvenile songs that have a higher similarity to tutor song.

4.3.6 Social interactions with other conspecifics

In the first chapter, I focused on investigating whether females contribute to song learning process. Although the reason for using females was to dissect social interactions between a juvenile and a conspecific from social interactions between a juvenile and a male tutor, the question remains how other social companions might aid song learning. Apart from the tutor, siblings could also be social companions of the learning juvenile. A previous study reported the more male siblings are learning from one tutor, the less accurate is the copy of the tutor song[12]. In contrast, blind-folded juvenile zebra finches learn a better copy of the tutor song, if accompanied with a female sibling[14]. In the first chapter, in one of the experimental groups juvenile birds were raised in social isolation, which could potentially affect their stress level and therefore impair the song learning process[126]. I aimed to do a controlled experiment for social interactions during vocal learning, which included not exposing the male juvenile bird to any other song than the tutor song playback. Therefore, it was not possible to accompany the learning juvenile with other male siblings. To disentangle how female-specific song preferences might influence song learning process, juvenile birds could be accompanied with adult females that were raised in song isolation and are naïve to song. These females would likely not show any preferences to acoustic features of song, since the early exposure to the song is missing[53]. The role of female song preference behavior in aiding juvenile song learning process should be further explored, but so far remains elusive.

4.4 REPRESENTATIONS OF SOCIAL FEEDBACK IN THE NEURAL CIRCUIT

4.4.1 Involvement of HVC in song learning

In singing adult birds $\mathrm{HVC}_{\mathrm{RA}}$ neurons produce a sparse burst of action potentials during a specific timepoint in song[16]. $\mathrm{HVC}_{\mathrm{Area\ X}}$ neurons in turn produce either one or several bursts during song and send the information of song timing to the anterior forebrain pathway[127]. The general function of HVC is the control of the timing of song[72], and although directed song in adults is stereotyped, variability in song timing exists in undirected song. As the adult bird is singing several motifs within the same song bout of undirected song, the duration of the motifs gradually increases towards the end of the song bout[128]. Within the variability of motif duration, the gap durations between syllables vary more than the syllable durations[128]. In a targeted auditory feedback paradigm adult males are also able to modify the duration of a specific song syllable[129,130], the lengthening in duration of the targeted syllable does not affect the duration of other syllables in song[129,130].

During development, the temporal features of the juvenile song are more variable than in adulthood[41]. As the song learning progresses during the plastic song phase, the gap duration decreases whereas the syllable duration remains stereotyped[41]. In the second chapter, I report shorter syllable durations in adults than in juveniles, whereas the gap duration was not distinguishable. Additionally, I also report a higher variability of burst occurrence in juveniles, which is accompanied by a lower instantaneous firing rate of bursts in HVC projection neurons in juveniles than in adults. These observations could be explained by a premature state of song production network within HVC. One hypothesis of how synaptic chain-like activity in HVC forms is through consistent growth and splitting of an early chain[19]. In this paradigm, HVC is at a random state of connectivity at the beginning of the development. During song production, upstream areas from HVC send excitatory input during syllable onsets, from which the chain-like activity starts to emerge. During subsequent emergence of protosyllables, the bursting activity in HVC becomes rhythmic. New syllables evolve from the previous protosyllables and the bursting activity within HVC becomes specific to a unique timepoint in song motif[19]. My observation of lower instantaneous firing rate of bursts could be an indication for slower temporal processing within HVC in juveniles than in adults. This hypothesis is further supported by the observation of longer duration excitatory post-synaptic events in HVC projection neurons of juveniles. This effect could either be due to a lower state of connectivity within HVC or different physiological properties of the neurons in a developing neural circuit[131]. Although the number of recorded neurons per each bird and the number of juvenile birds in these experiments was small, my findings provide an insight into the neural dynamics in HVC during song crystallization phase. Future large-scale extracellular recordings in HVC of singing juveniles would yield additional information about how the fine-tuning of song is represented in HVC interneurons. What influence do female calls have on HVC circuitry?

In the first chapter, I report auditory responses to female calls in HVC in juvenile zebra finches. I observed auditory responses in HVC in three different contexts: 1) during listening to female calls surrounded by silence, 2) during listening to female calls mixed in a playback of birds own song and 3) during song production. Additionally, I also report auditory responses to female calls surrounded by silence in awake, listening adult male birds. This finding is in line with a previous study, that found auditory responses to female calls in HVC of adult males during male-female call interactions[132], suggesting that female calls are a behaviorally relevant auditory signal. In contrast, when adult birds sing, HVC projection neurons are not responsive to female call perturbations[65]. The nature of auditory responses in HVC in zebra finches is state dependent[133–135]. In anesthetized adult birds, HVC shows auditory responses to birds own song, but these responses strongly depend on the temporal structure of song[136]. In awake adults, the spiking responses to a playback of birds own song are abolished[135], but auditory responses can be detected on a subthreshold level[137]. Pharmacological modulation of neural activity of HVC interneurons can reinstate previously absent auditory responses in HVC of adult birds[112]. This finding suggests, that auditory responses in HVC are gated by the local inhibitory interneurons. Since in juvenile birds the excitatory and inhibitory populations of neurons in HVC are at a premature state of connectivity, it would explain why I was able to observe auditory responses to female calls in the excitatory HVC projection neurons in singing juveniles.

A crucial question to investigate is how female calls may have a mechanistic effect on the neural circuitry for song production in juvenile birds. In the first chapter, I was recording from HVC projection neurons during song production, but their identity ($\mathrm{HVC}_{\mathrm{RA}}$ or $\mathrm{HVC}_{\mathrm{Area\ X}}$ projection neurons) in most cases was unfortunately not confirmed with an antidromic stimulation due to the complexity of the experiment itself. It is therefore not possible to distinguish, whether both projection neuron classes exhibited auditory responses to female calls during song production. In adult birds, $\mathrm{HVC}_{\mathrm{Area\ X}}$ projection neurons are thought to provide information about the immediate motor-related signal and not auditory information to the anterior forebrain pathway[137]. This finding is in line with the absence of auditory responses of $\mathrm{HVC}_{\mathrm{Area\ X}}$ neurons to white noise perturbation during singing in juveniles[63] and adults[137]. $\mathrm{HVC}_{\mathrm{RA}}$ neurons on the other hand are involved in relaying information about timing of the song to the downstream motor nucleus RA, however premotor $\mathrm{HVC}_{\mathrm{RA}}$ neurons in juveniles also show auditory responses to tutor song[112]. These findings suggest, that HVC projection neurons may also respond to socially relevant stimuli in juveniles.

To provide a perspective to my findings, I would like to hypothesize, that during vocal development, the information from upstream auditory areas to HVC is being relayed directly on HVC projection neurons due to lack of gating-off of the auditory responses of inhibitory interneurons. In turn, excitatory HVC projection neurons show responses to a female call during song production and may inform RA and the anterior forebrain pathway about a socially valent feedback, that was received at a particular timepoint of song. Within the anterior forebrain pathway, the activity in Area X might simultaneously be modulated by a release of dopamine during the female call as well as receive the information about the current variation of song from LMAN, allowing the neural circuitry of the song system to possibly adjust the song performance over next renditions of song.

Although this study does not provide any information about how the social feedback is directly integrated in the neural circuitry during development, it provides an interesting research area to explore. For example, functional Calcium imaging of HVC in juveniles during presentation of female call playbacks may yield more insight of how the female call is integrated in the population activity of HVC. Additionally, an interesting idea would be to distinguish whether a juvenile male is listening to a female call playback through the speaker or being exposed to a female bird emitting a call. I would expect, that exposure to a female emitting a call should have a greater auditory response in HVC than hearing a female call outside of the social context. Lastly, it has previously been shown that exposure to a singing male tutor leads to a release of dopamine in HVC, which ultimately leads to a better copy of the tutor song[4]. Although in this study the authors did not observe increased activity of dopaminergic neurons projecting to HVC when the juvenile was exposed to a female bird, they did not address the question of how these neurons might have responded to a female bird calling. Whether female calls can also trigger dopamine release in Area X or in HVC in juveniles remains unknown.

4.5 CONCLUSIONS AND OUTLOOK

The results of my thesis are novel in highlighting how social feedback could aid vocal learning process. Although previous studies had established, that social interactions aid vocal learning, they did not disentangle the social interactions from the abundance of a role model for tutor song. To distinguish whether social interactions alone can aid goal-directed vocal learning, this study explored the benefit of training juvenile male zebra finches with a song playback, while accompanied with non-singing adult females during the song learning phase. Additionally, in the first chapter of my thesis I suggest a potential social feedback mechanism that may aid juvenile song learning – female calls. I further show that female calls have the potential to modulate the neural activity of HVC projection neurons during singing in juveniles. Although more in-depth experiments are needed to pinpoint whether female calls alone are sufficient to aid juvenile song learning and whether female calls are a reinforcement signal to juveniles, this study for the first time illustrates a possible social feedback mechanism during vocal learning in zebra finches. In the second chapter, I describe the neural dynamics underlying song crystallization in juveniles, by for the first time describing both the spiking and subthreshold activity of intracellularly recorded HVC projection neurons during song production in juveniles. The findings of the second chapter suggest distinct temporal neural dynamics during the song crystallization phase and inspire more questions as to how these temporal dynamics may fine-tune spectral and temporal features of the song syllables.

4.5.1 Are female calls alone sufficient to guide juvenile song learning process?

One potential question to answer is whether female calls alone can aid juvenile song learning. To investigate this question, juvenile birds should be raised in social isolation, but at the same time trained with a tutor song playback. Based on my previous results, a playback of a female call should be elicited during juvenile song practice. The number of calls per juvenile song practice should be increased as the learning phase progresses. If female calls alone are sufficient to aid juvenile song learning, this type of auditory feedback may already lead to a more accurate copy of the tutor song.

4.5.2 Do female calls act as a positive reinforcement to juveniles?

Another interesting question is whether female calls serve as a reinforcement signal to the male juveniles. If female calls are rewarding, I would expect to observe a release of dopamine in HVC, since HVC receives dopaminergic projections from PAG[4]. To test this idea, I suggest to express a virus that encodes dopamine sensors in HVC. After the virus has been expressed, 2-photon imaging should be done in HVC in juveniles during the playback of a female call. An increase of dopamine release in HVC would indicate, that female calls can serve as a rewarding signal in juvenile males and potentially aid the song learning process.

4.5.2.1 Do female calls lead to strengthening of neural sequences in HVC?

Lastly it should be explored, whether female calls affect the forming of neural patterns in HVC during song learning. To answer this question, I propose to perform functional calcium imaging in nucleus HVC in singing juvenile birds. Juvenile males should be trained in social isolation with a tutor song playback. At 60 days post hatch, the calcium imaging in HVC should be performed during song production to identify consecutively activated HVC projection neurons[107]. Next, juvenile males should 1) either be exposed to female calls as a feedback during their song practice or 2) continue the song training without female calls. Daily calcium imaging in HVC should be performed during song production to detect, whether the consecutive activation of HVC projection neurons strengthens as a result of female call feedback, or whether the strengthening is weak, as presumably in the absence of female call feedback. The song of the juveniles should also be compared to the tutor song daily, to see whether female call feedback during song practice also leads to increased song similarity to tutor song. The findings of this experiment could potentially provide an insight into how female calls may affect the song learning circuitry within HVC.

In summary, my thesis on the influence of social interactions on vocal learning suggests the existence of a social feedback mechanism. This work stimulates new questions about the precise neural and behavioral mechanisms by which positive reinforcement shapes vocal production.

5 REFERENCES

1. Curtiss, S. *Genie: A Psycholinguistic Study of a Modern-Day Wild Child*. (Academic Press, 2014).
2. Janik, V. M. & Knörnschild, M. Vocal production learning in mammals revisited. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376**, 20200244 (2021).
3. Derégnaucourt, S., Poirier, C., Kant, A. V. der, Linden, A. V. der & Gahr, M. Comparisons of different methods to train a young zebra finch (Taeniopygia guttata) to learn a song. *Journal of Physiology Paris* **107**, 210–218 (2013).
4. Tanaka, M., Sun, F., Li, Y. & Mooney, R. A mesocortical dopamine circuit enables the cultural transmission of vocal behaviour. *Nature* **563**, 117–120 (2018).
5. Eales, L. A. *Song Learning in Zebra Finches: Some Effects of Song Model Availability on What Is Learnt and When*. *Anim Behav* 33 (1985).
6. Jones, A. E. & Slater, P. J. B. The Role of Aggression in Song Tutor Choice in the Zebra Finch: Cause or Effect? *Behaviour* **133**, 103–115 (1996).
7. Tchernichovski, O., Mitra, P. P., Lints, T. & Nottebohm, F. Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science* **291**, 2564–2569 (2001).
8. Williams, H. Models for song learning in the zebra finch: fathers or others? *Animal Behaviour* **39**, 745–757 (1990).
9. Immelmann K. Song development in the zebra finch and other estrildid finches | Semantic Scholar. *Bird Vocalizations* (1969).
10. Clayton, N. S. Song tutor choice in zebra finches. *Animal Behaviour* **35**, 714–721 (1987).
11. Böhner, J. Early acquisition of song in the zebra finch, Taeniopygia guttata. *Animal Behaviour* **39**, 369–374 (1990).
12. Tchernichovski, O. & Nottebohm, F. Social inhibition of song imitation among sibling male zebra finches. *Proceedings of the National Academy of Sciences* **95**, 8951–8956 (1998).

13. Carouso-Peck, S. & Goldstein, M. H. Female Social Feedback Reveals Non-imitative Mechanisms of Vocal Learning in Zebra Finches. *Current Biology* **29**, 631-636.e3 (2019).

14. Adret, P. Vocal imitation in blindfolded zebra finches (Taeniopygia guttata) is facilitated in the presence of a non-singing conspecific female. *Journal of Ethology* **22**, 29–35 (2004).

15. Zann, R. A. *The Zebra Finch - A Synthesis of Field and Laboratory Studies*. 197–247 (1996).

16. Hahnloser, R. H. R., Kozhevnikov, A. A. & Fee, M. S. An ultra-sparse code underliesthe generation of neural sequences in a songbird. *Nature 2002 419:6902* **419**, 65–70 (2002).

17. Picardo, M. A. *et al.* Population-Level Representation of a Temporal Sequence Underlying Song Production in the Zebra Finch. *Neuron* **90**, 866–876 (2016).

18. Lynch, G. F., Okubo, T. S., Hanuschkin, A., Hahnloser, R. H. R. & Fee, M. S. Rhythmic Continuous-Time Coding in the Songbird Analog of Vocal Motor Cortex. *Neuron* **90**, 877–892 (2016).

19. Okubo, T. S., Mackevicius, E. L., Payne, H. L., Lynch, G. F. & Fee, M. S. Growth and splitting of neural sequences in songbird vocal development. *Nature* **528**, 352–357 (2015).

20. Aronov, D., Andalman, A. S. & Fee, M. S. A Specialized Forebrain Circuit for Vocal Babbling in the Juvenile Songbird. *Science* **320**, 630–634 (2008).

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

22. Robinson, S. R. Antipredator behaviour and predator recognition in Belding's ground squirrels. *Animal Behaviour* **28**, 840–852 (1980).

23. Kim, E. J., Kim, E. S., Covey, E. & Kim, J. J. Social Transmission of Fear in Rats: The Role of 22-kHz Ultrasonic Distress Vocalization. *PLOS ONE* **5**, e15077 (2010).

24. Levréro, F. *et al.* Social bonding drives vocal exchanges in Bonobos. *Sci Rep* **9**, 711 (2019).

25. Janik, V. M. & Slater, P. J. B. *Vocal Learning in Mammals*. *ADVANCES IN THE STUDY OF BEHAVIOR* vol. 26 (1997).

26. Petkov, C. I. & Jarvis, E. D. Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience* **4**, (2012).

27. Janik, V. M. & Slater, P. J. B. The different roles of social learning in vocal communication. *Animal Behaviour* **60**, 1–11 (2000).

28. Jones, A. E., ten Cate, C. & Slater, P. J. B. Early Experience and Plasticity of Song in Adult Male Zebra Finches (Taeniopygia guttata). *Journal of Comparative Psychology* **110**, 354–369 (1996).

29. Williams, H., Kilander, K. & Sotanski, M. L. Untutored song, reproductive success and song learning. *Animal Behaviour* **45**, 695–705 (1993).

30. Riebel, K. Early exposure leads to repeatable preferences for male song in female zebra finches. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **267**, 2553–2558 (2000).

31. Zann, R. Structure, Sequence and Evolution of Song Elements in Wild Australian Zebra Finches. *The Auk* **110**, 702–715 (1993).

32. Rajan, R. & Doupe, A. J. Behavioral and Neural Signatures of Readiness to Initiate a Learned Motor Sequence. *Current Biology* **23**, 87–93 (2013).

33. Johnson, F., Soderstrom, K. & Whitney, O. Quantifying song bout production during zebra finch sensory-motor learning suggests a sensitive period for vocal practice. *Behavioural Brain Research* **131**, 57–65 (2002).

34. Sossinka, R. & Böhner, J. Song Types in the Zebra Finch Poephila guttata castanotis. *Zeitschrift für Tierpsychologie* **53**, 123–132 (1980).

35. Elie, J. E. & Theunissen, F. E. The vocal repertoire of the domesticated zebra finch: a data-driven approach to decipher the information-bearing acoustic features of communication signals. *Anim Cogn* **19**, 285–315 (2016).

36. Elie, J. E. & Theunissen, F. E. Zebra finches identify individuals using vocal signatures unique to each call type. (1234) doi:10.1038/s41467-018-06394-9.

37. Zann, R. Ontogeny of the Zebra Finch Distance Call: I. Effects of Cross-fostering to Bengalese Finches. *Zeitschrift für Tierpsychologie* **68**, 1–23 (1985).

38. Benichov, J. I. *et al.* The Forebrain Song System Mediates Predictive Call Timing in Female and Male Zebra Finches. *Current Biology* **26**, 309–318 (2016).

39. Leitão, A. & Gahr, M. Babbling opens the sensory phase for imitative vocal learning. *Proceedings of the National Academy of Sciences* **121**, e2312323121 (2024).

40. Lipkind, D. *et al.* Songbirds work around computational complexity by learning song vocabulary independently of sequence. *Nat Commun* **8**, 1247 (2017).

41. Glaze, C. M. & Troyer, T. W. Development of temporal structure in zebra finch song. *J Neurophysiol* **109**, 1025–1035 (2013).

42. Ravbar, P., Lipkind, D., Parra, L. C. & Tchernichovski, O. Vocal Exploration Is Locally Regulated during Song Learning. *J. Neurosci.* **32**, 3422–3432 (2012).

43. Fehér, O., Wang, H., Saar, S., Mitra, P. P. & Tchernichovski, O. De novo establishment of wild-type song culture in the zebra finch. *Nature* **459**, 564–568 (2009).

44. Araki, M., Bandi, M. M. & Yazaki-Sugiyama, Y. Mind the gap: Neural coding of species identity in birdsong prosody. *Science* **354**, 1282–1287 (2016).

45. James, L. S., Wang, A. S., Bertolo, M. & Sakata, J. T. Learning to pause: Fidelity of and biases in the developmental acquisition of gaps in the communicative signals of a songbird. *Developmental Science* **26**, (2023).

46. Tchernichovski, O., Lints, T., Mitra, P. P. & Nottebohm, F. Vocal imitation in zebra finches is inversely related to model abundance. *Proceedings of the National Academy of Sciences* **96**, 12901–12904 (1999).

47. Eales, L. A. The influences of visual and vocal interaction on song learning in Zebra finches. *Animal Behaviour* **37**, 507–508 (1989).

48. Vyas, A., Harding, C., Borg, L. & Bogdan, D. Acoustic characteristics, early experience, and endocrine status interact to modulate female zebra finches' behavioral responses to songs. *Hormones and Behavior* **55**, 50–59 (2009).

49. Nottebohm, F. & Arnold, A. P. Sexual dimorphism in vocal control areas of the songbird brain. *Science (New York, N.Y.)* **194**, 211–213 (1976).

50. Shaughnessy, D. W., Hyson, R. L., Bertram, R., Wu, W. & Johnson, F. Female zebra finches do not sing yet share neural pathways necessary for singing in males. *Journal of Comparative Neurology* **527**, 843–855 (2019).

51. Konishi, M. & Akutagawa, E. Neuronal growth, atrophy and death in a sexually dimorphic song nucleus in the zebra finch brain. *Nature* **315**, 145–147 (1985).

52. Woolley, S. C. & Doupe, A. J. Social context-induced song variation affects female behavior and gene expression. *PLoS Biology* **6**, 0525–0537 (2008).

53. Chen, Y., Clark, O. & Woolley, S. C. Courtship song preferences in female zebra finches are shaped by developmental auditory experience. *Proceedings of the Royal Society B: Biological Sciences* **284**, (2017).

54. Clayton, N. S. Song discrimination learning in zebra finches. *Animal Behaviour* **36**, 1016–1024 (1988).

55. Wei, J., Liu, Q. & Riebel, K. Generalisation of early learned tutor song preferences in female zebra finches (Taeniopygia guttata) 2. doi:10.1101/2022.05.05.490783.

56. Miller, D. B. Long-term recognition of father's song by female zebra finches. *Nature* **280**, 390–391 (1979).

57. Nordeen, K. W. & Nordeen, E. J. Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behavioral & Neural Biology* **57**, 58–66 (1992).

58. Lombardino, A. J. & Nottebohm, F. Age at Deafening Affects the Stability of Learned Song in Adult Male Zebra Finches. *J Neurosci* **20**, 5054–5064 (2000).

59. Leonardo, A. & Konishi, M. Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature* **399**, 466–470 (1999).

60. Gadagkar, V. *et al.* Dopamine neurons encode performance error in singing birds. *Science* **354**, 1278–1282 (2016).

61. Andalman, A. S. & Fee, M. S. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 12518–12523 (2009).

62. Tumer, E. C. & Brainard, M. S. Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. *Nature* **450**, 1240–1244 (2007).

63. Kozhevnikov, A. A. & Fee, M. S. Singing-related activity of identified HVC neurons in the zebra finch. *Journal of Neurophysiology* **97**, 4271–4283 (2007).

64. Roeser, A. *et al.* Dopaminergic error signals retune to social feedback during courtship. *Nature* (2023) doi:10.1038/s41586-023-06580-w.

65. Vallentin, D. & Long, M. A. Motor origin of precise synaptic inputs onto forebrain neurons driving a skilled behavior. *Journal of Neuroscience* **35**, 299–307 (2015).

66. Nottebohm, F., Stokes, T. M. & Leonard, C. M. *Central Control of Song in the Canary, Serinus Canarius*.

67. Kelley ' And, D. B. & Nottebohm, F. *Projections of a Telencephalic Auditory Nucleus-Field L-in the Canary*.

68. Stokes, T. M., Leonard, C. M. & Nottebohm, F. The telencephalon, diencephalon, and mesencephalon of the canary, Serinus canaria, in stereotaxic coordinates. *Journal of Comparative Neurology* **156**, 337–374 (1974).

69. Vicario, D. S. Organization of the zebra finch song control system: II. Functional organization of outputs from nucleus Robustus archistriatalis. *J Comp Neurol* **309**, 486–494 (1991).

70. Yu, A. C. & Margoliash, D. Temporal Hierarchical Control of Singing in Birds. *New Series* **273**, 1871–1875 (1996).

71. Kosche, G., Vallentin, D. & Long, M. A. Interplay of Inhibition and Excitation Shapes a Premotor Neural Sequence. *J. Neurosci.* **35**, 1217–1227 (2015).

72. Long, M. A. & Fee, M. S. Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* **456**, 189–194 (2008).

73. Roberts, T. F. *et al.* Identification of a motor to auditory pathway important for vocal learning. *Nat Neurosci* **20**, 978–986 (2017).

74. Spiro, J. E., Dalva, M. B. & Mooney, R. Long-range inhibition within the zebra finch song nucleus RA can coordinate the firing of multiple projection neurons. *Journal of Neurophysiology* **81**, 3007–3020 (1999).

75. Vu, E. T., Mazurek, M. E. & Kuo, Y.-C. *Identification of a Forebrain Motor Programming Network for the Learned Song of Zebra Finches*. *The Journal of Neuroscience* vol. 14 6924–6934 (1994).

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

77. Hessler, N. A. & Doupe, A. J. Singing-Related Neural Activity in a Dorsal Forebrain–Basal Ganglia Circuit of Adult Zebra Finches. *J. Neurosci.* **19**, 10461–10481 (1999).

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

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

80. Fee, M. S. & Goldberg, J. H. A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience* **198**, 152–170 (2011).

81. Sohrabji, F., Nordeen, E. J. & Nordeen, K. W. Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. *Behav Neural Biol* **53**, 51–63 (1990).

82. Scharff, C. & Nottebohm, F. *A Comparative Study of the Behavioral Deficits Following Lesions of Various Parts of the Zebra Finch Song System: Implications for Vocal Learning*. *The Journal of Neuroscience* vol. 7 2898–2913 (1991).

83. Hessler, N. A. & Doupe, A. J. Social context modulates singing-related neural activity in the songbird forebrain. *Nat Neurosci* **2**, 209–211 (1999).

84. Kao, M. H., Wright, B. D. & Doupe, A. J. Neurons in a forebrain nucleus required for vocal plasticity rapidly switch between precise firing and variable bursting depending on social context. *J Neurosci* **28**, 13232–13247 (2008).

85. Kao, M. H. & Brainard, M. S. Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *J Neurophysiol* **96**, 1441–1455 (2006).

86. Brainard, M. S. & Doupe, A. J. Interruption of a basal ganglia–forebrain circuit prevents plasticity of learned vocalizations. *Nature* **404**, 762–766 (2000).

87. Vates, G. E., Broome, B. M., Mello, C. V. & Nottebohm, F. Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches. *J Comp Neurol* **366**, 613–642 (1996).

88. Coleman, M. J. & Mooney, R. Synaptic transformations underlying highly selective auditory representations of learned birdsong. *J Neurosci* **24**, 7251–7265 (2004).

89. Cardin, J. A., Raksin, J. N. & Schmidt, M. F. Sensorimotor Nucleus NIf Is Necessary for Auditory Processing But Not Vocal Motor Output in the Avian Song System. *Journal of Neurophysiology* **93**, 2157–2166 (2005).

90. Mello, C. V., Vates, G. E., Okuhata, S. & Nottebohm, F. *Descending Auditory Pathways in the Adult Male Zebra Finch (Taeniopygia Guttata)*. *J. Comp. Neurol* vol. 395 137–160 (1998).

91. Schneider, D. M. & Woolley, S. M. N. Sparse and Background-Invariant Coding of Vocalizations in Auditory Scenes. *Neuron* **79**, 141–152 (2013).

92. Stripling, R., Volman, S. F. & Clayton, D. F. Response modulation in the zebra finch neostriatum: relationship to nuclear gene regulation. *J Neurosci* **17**, 3883–3893 (1997).

93. Meliza, C. D. & Margoliash, D. Emergence of Selectivity and Tolerance in the Avian Auditory Cortex. *J. Neurosci.* **32**, 15158–15168 (2012).

94. Akutagawa, E. & Konishi, M. New brain pathways found in the vocal control system of a songbird. *The Journal of comparative neurology* **518**, 3086–3100 (2010).

95. Moll, F. W. *et al.* Thalamus drives vocal onsets in the zebra finch courtship song. *Nature* **616**, 132–136 (2023).

96. Lewandowski, B., Vyssotski, A., Hahnloser, R. H. R. & Schmidt, M. At the interface of the auditory and vocal motor systems: NIf and its role in vocal processing, production and learning. *J Physiol Paris* **107**, 178–192 (2013).

97. Cardin, J. A. & Schmidt, M. F. Auditory responses in multiple sensorimotor song system nuclei are co-modulated by behavioral state. *J Neurophysiol* **91**, 2148–2163 (2004).

98. Hahnloser, R. H. R. & Fee, M. S. Sleep-Related Spike Bursts in HVC Are Driven by the Nucleus Interface of the Nidopallium. *Journal of Neurophysiology* **97**, 423–435 (2007).

99. Bauer, E. E. *et al.* A Synaptic Basis for Auditory–Vocal Integration in the Songbird. *Journal of Neuroscience* **28**, 1509–1522 (2008).

100. Lewandowski, B. C. & Schmidt, M. Short Bouts of Vocalization Induce Long-Lasting Fast Gamma Oscillations in a Sensorimotor Nucleus. *J. Neurosci.* **31**, 13936–13948 (2011).

101. Mackevicius, E. L., Happ, M. T. L. & Fee, M. S. An avian cortical circuit for chunking tutor song syllables into simple vocal-motor units. *Nat Commun* **11**, 5029 (2020).

102. Norton, P., Benichov, J. I., Pexirra, M., Schreiber, S. & Vallentin, D. A feedforward inhibitory premotor circuit for auditory–vocal interactions in zebra finches. *Proceedings of the National Academy of Sciences of the United States of America* **119**, (2022).

103. Naie, K. & Hahnloser, R. H. R. Regulation of learned vocal behavior by an auditory motor cortical nucleus in juvenile zebra finches. *J Neurophysiol* **106**, 291–300 (2011).

104. Roberts, T. F. *et al.* Motor circuits are required to encode a sensory model for imitative learning HHS Public Access Author manuscript. *Nat Neurosci* **15**, 1454–1459 (2012).

105. Roy, A. & Mooney, R. Song Decrystallization in Adult Zebra Finches Does Not Require the Song Nucleus NIf. *J Neurophysiol* **102**, 979–991 (2009).

106. Roberts, T. F., Tschida, K. A., Klein, M. E. & Mooney, R. Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning. *Nature* **463**, 948–952 (2010).

107. Mackevicius, E. L., Gu, S., Denisenko, N. I. & Fee, M. S. Self-organization of songbird neural sequences 1 during social isolation 2. doi:10.1101/2022.02.18.480996.

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

109. Ölveczky, B. P., Andalman, A. S. & Fee, M. S. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS biology* **3**, 0902–0909 (2005).

110. Aronov, D., Veit, L., Goldberg, J. H. & Fee, M. S. Two Distinct Modes of Forebrain Circuit Dynamics Underlie Temporal Patterning in the Vocalizations of Young Songbirds. *J. Neurosci.* **31**, 16353–16368 (2011).

111. Veit, L., Aronov, D. & Fee, M. S. Learning to breathe and sing: development of respiratory-vocal coordination in young songbirds. *J Neurophysiol* **106**, 1747–1765 (2011).

112. Vallentin, D., Kosche, G., Lipkind, D. & Long, M. A. Neural circuits: Inhibition protects acquired song segments during vocal learning in zebra finches. *Science* **351**, 267–271 (2016).

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

114. Duffy, A., Latimer, K. W., Goldberg, J. H., Fairhall, A. L. & Gadagkar, V. Dopamine neurons evaluate natural fluctuations in performance quality. *Cell reports* **38**, (2022).

115. Leonardo, A. Experimental test of the birdsong error-correction model. *Proceedings of the National Academy of Sciences* **101**, 16935–16940 (2004).

116. Veit, L., Tian, L. Y., Monroy Hernandez, C. J. & Brainard, M. S. Songbirds can learn flexible contextual control over syllable sequencing. doi:10.7554/eLife.61610.

117. McGregor, J. N. *et al.* Shared mechanisms of auditory and non-auditory vocal learning in the songbird brain. *eLife* **11**, e75691 (2022).

118. Margoliash, D. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *Journal of Neuroscience* **3**, 1039–1057 (1983).

119. Margoliash, D. Preference for autogenous song by auditory neurons in a song system nucleus of the white-crowned sparrow. *J. Neurosci.* **6**, 1643–1661 (1986).

120. Margoliash, D. & Fortune, E. S. Temporal and harmonic combination-sensitive neurons in the zebra finch's HVc. *J Neurosci* **12**, 4309–4326 (1992).

121. Theunissen, F. E. & Doupe, A. J. Temporal and Spectral Sensitivity of Complex Auditory Neurons in the Nucleus HVc of Male Zebra Finches. *J. Neurosci.* **18**, 3786–3802 (1998).

122. Mooney, R. *Different Subthreshold Mechanisms Underlie Song Selectivity in Identified HVc Neurons of the Zebra Finch*. (2000).

123. Lewicki, M. S. Intracellular Characterization of Song-Specific Neurons in the Zebra Finch Auditory Forebrain. *J Neurosci* **16**, 5854–5863 (1996).

124. Mets, D. G. & Brainard, M. S. Genetic variation interacts with experience to determine interindividual differences in learned song. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 421–426 (2017).

125. Alam, D., Zia, F. & Roberts, T. F. The hidden fitness of the male zebra finch courtship song. *Nature* **628**, 117–121 (2024).

126. Perez, E. C. *et al.* The acoustic expression of stress in a songbird: does corticosterone drive isolation-induced modifications of zebra finch calls? *Horm Behav* **61**, 573–581 (2012).

127. Long, M. A., Jin, D. Z. & Fee, M. S. Support for a synaptic chain model of neuronal sequence generation. *Nature* **468**, 394–399 (2010).

128. Glaze, C. M. & Troyer, T. W. Temporal structure in zebra finch song: Implications for motor coding. *Journal of Neuroscience* **26**, 991–1005 (2006).

129. Sakata, J. T. & Brainard, M. S. Real-Time Contributions of Auditory Feedback to Avian Vocal Motor Control. *J. Neurosci.* **26**, 9619–9628 (2006).

130. Pehlevan, C., Ali, F. & Ölveczky, B. P. Flexibility in motor timing constrains the topology and dynamics of pattern generator circuits. *Nat Commun* **9**, 977 (2018).

131. Ross, M. T., Flores, D., Bertram, R., Johnson, F. & Hyson, R. L. Neuronal Intrinsic Physiology Changes During Development of a Learned Behavior. *eNeuro* **4**, ENEURO.0297-17.2017 (2017).

132. Ma, S., Maat, A. ter & Gahr, M. Neurotelemetry Reveals Putative Predictive Activity in HVC during Call-Based Vocal Communications in Zebra Finches. *J. Neurosci.* **40**, 6219–6227 (2020).

133. Rauske, P. L., Shea, S. D. & Margoliash, D. State and neuronal class-dependent reconfiguration in the avian song system. *J Neurophysiol* **89**, 1688–1701 (2003).

134. Nick, T. A. & Konishi, M. Neural song preference during vocal learning in the zebra finch depends on age and state. *Journal of Neurobiology* **62**, 231–242 (2005).

135. Nick, T. A. & Konishi, M. Dynamic control of auditory activity during sleep: Correlation between song response and EEG. *Proceedings of the National Academy of Sciences* **98**, 14012–14016 (2001).

136. Lewicki, M. S. & Arthur, B. J. Hierarchical Organization of Auditory Temporal Context Sensitivity. *J. Neurosci.* **16**, 6987–6998 (1996).

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

6 ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisor, Daniela Vallentin. I feel truly lucky to have you as my PhD supervisor and mentor. Apart from my PhD project, there are many things I have learned from you and I will always remember my time in the lab with fondness and gratitude. Thank you, for guiding me through this amazing journey, always believing in me from the very start and encouraging me when things did not quite go as expected. Thank you for introducing me to the songbird field and teaching me my favorite experimental method – intracellular recordings. Thank you, for providing me with opportunities to go to all the conferences in so many places and present my project. Thank you for the progress reports, where I felt comfortable to tell you everything that was going on my mind and you took the time to listen to me and provide me with advice. Thank you for reading my abstracts and manuscripts on your train and airplane journeys, when there was no time otherwise. Thank you for encouraging me to not be too shy to ask questions from the very start. Thank you for making my dream of ‘’living in the South of Germany’’ possible and for all the car journeys we shared. Most importantly, thank you for being a constant source of support and trust in my life.

My heartfelt thanks go to Jon – my favorite office mate in the smallest office of our corridor. Thank you, for always being there for me and being the “fairy cricket”. Thank you for all the scientific discussions we have had and for encouraging me to think more critically. Thank you, for being the SAP 2011 expert and always being open to help with yet another problem. Thank you for making my wish come true, to have ‘’you as my office mate until I am done with my PhD’’. Thank you for always being there to listen to me and for reminding me about my ‘’hundreds of vacation days’’. Above all, thank you for encouraging me to relax sometimes and not take things too seriously.

I would like to thank the Vallentin lab for the amazing years we have spent together. Thank you, Giacomo for all the lunch and coffee conversations we had and for all the comments and suggestions we had for each other’s work. Thank you for always being there and lending an open ear. Thank you, Carlos, for teaching me how to dance and being the next best friend to my pet birds. Thank you, Aditi, for all the entertaining lunch time stories. Thank you, Meera, for a sense of calmness and for letting me try your very tasty-looking lunch time food. Thank you, Tamir, for taking the time and providing me with advice when I needed it. Thank you, Franzi and Niels, for explaining to me how one can also be happy while listening to heavy metal. Thank you, Susi, for dealing with all the animal protocols and helping me find my first home in Bavaria. Thank you, Neetash, Anna and Margarida for all the fun we had together at our Sunday brunches. Thank you, Fabi for all the car journeys and life stories we shared. Thank you, Stefan, for being supportive and reminding me, that there are also other things in life.

I am grateful to my parents, who showed me early on that everything in life is possible, if only you try. Thank you, for providing me with the opportunity to ever even get this far and for being my ‘’miera osta’’, that I always knew I could return to. I would like to thank my siblings for all the silliness we do, when I go home. Finally, I would like to thank my partner, for always being there for me, inspiring me to reach for my dreams and reminding me of all the nice things around me.

7 LIST OF PUBLICATIONS

Bistere, Gomez-Guzman, Xiong, Vallentin (accepted in *Nature Communications*) “Female vocal feedback promotes song learning in male juvenile zebra finches”

Bistere, Wilczek, Vallentin (under review in BMC Neuroscience) “Temporal dynamics of song crystallization in zebra finches”

8 AFFIDAVIT

I hereby confirm that the dissertation “Neural circuits for socially guided vocal learning” is the result of my own work and that I have only used sources or materials listed and specified in the dissertation.

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation “ Neural circuits for socially guided vocal learning” selbstständig angefertigt habe, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Munich, May 13th, 2024

Linda Bistere

9 DECLARATION OF AUTHOR CONTRIBUTIONS

First Study: ‘’Female vocal feedback promotes song learning in male juvenile zebra finches’’

Authors: Linda Bistere, Carlos M. Gomez-Guzman, Yirong Xiong, Daniela Vallentin

Author Contributions: L.B. and D.V. conceived the study and designed the experiments, L.B. and C.M.G.G. conducted the experiments, L.B., Y.X., C.M.G.G. and D.V. analyzed the data, L.B. and D.V. wrote the manuscript, D.V. acquired funding and supervised the project.

My contribution to this study: Together with my supervisor, Dr. Daniela Vallentin, we planned the study and designed the experiments. I trained the juveniles birds with a song playback and tracked female vocalizations during song learning. Additionally, I performed intracellular recordings of HVC projection neurons in both head-fixed and singing juvenile birds. I analyzed the behavioral data and the data from intracellular recordings. I wrote the first version of the manuscript and revised it together with my supervisor.

Second study: “Temporal dynamics of song crystallization in zebra finches”

Authors: Linda Bistere, Stefan Wilczek, Daniela Vallentin

Author Contributions: L.B. and D.V. designed the study and experiments, L.B., S.W. and D.V. conducted the experiments, L.B. analyzed the data, L.B. and D.V. wrote the manuscript, D.V. acquired funding and supervised the project.

My contributions to this study: My supervisor Dr. Daniela Vallentin and I designed the study and the experiments. I preformed the intracellular recordings from three out of four freely moving, singing juvenile birds. I analyzed the data from intracellular recordings and described the activity of HVC projection neurons during singing in juveniles and adults. I wrote the first version of the manuscript and revised it together with my supervisor.