Mechanisms of Vocal Coordination in Zebra Finches

Jonathan I. Benichov

Graduate Center, City University of New York

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu(gc_etds

Part of the Biology Commons

Recommended Citation

https://academicworks.cuny.edu(gc_etds/858

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
Mechanisms of Vocal Coordination in Zebra Finches

By

Jonathan I. Benichov

A dissertation submitted to the Graduate Faculty in Biology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2015
This manuscript has been read and accepted for the Graduate Faculty in Biology in satisfaction of the dissertation requirements for the degree of Doctor of Philosophy.

Date

Chair of Examining Committee
Dr. Ofer Tchernichovski

Date

Executive Officer
Dr. Laurel A. Eckhardt

Dr. Mark Hauber, Hunter College

Dr. Christopher Braun, Hunter College

Dr. Adrian Rodriguez-Contreras, CCNY

Dr. Michael A. Long, NYU School of Medicine

Supervising Committee

The City University of New York
Mechanisms of Vocal Coordination in Zebra Finches

Jonathan I. Benichov

Advisor: Dr. Ofer Tchernichovski

Abstract

Social animals frequently emit communication calls. Although these calls are often innate in their acoustic structure, they can be used adaptably to coordinate behavior with other individuals. It is not known, however, what each animal needs to learn in order to achieve and maintain synchronized call patterns with others. To study this process, we have developed a vocal robot that can be programmed to generate call patterns or to sense a bird’s contact (short) calls and respond with precisely timed call answers. By varying the robot’s vocal behavior, including call timing and rhythm, we tested how interacting zebra finches adapt to different call patterns produced by a partner robot bird. This approach allows us to assess engagement and the capacity to synchronize calls between females (vocal non-learners) and males (vocal learners) as well as birds with different levels of developmental social experience. We also tested if forebrain structures that are known to be involved in song learning are required for the coordination of calls. We discovered that zebra finches can learn to adjust the timing of their responses to a robot bird partner within minutes. Further, when challenged with complex rhythms containing jamming elements, birds dynamically adjusted the timing of their calls in anticipation of jamming. Blocking the song system cortical output dramatically reduced the precision of birds’ response timing and abolished their ability to avoid jamming. Surprisingly, we observed this effect in both males and females, indicating that the female song system is functional rather than vestigial. We then tested if social interactions during development are necessary for birds to acquire the capacity to synchronize and adapt their call timing to those of a partner bird robot. We found that socially isolated birds were extremely imprecise in the timing of their responses. Further, they were unable to avoid disruptive jamming. Interestingly, these results were very similar to those observed after blocking the forebrain song system in socialized
birds. We conclude that social interactions during development are necessary for zebra finch males to develop the capacity to precisely adapt the timing of their calls. Further, the capacity to synchronize calls must be acquired independently from that of song learning. Finally, we investigated if, and to what extent, birds can take into account the behavior of a third party while interacting with a partner. Using miniaturized wireless audio transmitters, we found that when two birds are interacting simultaneously with the vocal robot and with each other, they can avoid jamming with each other and with robot by cooperatively changing the latencies of their answer calls. These qualitative results suggest that birds are capable of adjusting the timing of their calls with respect to more than a single partner bird. Together, our results uncover behavioral and physiological mechanisms that give rise to vocal coordination, bridging a functional gap between innate and learned vocalization abilities.
Acknowledgements

This work would not have been possible without the contributions, support, and guidance of others. First and foremost, I would like to thank Dr. Ofer Tchernichovski for cultivating a lab environment that provided a balance of freedom, rigor, and lunchtime. Your mentoring style managed to put many of my concerns at ease while still helping me to stay motivated and excited about my work. Thank you to all the members (past and present; human and avian) of Ofer’s Lab, particularly my fellow occupants of “The Nook”, whose collection of “Oferisms” will one day provide the basis of a great religion. Although we kvetched about having to move the lab from CCNY to Hunter, I am glad I was there to set up and benefit from the lab’s new home. I also ended up with the best commute in the city; Biking through Central Park energized me on the way to the lab and allowed me to blow off steam after a long day of wrestling with SAP.

I am fortunate to have also been able to work with outstanding collaborators. I am grateful to Dr. Michael Long and his team for welcoming me into their lab and providing support, guidance, and laughs. In particular, thanks to Sam Benezra for his skilled contributions to the anatomical data and Dr. Daniela Valentin for her help with analysis and electrophysiology. Dr. Eitan Globerson provided insights that could have only come from the unique perspective of a master musician with a Ph.D. in neuroscience. Many thanks to Dr. Andries Ter Maat and colleagues at the Max Planck Institute of Ornithology for their hospitality and generosity in sharing their remarkable transmitter technology.

Thank you to the faculty and staff of the Biology program at the Graduate Center and the Hunter Psychology Department. Many thanks to my Doctoral Committee - Drs. Chris Braun, Mark Hauber, Michael Long, and Adrian Rodriguez-Contreras - for your time and valuable feedback. Thank you to my students for helping me continue to view the world with fresh eyes and a sense of wonder while also enabling me to develop a more confident voice. I gratefully acknowledge the funding sources that supported this work: NIDCD, NSF, PSC CUNY, and the CUNY Science Scholarship. Last, but certainly not least, thank you to my friends and family, especially my parents, for your seemingly endless supply of encouragement, patience, and love.
# Table of Contents

Abstract ......................................................................................................................... iii

Acknowledgements .......................................................................................................... v

Table of Contents ............................................................................................................. vi

List of Figures .................................................................................................................. viii

Chapter 1: Introduction ..................................................................................................... 1
  1.1 Background .................................................................................................................. 1
  1.2 Ecology and Group Dynamics ..................................................................................... 1
  1.3 Individual recognition and the neural basis of call perception .................................. 3
  1.4 Plasticity in the timing of calls and the neural basis of call production ................. 4

Chapter 2: The Forebrain Song System Mediates Predictive Call Timing in Female and Male Zebra Finches ............................................................................................................ 9
  2.1 SUMMARY .................................................................................................................. 9
  2.2 BACKGROUND AND RATIONALE ........................................................................... 10
  2.3 METHODS ................................................................................................................ 12
    2.3.1 Animal care ......................................................................................................... 12
    2.3.2 Establishing vocal synchronization in a pair of birds ........................................... 12
    2.3.3 Adaptive Vocal Interface (AVI) vocal robot system .............................................. 13
    2.3.4 Call response analysis ......................................................................................... 14
    2.3.7 Electrolytic Lesions and RA Inactivation ............................................................. 16
  2.4 RESULTS ................................................................................................................... 18
    2.4.1 Call Interactions .................................................................................................. 18
    2.4.2 Jamming Avoidance ......................................................................................... 20
    2.4.3 Dynamic Adjustments to Complex Rhythms .................................................... 24
    2.4.4 Effects of Blocking the Output of the Forebrain Song System ...................... 28
  2.5 DISCUSSION ............................................................................................................. 35
    2.5.1 The Forebrain Song System and Anticipatory Call Timing ............................. 35
    2.5.2 Phylogenetic Perspective on Vocal Learning .................................................. 37
  2.6 AUTHOR CONTRIBUTIONS .................................................................................... 39
Chapter 3: How Developmental Social Experience Affects the Capacity for Vocal Coordination

3.1 SUMMARY ........................................................................................................................................40
3.2 BACKGROUND AND RATIONALE ...............................................................................................40
3.3 METHODS.........................................................................................................................................40
3.3.1 Social Isolation During the Sensitive Period for Vocal Learning .............................................42
3.3.2 Assessing Call Coordination Capacity in Social Isolates .........................................................42
3.4 RESULTS..........................................................................................................................................43
3.4.1 Effects of social isolation on call response timing.................................................................43
3.4.2 The effect of social isolation on call timing precision is similar to that of blocking the forebrain song system .........................................................................................46
3.5 DISCUSSION ...................................................................................................................................50

Chapter 4: Vocal coordination beyond the dyad .................................................................................52

4.1 SUMMARY.......................................................................................................................................52
4.2 BACKGROUND AND RATIONALE ...............................................................................................52
4.3 METHODS.........................................................................................................................................52
4.3.1 Wireless audio transmitters for call source separation ..........................................................55
4.3.2 Intrusive Robot Calls ..............................................................................................................57
4.4 RESULTS..........................................................................................................................................59
4.4.1 Birds avoid jamming with each other as they respond to robot calls. .................................59
4.4.2 Jamming robot calls are incorporated into a sequence of group calling ............................59
4.5 DISCUSSION ...................................................................................................................................61

Appendix ................................................................................................................................................63

A1 ADDITIONAL METHODS DEVELOPED ....................................................................................63
A1.1 Responsive Robotic Vocal Interface ..........................................................................................63
A1.2 Robot responds to one bird in a pair .......................................................................................63
A1.3 Call modulation in response to predictable environmental noise ........................................64
A1.4 3D accelerometer & Heart rate ..............................................................................................66

References.............................................................................................................................................68
List of Figures

Fig. 2.1: Schematics of the vocal robot system and an isochronous call (IC) session. ..........14
Fig. 2.2: Determining the jamming window and a schematic of a jamming call session. ..........15
Fig. 2.3: Rhythm patterns of alternating single calls and jamming call pairs. .........................16
Fig. 2.4: Forebrain lesions and RA inactivation. ......................................................................18
Fig. 2.5: Call exchanges in a pair of zebra finches. .................................................................19
Fig. 2.6: Responses to the vocal robot. .....................................................................................20
Fig. 2.7: Jamming avoidance. .................................................................................................21
Fig. 2.8: Jamming avoidance strategies and response probability across birds.......................22
Fig. 2.9: Assessing jamming avoidance with catch trials. .......................................................23
Fig. 2.10: Jamming avoidance across a session. ......................................................................24
Fig. 2.11: Rhythm adaptation. ...............................................................................................26
Fig. 2.12: Responses to rhythmic jamming calls. ....................................................................27
Fig. 2.13: Response rates across experimental conditions.......................................................28
Fig. 2.14: Effects of RA lesions on responses to ICs. ..............................................................29
Fig. 2.15: Effects of RA and control lesions on response precision and skewness...................30
Fig. 2.14: Effects of RA inactivation on responses to ICs........................................................31
Fig. 2.17: Effects of RA lesions on precision and jamming avoidance. .....................................32
Fig. 2.18: Jamming avoidance in 4 control-lesioned birds. ......................................................33
Fig. 2.19: Jamming avoidance is not correlated with response rates. .......................................34
Fig. 3.1: Song learning and call responses in a developmentally isolated adult male. ...............44
Fig. 3.2: Social isolates’ call response timing compared to socially reared birds....................45
Fig. 3.3: Features of isolate call timing resemble birds with RA lesions. .................................47
Fig. 3.4: Social isolates exhibit impaired jamming avoidance..................................................48
Fig. 3.5: Response distributions across social conditions.......................................................49
Fig. 4.1: Schematic of coupled timing representations for the detection of a signal in noise.....55
Fig. 4.2: Source separation with miniature audio transmitters. ..........................................................57
Fig. 4.3: Recording from a pair of birds interacting with the vocal robot........................................58
Fig. 4.4: Bird responses to the vocal robot alone and in a pair............................................................60
Fig. A1: Vocal robot design and operation in responsive mode............................................................63
Fig. A2: Vocal robot selectively response to one bird within a pair....................................................64
Fig. A3: Spectrogram of a periodic white noise pattern with a phase shift........................................65
Fig. A4: Recording EKG with a Neurologger.......................................................................................66
Fig. A5: Recording EKG across an interactive call bout with a vocal-robot........................................67
Chapter 1: Introduction

1.1 Background

The zebra finch vocal repertoire primarily consists of short vocal sounds known as calls. The spectral structure of most calls is simple and considered to be unlearned (or innate). Both males and females call often in a variety of social contexts that extend beyond territoriality and courtship (Marler, 2004; Zann, 1996). Short contact calls, such as “tets” and “stacks”, in particular are frequently exchanged between birds that are in close proximity. Individual birds, however, do not call with random timing and are not limited to reflexively responding to calls in an indiscriminate manner. Even within a seemingly chaotic stream of chatter, call interactions can exhibit emergent structure that is specific to particular pairs as well as groups (Anisimov et al., 2014; Elie et al., 2010; Elie, Soula, Mathevon, & Vignal, 2011; Ter Maat, Trost, Sagunsky, Seltmann, & Gahr, 2014). It is unclear, though, how call and response interactions develop from moment to moment and over developmental timescales: can birds learn to predict the timing of other bird’s calls? Can they perceive and adapt their answers to complex patterns of calls produced by other birds? What is the role, if any, of development and lifetime experience of interacting with birds in the capacity to coordinate calls? And finally, is there any mechanistic link between the ability to coordinate call timing and song learning? To guide our investigation of call coordination, we will consider the contexts in which calls are produced, the plasticity of their timing and structure, as well as the neural pathways that govern their perception and production.

1.2 Ecology and Group Dynamics

In the wild, zebra finches live in fluid colonial groups that vary widely in composition, density, and size between breeding cycles. Most adults form lifelong pair bonds that are
maintained within this broader fission-fusion flocking system (Zann, 1996). These social phenotypes and associated vocal communication are tightly coupled to specific structures and activity within the mesolimbic dopaminergic system (Goodson, Kabelik, Kelly, Rinaldi, & Klatt, 2009). In general, avian flocking behavior is driven by the individual’s ability to continuously monitor the state of the group via local interactions and respond in a context-dependent manner. Such collective behaviors are often robust enough to withstand large variations in group size and density while maintaining or even enhancing responsiveness to social and environmental stimuli (Ballerini et al., 2008; Cavagna et al., 2010; Couzin, 2007; N. Miller, Garnier, Hartnett, & Couzin, 2013). Events of formation and dissolution of zebra finch flocks, across spatial and temporal scales, are likely to be regulated by an adaptive calling mechanism for establishing and maintaining group and pair cohesion while coupling behavioral states (Goodson et al., 2009; Zann, 1996).

Groups of zebra finches can both synchronize vocal behavior and respond to changes in the group vocal signal. Elie et al (2011) have demonstrated that zebra finches can generate distinguishable patterns of vocal activity at the group level. She formed heterogeneous groups of birds that differed in size and composition (e.g. sex ratios and age) and recorded group vocal activity daily for several weeks. After two weeks, the vocal output of each group was remarkably distinct in vocalization rates and rhythms. The group-specific patterns ranged from those with relatively uncoordinated chatter to others characterized by dramatic collective bursts of vocalization interleaved with periods of quiet. Although this vocal plasticity is interesting to consider in light of the bird’s flexible grouping in the wild, it is unknown how this form of group-level vocal structure develops or whether it has a particular function. However, in another study on group vocal behavior, Waas, et al (2005) reported that playing colony audio recordings to
breeding groups resulted in the advancing and synchronizing of breeding schedules within the group. In addition, clutch sizes increased compared to controls that were not exposed an intensified group signal. Together this preliminary evidence suggests the ability of individuals to vocally coordinate within a group and to use higher-order group vocal signals, possibly conveying information about flock size and state, to influence behavior in a socially - and perhaps evolutionarily - adaptive manner.

1.3 Individual recognition and the neural basis of call perception

Benefits gained from modifying behavior based on group activity often must be balanced with benefits derived from acting based on individual experience (Miller, 2013). For the pair bonding zebra finch, this tension is reflected in the ability to selectively communicate with certain individuals while remaining unresponsive to others nearby. Ter Maat et al (2014) and Anisimov et al (2014) were able to precisely track the calls of individual birds within semi-naturalistic colonies and both found striking specificity in the symmetry of call-response patterns across pairs. This suggests that although call signals are shared with and perhaps used by the whole group, birds also possess the capacity for individual recognition and selective (i.e. learned) coordination of their call responses.

The auditory mechanisms underlying both song and call perception are highly sensitive to social context. Neural responses and immediate early gene expression are modulated by the sex of the signaler, phase of breeding cycle, social experience, and asexual vs. sexual nature of a relationship (Cousillas et al., 2013; Gobes et al., 2009; Maul et al., 2010; Vignal, Andru, & Mathevon, 2005; S. C. Woolley & Doupe, 2008). In particular the caudomedial nidopallium (NCM), a secondary forebrain auditory region, shows specificity of neuronal activity that is
believed to provide a modifiable representations of the familiar calls and songs of multiple individuals (Chew & Vicario, 1996; Menardy et al., 2012). Furthermore, auditory responses within the song system to a variety of short calls, alarm calls, and distance calls may encode information about “semantic” contexts in which they are used (Elie & Theunissen, 2015a). The same song system auditory regions, it is argued, also partially encode the neural representation of auditory memories or “template” necessary for song learning (Bolhuis & Gahr, 2006).

The specificity of call responses can arise early in development. For instance, chicks exhibit preferential responses to parent calls, although males eventually lose this preference as they mature (Mulard, Vignal, Pelletier, Blanc, & Mathevon, 2010). As adults, pair bonded individuals are also known to produce specific vocal duets at the nest (Elie et al., 2010). In addition, the ability to discriminate between the longer distance calls of males and females also requires an intact vocal motor pathway (David S Vicario, Naqvi, & Raksin, 2000). Interestingly, changes to individual interactions may propagate, with ramifications for the broader group structure. Macguire et al (2013) found that inactivating a female cowbird’s song system nucleus HVC (formal name) reduces the specificity of her responses to individual vocalizations within a group, which unexpectedly triggers the restructuring the entire social network. Taken together, these findings demonstrate that the forebrain structures involved in song learning are also likely to play a critical role in the processing of unlearned calls to modulate responses in a socially relevant way.

1.4 Plasticity in the timing of calls and the neural basis of call production

Individual-specific call interactions that occur within a larger set of group vocalizations face the problem of interfering signals or masking. Although it is known that vocal amplitude can
be modulated by both males and females based on the proximity of the receiver and levels of environmental noise (Brumm & Slater, 2006; J Cynx, Lewis, Tavel, & Tse, 1998; Jeffrey Cynx & Gell, 2004), if all birds in a group relied solely on this mechanism, an amplification arms race would at best provide a temporary advantage in signal to noise ratio. In this case, a basic jamming avoidance response can provide a means of coordinating signals within a noisy environment. Jamming avoidance is exhibited at different timescales by phylogenetically diverse species such as electric fish, anurans, and bats (Carr, 1993; Gillam, Ulanovsky, & McCracken, 2007; Heiligenberg, Wong, Metzner, & Keller, 1996; Jones, Jones, & Ratnam, 2014; Stamper, Madhav, Cowan, & Fortune, 2012; Ulanovsky, Fenton, Tsoar, & Korine, 2004) and is likely to have evolved within various highly competitive signaling environments. This remarkably precise control of signaling can be achieved largely in the brainstem (Heiligenberg et al., 1996; C. P. Schwartz & Smotherman, 2011). A similar response mechanism in zebra finches would allow call interactions to occur simultaneously over different channels (i.e. multiplexing). Antiphonal or phase-shifted exchanges of calls may be an example of this.

Coupled oscillator models have been proposed to characterize how the rhythmic calls of individuals are adjusted in their rate and phase to produce antiphonal calling (Aihara et al., 2011; Ghazanfar, Smith-Rohrberg, & Pollen, 2002; Jones et al., 2014; Takahashi, Narayanan, & Ghazanfar, 2013). In plain-tailed wrens, antiphonal duets are associated with temporally specific neural representation of both partners’ combined vocalization in the activity of HVC. Similar activity is present even when one bird is singing their part in isolation. Furthermore, the activity of HVC neurons is highly sensitive to inter-syllable intervals of heard songs, such that they respond less to manipulated playbacks of their own duet with intervals differ from the natural timing, even when the syllables are identical in acoustic structure (Fortune, Rodríguez, Li, Ball,
& Coleman, 2011). This is suggestive of a central pattern generator that can be aligned to a partner’s through auditory feedback. Developmentally, starlings begin to exhibit similar antiphonal calling behavior even before fledging: first between brood mates and then with parents (Chaiken, 1990).

Social context is known to affect the rhythm and timing of vocalizations, specifically in the case of conspecific-directed song (Hessler & Doupe, 1999; Zann, 1996). The presence of conspecifics can induce physiological changes in heart rate, respiration, dopaminergic signaling, and neural activity in the forebrain song system (Cooper & Goller, 2006; Hessler & Doupe, 1999; Kao, Wright, & Doupe, 2008; Sasaki, Sotnikova, Gainetdinov, & Jarvis, 2006). In addition, song timing can be specifically manipulated, independently of other song features, by locally cooling or heating HVC. The same procedure, applied to the downstream premotor nucleus robustus archipalialis (RA) has no affects on song timing (Long & Fee, 2008). Evidence from lesioning studies, however, suggests that RA, in combination with ascending brainstem activity, can influence song timing independently of motor commands from upstream nuclei such as HVC (Ashmore, Bourjaily, & Schmidt, 2008; Ashmore, Renk, & Schmidt, 2008). These results provide strong evidence that activity in the vocal pre-motor pathway governs the precise timing of vocal sequences produced by the song system.

On the temporal scale of individual calls, it has been shown that some male long “distance” calls can include learned frequency modulation patterns. These learned features are dependent upon the activity of both HVC and RA (Simpson & Vicario, 1990). When either of these nuclei are lesioned, calls can still be produced, however male calls become less modulated and more similar to those of females. These findings were expanded upon by
electrically stimulating the dorsomedial nucleus (DM) of the brainstem, which evoked simple calls, where as stimulation of either HVC or RA evoked more complex vocal structures (Vicario & Simpson, 1995). It is worth noting that time-locked premotor activity in RA also precedes the production of unlearned calls (Ter Maat et al., 2014). It remains to be seen how the activity of RA or HVC controls specific temporal aspects of calls that are acoustically unlearned yet socially modified with respect to the timing of their usage.

Here we go beyond existing - predominantly observational - approaches to studying call coordination by virtually participating in call interactions. Namely, we exert experimental control over the parameters of the “conversation” using a vocal robot. First, we address the question of whether birds can anticipate the calls of others by challenging them to interact with a not-so-cooperative calling partner with a range of temporal complexity and by measuring how birds adapt their response timing. We also take advantage of the behavioral and anatomical similarities and differences between male and female zebra finches to ascertain whether call coordination has mechanistic links with song learning. Previously, activity within the song system has been associated call interactions, based on passive recording (Ter Maat et al., 2014). However, in order to determine whether song system activity is necessary for specific aspects of call coordination, we blocked the output of the song system and measured the effects within controlled call interactions. This approach also extends to our examination of the effects of developmental social experience on the capacity to coordinate calls in adulthood.

Finally, the ability to monitor specific individuals within a calling group has unveiled a wealth of information and experimental possibilities (Anisimov et al., 2014; Ter Maat et al., 2014; Vyssotski, Anisimov, Latanov, & Hahnloser, 2013). But because group call interactions
inherently contain multiple sources of variability, with effects that can echo within complex feedback loops, it is difficult to gain a foothold on specific parameters that drive moment-to-moment changes in each bird’s behavior. We demonstrate that it is enormously advantageous to introduce an experimentally controlled point of reference within these interactions (i.e. the vocal robot) in order to identify causal factors. We can then determine that birds adapt their behavior in response to but also in anticipation of the behavior of others. We believe that this collection of approaches is both novel and fruitful for understanding the calling behavior of zebra finches, adding to a growing set of tools for probing the dynamics of interactive behavior across biological systems (Kelso, Guzman, & Reveley, 2009; Lerch, Roy, Pachet, & Nagle, 2011; C. T. Miller, Beck, Meade, & Wang, 2009; Patricelli, Coleman, & Borgia, 2006; Prinz, Abbott, & Marder, 2004; J. J. Schwartz, Buchanan, & Gerhardt, 2002; Smith, 1996; Swain, Couzin, & Ehrich Leonard, 2012).
2.1 SUMMARY

The dichotomy between vocal learners and non-learners is a fundamental distinction in the study of animal communication. Male zebra finches (*Taeniopygia guttata*) are vocal learners that acquire a song resembling their tutors’, whereas females can only produce innate calls. The acoustic structure of short calls, produced by both males and females, is not learned. However, these calls can be precisely coordinated across individuals. To examine how birds learn to synchronize their calls, we developed a vocal robot that exchanges calls with a partner bird. Because birds answer the robot with stereotyped latencies, we could program it to disrupt each bird’s responses by producing calls that are likely to coincide with the bird’s. Within minutes, the birds learned to avoid this disruptive masking (jamming) by adjusting the timing of their responses. Notably, females exhibited greater adaptive timing plasticity than males. Further, when challenged with complex rhythms containing jamming elements, birds dynamically adjusted the timing of their calls in anticipation of jamming. Blocking the song system cortical output, using either electrolytic lesions or transitory inactivation, dramatically reduced the precision of birds’ response timing and abolished their ability to avoid jamming. Surprisingly, we observed this effect in both males and females, indicating that the female song system is functional rather than vestigial. The forebrain song-production pathway, then, appears to be a general-purpose sensorimotor communication system that facilitates social interactions. In the case of calls, it enables plasticity in vocal timing, whereas in the case of songs, plasticity extends to developmental changes in vocal structure.
2.2 BACKGROUND AND RATIONALE

The acoustic structures of most animal vocalizations are not learned. The ability to match vocal sounds to sensory templates is rare, although recent studies show evidence for vocal learning in a growing number of species (Knörnschild, 2014; Poole, Tyack, Stoeger-Horwath, & Watwood, 2005). Vocal coordination is much more common, having been recognized in many species that are vocal non-learners such as cicadas (Hedwig, 2014), frogs (Jones et al., 2014), and marmosets (Takahashi et al., 2013). Here we investigate the hypothesis that vocal coordination is mechanistically, and perhaps evolutionarily (Liu, Wada, Jarvis, & Nottebohm, 2013; Pfenning et al., 2014), linked to vocal learning.

We examine call coordination in zebra finches, taking advantage of their sexual dimorphism. Males are vocal learners that acquire their songs by imitation, but females are vocal non-learners that cannot sing and produce only innate calls (Simpson & Vicario, 1990; Zann, 1996). Although the forebrain song system is not required for producing short calls (D S Vicario & Simpson, 1995) or responding to them (Simpson & Vicario, 1990), birds appear to show plasticity in the timing of their calls. As in many social species that are vocal non-learners, the temporal patterns of calls produced by a group of zebra finches can be highly structured (Elie et al., 2010, 2011) which may reflect social bonds (Anisimov et al., 2014; Elie & Theunissen, 2015b; Ter Maat et al., 2014) and social hierarchy (J. Benichov & Tchernichovski, 2012; Lerch et al., 2011).

To test how male and female zebra finches learn to synchronize their calls, we presented them with a vocal robot that we initially programmed to produce calls in a simple isochronous pattern. This allowed the bird to exchange back-and-forth (antiphonal) calls with the
robot. We then programmed the robot to produce call patterns designed to disruptively interfere with this communication channel (Movie S1), prompting the birds to adjust the timing of their responses. Challenging birds with more complex call patterns allowed us to further test for a capacity to dynamically adjust call timing and predictively avoid jamming by calling in patterns that conform to the presented rhythm. Finally, we tested if the forebrain song system, which is thought to be dedicated to vocal learning, is required for the control and for the predictive adjustment of call timing. We compared the effects of blocking or inactivating the song system on vocal timing plasticity between males and females. Results prompt a reevaluation of the context in which forebrain mechanisms enable vocal plasticity.
2.3 METHODS

2.3.1 Animal care

All animal care and experimental procedures were performed according to the guidelines of the US National Institutes of Health and were reviewed and approved by the Institutional Animal Care and Use Committees of Hunter College of the City University of New York and the New York University Langone Medical Center. For behavioral experiments we used 10 adult (>90 days post-hatch) zebra finches (*Taeniopygia guttata*) bred at Hunter College of the City University of New York. For combined behavioral and lesion studies, which were performed at NYU, 10 adult birds were purchased from a breeder. All birds were maintained in temperature- and humidity-controlled environments with a 12/12 h light/dark schedule. During vocal-robot experiments birds were kept singly in sound attenuation chambers. Birds were provided with unlimited access to seeds and water throughout the experiments.

2.3.2 Establishing vocal synchronization in a pair of birds

Two adult birds (1 male, 1 female) were kept separately in sound attenuation chambers for 5 days prior to pairing. We then moved the female to the male’s chamber and continuously recorded audio using Sound Analysis Pro (SAP) for four additional days. Birds’ call onsets and durations were identified semi-automatically and analyzed using Matlab 7. Call feature calculation and cluster analysis were performed using SAP. Cluster information was used to elucidate bird identity. We then used Matlab 7 for analysis of call answers (calls within a 500ms window) for 500 consecutive calls, selected pseudorandomly from each of the four days.
2.3.3 Adaptive Vocal Interface (AVI) vocal robot system

To examine how birds learn to synchronize their calls, we employed a ‘vocal robot’ system (J. Benichov & Tchernichovski, 2012) that can exchange calls with a partner bird (Lerch et al., 2011). We developed the Adaptive Vocal Interface (AVI) using Labview 2012 (code to be made available). The system can be programmed to deliver calls in modes that are either “Leading” and predictable or “Responsive” and adaptive to specific calls produced by the bird (Figure 2.1 and Movie S1: Video demonstration.). Robot calls can be delivered as repeated patterns or rhythms, with little or no contingency on the bird’s calls while measuring the bird’s responses (leading mode), as used here.

The visual component can be easily modified using a 3D printer and an Arduino Uno controller that can translate amplitude of robot call audio into synchronized movement of the plastic model. Vocal robot call audio files were composed of natural calls recorded at 44.1kHz sampling rate from an interacting pair-bonded male in a sound-attenuated chamber. These calls were representative of an average “stack” call and were tested for the ability to elicit call responses in male and female birds. A10kHz pure tone marker (inaudible to birds at this level (Amin, Doupe, & Theunissen, 2007)), of the same duration and RMS amplitude, was added to the call for identifying onsets/offsets in case of jamming. Calls were delivered at 70dB through a mono-channel speaker. Call patterns generated were isochronous (rate of 1Hz) or consisted of jamming call pairs (one jamming pair per second) (Figure 2.2).
Fig. 2.1: Schematics of the vocal robot system and an isochronous call (IC) session. Each block consists of isochronous calls (ICs) delivered by the vocal robot at the rate of 1Hz for 30 seconds. The bird receives one block of calls per minute for a 10 minute session.

2.3.4 Call response analysis

Each bird’s calls were recorded over a ten-minute session of isochronous robot calls. Responses were automatically segmented with SAP and manually segmented in case of overlap. Call response onsets and offsets were coded relative to robot call onsets (or first call in a jamming pair) for each cycle. Coded responses were used to generate raster plots and probability distributions in Matlab 7. To calculate the jamming window, responses were summed across all cycles in a session to produce a response probability distribution. The jamming window was defined as the 100ms interval with the highest response density. The window onset is the latency of the jamming call delivered in each 1s cycle during the next session of jamming calls. Jamming percentages were calculated as the proportion of total calling activity falling with the bounds of the jamming window. Catch trials were calculated as above. Response latencies and skewness were calculated in Matlab 7 using the onsets of responses, relative to the previous robot call. Precision scores were calculated for each session using the proportion of all response onset latency differences that were within ±50ms (approx. duration of a call). This proportion was used to compute a Z-score, relative to a distribution of proportions from 1000
simulated sessions containing an equal number of uniformly distributed pseudorandom latencies. The precision score is expressed as the square root of this Z-score.

![Diagram](image)

**Fig. 2.2: Determining the jamming window and a schematic of a jamming call session.**

(A) The distribution of a bird’s call responses to the robot’s ICs during a 10-minute session are used to compute an optimal jamming window. The robot then produces jamming calls during this window in the following session. (B) A session of robot jamming call pairs, delivered at a rate of 1Hz in blocks of 30 seconds, separated by 30 seconds of silence. Catch trials (cycles containing a single call) occur with 10% probability. The bird receives one block of calls per minute for a 10 minute session.
2.3.6 Robot rhythm patterns

(a) Rapidly alternating single calls and jamming pairs (1s cycle, 200ms jamming latency). Robot generates 10-second blocks followed by 10 seconds of silence, every 20 seconds for 10 minutes.

(b) Slowly alternating rhythm pattern (2s cycle, 250ms jamming latency). Robot generates a 30-second block of calls every 60 seconds for a 10 minutes.

For responses to single calls and jamming calls in a rhythm, we calculated median response latencies for those responses with onsets prior to the expected jamming window.

2.3.7 Electrolytic Lesions and RA Inactivation

In order to determine if the song system is necessary for controlling call timing, we bilaterally lesioned the final common output nucleus of the forebrain vocal motor pathway, the Nucleus Robustus of the Archopallium (RA). Blocking this structure leaves only midbrain, brainstem, and other non-specific influences on vocal production.

Birds were anesthetized with 1.5-3% isoflurane in oxygen and head-fixed with ear bars on a heated stereotaxic stage. The skull was rotated forward 80 degrees from horizontal, as measured at the anterior surface. Bilateral craniotomies were performed 2.3mm lateral of the
midline and 1.85 mm posterior of the bifurcation of the mid-sagittal sinus, creating an ~600 x 600 μm window above each hemisphere. Dura was removed and a 5 μm carbon fiber microelectrode (Kation Scientific, Minneapolis, MN) was lowered at an angle of 15 degrees to locate RA via extracellular spike monitoring. Bilateral electrolytic lesions were produced with a bipolar stimulating electrode applying 100 μA of current (60s per location) at 3 sites spanning ~600 μm across RA at a depth corresponding to the center of the nucleus (Figure 2.4).

Control lesioned birds underwent identical procedures with the stimulating electrode inserted 150μm into the pallial surface directly dorsal to RA. Craniotomies were sealed with Kwik-Cast (World Precision Instruments). Anti-bacterial ointment (Neosporin) was applied to incised scalp and birds were allowed to recover for at least 24 hours in an isolated heated chamber where they were provided with antibiotic (Baytril) water solution and food ad libitum. After recovery the lesioned bird was transferred to the same pre-lesion testing chamber. Lesioned birds remained isolated and were tested for up to 8 days post-lesion. Control lesioned birds underwent identical procedures with the stimulating electrode inserted into the pallial surface directly dorsal to RA.

To verify lesioning after behavioral testing, birds were deeply anesthetized with an intramuscular injection of pentobarbital (40 mg/kg in saline) and intracardially perfused with 4% PFA. The bird was then decapitated and the brain was removed from the skull and post-fixed in PFA overnight. For histological imaging, sagittal sections (100μm) were cut with a vibrotome (Leica VT1000S) and treated with a blue fluorescent Nissl stain (Life Technologies; 1:200 in PBS) for 2 hours. Slices were mounted on slides with Vectashield (Vector Labs) and imaged with a Zeiss LSM 510 confocal microscope using a 5x objective.
Using the same procedures as those preceding electrolytic lesions, we delivered bilateral injections 23nL, 0.5μM TTX or saline into RA of an adult male using a microinjector (Nanoject) (Figure 2.4C). Post-injection behavioral testing was performed within 6 hours and after 24 hours.

2.4 RESULTS

2.4.1 Call Interactions

While tracking the development of call interactions between a newly formed pair of birds, we observed that, initially, only a small proportion of calls were answered. Within a few days, however, tightly synchronized vocal coordination emerged, characterized by precise antiphonal calling and higher acoustic stereotypy (Figure 2.5). Exposing birds to a vocal robot that produces isochronous calls at a rate of 1Hz (ICs) induced prompt engagement, with birds answering 34±3.7% of ICs during a 10-minute session (means ± s.e.m. hereafter, n=7 males, 7 females, (Figure 2.6). These rates are comparable to those of established pairs within social groups(Ter Maat et al., 2014). Response latencies varied across birds (medians ranging
between 249-466ms), but were fairly stereotyped for each bird, with 50% of answers occurring within an interval of $249 \pm 23$ms. We observed no gender effect on answer latency or stereotypy (n=7 males & 7 females).

**Fig. 2.5: Call exchanges in a pair of zebra finches.**

**(A)** Calls exchanged between a male (blue) and a female zebra finch (red) over the first 4 days housed together. Dots represent the pitch vs. amplitude of short calls. Left: Green lines connect female calls to male answers (when answered within 500ms). Right: Green lines connect male calls to female answers. Inset: Proportion of partner calls answered. **(B)** Representative sonograms by day. Intervals between calls and answers are shaded in green.
Fig. 2.6: Responses to the vocal robot.

(A) Left: A male answering robot’s 1Hz isochronous calls (ICs). Right: a female answering robot’s ICs. Green lines connect robot calls to bird answers. Inset: Proportion of robot calls answered. (B) Male and female response latencies to ICs. Boxes denote interquartile ranges between first and third quartiles and inner lines represent the median latency for each bird. Whiskers: 5-95 percentiles.

2.4.2 Jamming Avoidance

The stereotypy of response latencies provided an opportunity to disruptively mask or “jam” call interactions and determine whether birds can adapt their call timing. After obtaining a reliable estimate of a bird’s response latency distribution, we programmed the robot to produce call patterns in which each repeating call cycle contains a second jamming call exactly when the bird is most likely to respond to the first call (Figure 2.2). Within a single 10-minute session, birds changed the latency of their responses to avoid
jamming (Figure 2.7): They produced calls with shorter latencies, longer latencies, or a combination of shorter and longer latencies (Figure 2.8).

---

**Fig. 2.7: Jamming avoidance.**

(A) A male’s call responses (blue) to robot ICs (gray). Top: Responses are aligned by IC cycle onset and presented in sequential rows over a 10-minute session. Middle: The male’s responses (red) to robot call cycles (grey) in which the robot produces jamming calls (yellow). Bottom: Distribution of the bird’s responses during the IC session (blue) and during the jamming call session (red). (B) As in (A) for a female.
Fig. 2.8: Jamming avoidance strategies and response probability across birds.

(A) Examples of jamming avoidance strategies in three birds. Birds' call timing for IC's (blue) and jamming call sessions (red) is plotted relative to a normalized jamming window (yellow). (B) as in (A) pooled across 12 birds.

To confirm that this effect is predictive rather than reactive, we incorporated catch trials (Tumer & Brainard, 2007), where the vocal robot intermittently produced only one call within a cycle (with 10% probability) during a session containing jamming calls (Figure 2.2). We could then estimate predictive jamming avoidance by measuring the proportion of responses that fall within the expected jamming window during catch trials (Figure 2.9). In 12 birds we found that the mean proportion of responses within the expected jamming window decreased from 35±2% for ICs to 13±3% during catch trials (Figure 2.9C, paired-t test, p<0.001). This avoidance was statistically significant in males (34±3% to 20±5%, n=6, paired-t test, p = 0.014) but greater in females (36±3.5% to 6±2.8%, n=6, paired-t test, p<0.001; t-test females vs. males, p=0.014), indicating more adaptive timing plasticity in the calling behavior of females.
Fig. 2.9: Assessing jamming avoidance with catch trials.

(A) As in 2.6B comparing call responses across 12 birds for ICs (blue) and catch trials during the jamming session (green). (B) Across sessions, the percent of responses that overlap with the jamming call is correlated with the percent of responses within the jamming window for catch trials (n=20 sessions, $R^2=0.29$, Slope=1.46 ± 0.54, *P=0.014). (C) Birds predictively reduce the proportion of calling within the jamming window during catch trials compared to ICs (n=12, paired t-test, ***P<0.001; % decrease in 6 males (red) vs. 6 female (blue), t-test, *P<0.05).

Investigating the time course of changes in answer latencies, we found that birds were able to avoid jamming even during the first 30-second block in a session (Figure 2.10A), suggesting that, like humans, they can detect temporal regularities and predict the timing of upcoming beats even after hearing only a small number of repetitions (Aa, Honing, & ten Cate, 2015; Repp, London, & Keller, 2011). However, we also observed a slower refinement of response timing across a session as some birds learned to fit an increasing proportion of their calls precisely in between jamming call pairs (Figure 2.10B).
Fig. 2.10: Jamming avoidance across a session.

(A) Across 12 birds, the average proportion of jamming was lower than expected by chance over the course of a session. (B) A female gradually changes response latencies over the course of a session (First third of the session: bright red. Last third: Dark red) to predictively call between the robot’s jamming call pairs (grey and yellow).

2.4.3 Dynamic Adjustments to Complex Rhythms

To determine whether birds adjust their response timing in a generalized way or can alter it dynamically in a context-dependent fashion, we programmed the vocal robot to produce jamming calls embedded in complex rhythm patterns. Presenting birds with non-isochronous call patterns, containing alternating jamming and non-jamming elements, we found that birds specifically decrease their answer latencies prior to jamming calls (Figure 2.11-2.12). We observed these dynamic adaptations of latencies within a single 10-minute session and over a range of tempos, with cycles of 1-2 seconds (each cycle is a repeating unit of a single call followed by a pair containing a jamming call, Figure 2.3).
For example, in a bird trained with a cycle of one second we can assess its anticipation of a jamming call by comparing responses within the 200ms after the robot calls, immediately before the jamming window. In that 200ms interval, the bird’s median response latency for single calls in the rhythm was similar to that for ICs (singles in rhythm: 140±1.1ms; ICs: 147±10.6ms). For those rhythmic calls that were followed by a jamming call, however, response latencies were shorter (125± 0.8ms; Figure 2.11). Therefore, the bird appeared to alternate between faster and slower responses, conforming to the rhythm produced by the vocal robot.

To rule out the possibility that this difference between latencies is due to a refractory effect from previous responses to jamming calls, we tested six birds for dynamic adaptation in cycles that were twice as long (Figures 2.3 and 2.12): Following single calls in the rhythm, median latencies were similar to those for ICs (single calls in rhythm: 203.6±4.9, ICs: 213.8±4.2, paired t-test, n=6, NS, p=0.27). For rhythmic calls that were followed by a jamming call, however, birds accelerated their responses during the same interval, in anticipation of jamming calls (Figure 2.12B, 193.5±4.0, ICs vs. jamming calls in rhythm, paired t-test, n=6, p=0.022; Singles vs. jamming, p=0.026). This predictive response acceleration occurred in both males and females (Figure 2.12), indicating that both can adapt their vocal response latencies in real-time to coordinate with complex rhythms.
Fig. 2.11: Rhythm adaptation.

(A) Robot call rhythm pattern shown in 2.3A (jamming calls in yellow). (B) A male’s call responses. Green: Responses to the single robot calls. Red: Responses to the jamming call pairs. Middle: Distribution of bird’s responses. Bottom: Cumulative responses within the 200ms following robot call onsets for single calls (green) and the first calls in jamming pairs (red), showing shorter answer latencies prior to jamming.
Fig. 2.12: Responses to rhythmic jamming calls.

(A) Robot rhythm pattern as in 2.11A but with slower 2s cycle and 250ms jamming latency (2.3B). (B) Cumulative responses to the slowly alternating rhythm in 6 birds (green, single calls; red, jamming pairs). (C) Median response latencies for each bird following ICs (blue), single calls in a rhythm (green), and the calls that precede jamming calls in a rhythm (red; Paired t-tests, n=6, *p <0.05).

In contrast to reflexive jamming avoidance responses, such as those seen in electric fish (Heiligenberg et al., 1996), individual zebra finches can use varied strategies to adapt their call latencies in an anticipatory way. This surprisingly high degree of sensory-motor flexibility may suggest pallial (i.e. cortical) involvement.
2.4.4 Effects of Blocking the Output of the Forebrain Song System

The forebrain song system exhibits premotor activity when males sing, but also prior to the production of unlearned calls (Ter Maat et al., 2014). We therefore suspected that the song system is involved in learned timing adaptation. This could be the case in females too: Zebra finch females do not sing and their song system is assumed to be vestigial (Wade & Arnold, 2004). However, because their jamming avoidance and rhythm adaptation performances were similar to - and often better than - those of males, their less anatomically developed song system could perhaps function to guide the coordination of their vocal behavior. In two males and three females we performed bilateral electrolytic lesioning of the robust nucleus of the arcopallium (RA), which is the common final output of the song system (Nottebohm, 2005). Lesioning RA profoundly affected the timing of responses in all 5 birds. Although the birds remained responsive to the robot calls (Figure 2.13A,B), the latency of their responses became much less stereotyped (Figures 2.13).

![Fig. 2.13: Response rates across experimental conditions.](image)

(A) Average calls per minute by condition (means±s.e.m). (B) Call rate differences between blocks of robot calls and intervening silent periods from sessions before and after RA lesions.
To quantify this effect on response time, we calculated a precision measure (Kosche, Vallentin, & Long, 2015) that estimates deviation from a uniform distribution of response latencies as a z-score. After lesioning, precision of responses to ICs decreases from $4.43 \pm 0.37$ to $1.91 \pm 0.25$ (paired t-test, $n=5$ birds, $p<0.01$, Figure 2.15-2.16). Response latency skewness also decreased, from $1.31 \pm 0.17$ to $0.34 \pm 0.13$ (paired t-test, $n=5$, $p<0.01$, Figure 2.15A, bottom), indicating higher symmetry and uniformity of response distributions within a cycle. Both effects were observed in all three females (Figure 2.14 & 2.15).

Fig. 2.14: Effects of RA lesions on responses to ICs.
Responses and probability distributions for ICs before (blue) and after RA lesions (orange) for 5 birds.

In sum, RA lesions appeared to impair the ability to couple the timing of calls to those of the robot, independent of the overall ability to respond. Birds that underwent identical surgical procedures, but received electrolytic lesions outside the song system, at the pallial surface (Figure 2.4, $n = 2$ females, 2 males) showed no changes in timing of their responses (Figures 2.16).
Fig. 2.15: Effects of RA and control lesions on response precision and skewness.

(A) Precision and skewness of responses to ICs in 2 males and 2 females before and after control lesions (n=4, paired t-tests, NS, P>0.65); in 2 males and 3 females after bilateral RA lesions (n=5, paired t-tests, **P<0.01). (B) Effects of control lesions in four birds. As in Figure S3C, before (blue) and after control lesions (purple) in 4 birds.
To determine whether we could obtain a similar effect reversibly, in one male we delivered bilateral microinjections of 23nL, 0.5μM TTX into nucleus RA. This caused a dramatic drop in response precision from 6.11 to 0.71 with a partial (21%) recovery of precision after 24hrs (Figure 2.16).

![Figure 2.14](image)

**Fig. 2.14: Effects of RA inactivation on responses to ICs.**

Bird’s responses as in Figs. 2.14, 2.15 before (blue), during (orange), and 24h after (light blue) bilateral RA inactivation with TTX.

The lesion-induced loss of response time precision made it more difficult to target responses for jamming. However, we were still able to test whether the birds exhibited any adaptation to jamming calls by comparing IC responses to those from jamming catch trials. The response probability distributions for ICs vs. jamming catch trials, across 5 lesioned birds, show no decrease in the proportion of calling within the expected jamming window (Figure 2.17). All five birds showed jamming avoidance prior to but not after lesioning (decrease from ICs to catch before: 22.6±3.8% vs. after RA lesions: -1.2±0.6%, paired t-test, n=5, p<0.01).
Fig. 2.17: Effects of RA lesions on precision and jamming avoidance.

Response distributions for ICs and catch trials pooled over 5 birds (as in Figure 2.9A) before (A) and after (B) RA lesions. (C) Percentage of calling within the jamming window for ICs vs. catch trials before (blue, n=5, *P<0.01) and after RA lesions (orange, NS). (D) RA lesions abolish jamming avoidance, measured as the difference in percent of calling within jamming window for ICs and catch trials (n=5, means, s.e.m, **P<0.01).
Fig. 2.18: Jamming avoidance in 4 control-lesioned birds.

(A) As in Figure 4F, cumulative response probability distributions for IC’s (purple) and jamming catch trials (green). (B) As in Figure 4G, percentage of calling within the jamming window for ICs and catch trials before (blue) and after control lesions (purple, 2 females and 2 male, paired t-test, P<0.05). 12 baseline birds in grey. (C) As in Figure 4H, difference in percent of calling within the jamming window for single and catch trails, before and after control lesions (n=4, paired t-test, NS, P>0.1).

Taken together, these results indicate that although the song system is not required for producing call responses, it plays an essential role in regulating the precise timing needed to predictively coordinate innate calls in both female and male zebra finches.
Fig. 2.19: Jamming avoidance is not correlated with response rates.

(A) Percentage of calls in the jamming window for ICs (left) and catch trials (right) is independent of percent of robot calls answered (N=10 birds). (B) As above, for 5 birds after bilateral RA lesioning (NS in all cases, P>0.1).
2.5 DISCUSSION

2.5.1 The Forebrain Song System and Anticipatory Call Timing

Three lines of evidence support the involvement of the song system in driving anticipatory call timing: First, innate short calls are preceded by pre-motor activity in song nucleus RA (Ter Maat et al., 2014). Second, electrolytic lesions of RA, as well as transitory inactivation of that structure, profoundly reduce the timing precision of call responses without reducing response rates. Third, call timing plasticity for jamming avoidance is abolished by RA lesions. Ter Maat et al. found that RA neurons fire just milliseconds before short calls are produced even though midbrain and brain stem nuclei can generate calls independently of the forebrain (Simpson & Vicario, 1990; D S Vicario & Simpson, 1995). Together with our current results, these findings indicate that the precise timing required for coordinated calling is regulated either by song nucleus RA or by upstream song nuclei.

RA lesions or inactivation are likely to compromise the surrounding ‘RA cup’ (e.g., ventral intermediate archopallium, AIV) and associated fibers. However, the RA cup alone cannot regulate call timing with the milliseconds precision we observed, because there are no projections from the RA cup to the brainstem vocal nuclei (Poirier, Boumans, Verhoeye, Balthazart, & Van der Linden, 2009). While the direct regulation of call timing is likely to be driven by nucleus RA, the descending forebrain auditory pathway, which converges on RA cup, is also closely associated with vocal learning (S. W. Bottjer & Altenau, 2010; Mandelblat-Cerf, Las, Denisenko, & Fee, 2014; Vates, Broome, Mello, & Nottebohm, 1996). Therefore, RA cup may still be involved in the evaluation of timing error, or in predicting the timing of the robot’s calls. In sum, bilateral RA lesions eliminated the most direct path between the forebrain vocal centers and the primary brainstem vocal motor nuclei (nXIIIts, RAm, PAm), hence leaving only
non-specific tracts that lack the speed and precision of the monosynaptic drive from the forebrain through RA. In this case, midbrain motor centers (most likely the dorsal medial nucleus, DM) and ascending auditory pathways (dorsal lateral nucleus of the mesencephalon, MLd) can apparently still trigger the production of call responses albeit without the predictive timing control provided by the forebrain (Fukushima & Aoki, 2000; Simpson & Vicario, 1990).

The song system’s involvement in adaptive call timing suggests that vocal learning and call coordination rely on some of the same mechanisms. In both song learning and call coordination vocal production changes in reference to a sensory ‘template’. In coordinated calling the bird does not reproduce the template, but, as with syncopation in music (Fitch & Rosenfeld, 2007), timing is predictively shifted in relation to a pattern. Interestingly, the differences between learning to imitate vocal sounds and coordinating calls can explain specific anatomical differences (S. Bottjer, Miesner, & Arnold, 1984; Maguire, Schmidt, & White, 2013; Poirier et al., 2009; Roberts, Klein, Kubke, Wild, & Mooney, 2008; Wade & Arnold, 2004) between the male and female zebra finch song system: We propose that the female song system functions in call coordination, namely in predictive timing. The male song system functions in vocal learning, which includes both timing and spectral plasticity. Accordingly, song nuclei that are involved exclusively in learning spectral structure should not exist in the female song system. Indeed, the largest song system nucleus, Area X – which is absent from the female brain- is not required for learning time structure in males (Ali et al., 2013; Wade & Arnold, 2004). In contrast, nucleus HVC (used as a proper name), nucleus RA, and RA cup are present in both males and females (Roberts et al., 2008; Simpson & Vicario, 1990; D S Vicario & Simpson, 1995; Wade & Arnold, 2004) and have all been specifically associated with temporal aspects of song learning (Ashmore, Renk, et al., 2008; Long & Fee, 2008; Mandelblat-Cerf et al., 2014; Vates et al., 1996). The song system of the female zebra finch, therefore, appears to
be specialized for vocal plasticity with respect to timing.

2.5.2 Phylogenetic Perspective on Vocal Learning

The finding that the female song system can support vocal timing plasticity has several implications for our understanding of the origins of vocal learning. Although females have weak HVC to RA projections they do have descending connections from HVC shelf to RA cup (Vates et al., 1996). This descending auditory pathway is common to vocal non-learners that lack RA (e.g. pigeons) and vocal learners with RA homologues (e.g. budgerigars). However the descending auditory pathway does not directly project to subcortical vocal production centers in any of these cases. In at least one sub-oscine vocal non-learning species, the eastern phoebe, this pathway exhibits RA-like auditory-motor properties and does project to brainstem vocal production centers. Lesions to this intermediate structure produce subtle but reliable changes to the unlearned song (Liu et al., 2013). In this comparative light, it appears that the development of a rudimentary RA, in addition to the descending auditory pathway, was a pivotal step in the evolution of vocal learning (Liu et al., 2013). Evidence here that the rudimentary female RA is necessary for the coordination and plasticity of call timing is consistent with the view that this premotor nucleus was involved in the cortical modulation of unlearned vocalizations before it functioned in song learning.

Vocal learning and auditory-motor synchronization (e.g., dancing to music) are considered to be evolutionarily and mechanistically linked, perhaps stemming from shared social origins (Fitch, 2011; Patel & Iversen, 2014; Schachner, Brady, Pepperberg, & Hauser, 2009). Our results provide direct evidence for a similar link between song learning and call synchronization except that here both phenomena are vocal and are supported by the same sensorimotor pathway. Zebra finches can distinguish between various isochronous as well as
non-isochronous temporal patterns (Van Der Aa et al., 2015). In human beat perception, pre-motor cortex is implicated in the analysis and prediction of sound timing regularities (Fitch & Rosenfeld, 2007; Patel & Iversen, 2014; Wang, Uhrig, Jarraya, & Dehaene, 2015). Evidence for an analogous process in the song system comes from the coordinated duets of male and female plain-tailed wrens, in which pre-motor activity may encode information about the rapidly alternating partner’s timing intervals (Fortune et al., 2011). The loss of timing precision and jamming avoidance after RA lesioning in zebra finches may be the consequence of disrupting a similar auditory-motor timing prediction mechanism that guides call coordination. Furthermore, comparative evidence that female singing is ancestral in songbirds (Odom, Hall, Riebel, Omland, & Langmore, 2014) suggests that the song system did not initially evolve for sexually dimorphic courtship, guided primarily in males by female mate choice. Rather, the song system was more likely to have functioned in coordinating vocalizations in a more general social context.

Considering that coordination of unlearned vocalizations is relatively widespread across taxa, we propose that there is a mechanistic continuum from vocal coordination to vocal learning, spanning a functional gap that has separated vocal non-learners from vocal learners. In this view, the primitive function of the forebrain song system was to enable plasticity in the timing of social behavior. Such a general-purpose sensorimotor communication system could facilitate vocal interactions on different time scales: In the case of calls, it enables predictive modulation of vocal timing, where as in the case of learned songs, plasticity extends to developmental changes in vocal structure. This social coordination hypothesis can explain some of the evolutionary convergence in the mechanisms of vocal learning and auditory-motor synchronization seen across mammalian and avian brains (Fitch & Rosenfeld, 2007; Fitch,
2011; Liu et al., 2013; Patel & Iversen, 2014; Pfenning et al., 2014; Schachner et al., 2009; Takahashi et al., 2013).

2.6 AUTHOR CONTRIBUTIONS

3.1 SUMMARY

All vocal learners, from songbirds to human infants, require social and vocal interactions during development for achieving basic communication skills, such as spoken language in humans. In songbirds, complete social isolation during the sensitive period for vocal learning results in irreversibly abnormal (isolate) songs. Here we tested if social interactions during development are required for birds to develop the capacity to synchronize and adapt their call timing to that of a partner bird robot. We raised male zebra finches in complete social isolation, but we trained them to imitate a song using operant song playbacks. As expected, these birds produced fairly normal songs. However, they never had the opportunity to interact with other birds via short contact calls during development. We found that, as adults, these birds were responsive to the robot contact calls, but were extremely imprecise in the timing of their responses. Further, these birds were unable to avoid disruptive jamming. Interestingly, these effects are extremely similar to those observed after blocking the forebrain song system in normal birds. We conclude that social interactions during development are necessary for zebra finch males to develop the capacity to match and adapt the timing of their calls. Further, the capacity to synchronize calls must be acquired independently from that of song learning.

3.2 BACKGROUND AND RATIONALE

One of the hallmarks of vocal learning is the developmental sensitive period in which early auditory and social experience, combined with vocal practice, has a strong impact on the
ontogeny of vocal behavior. In zebra finches, exposure to song playbacks during this sensitive period in socially isolated birds is sufficient to induce vocal learning, often resulting in fairly accurate imitation of the presented song playbacks (Hultsch & Todt, 1982; Nelson & Marler, 1994; Tchernichovski, Mitra, Lints, & Nottebohm, 2001). In addition to this auditory experience, however, social experience may also play an important role in guiding vocal development (Derégnaucourt & Gahr, 2013; Goldstein, King, & West, 2003; Honarmand, Riebel, & Naguib, 2015; Kao et al., 2008; Kojima & Doupe, 2011; Tchernichovski & Nottebohm, 1998; S. C. Woolley & Doupe, 2008; S. M. N. Woolley, Hauber, & Theunissen, 2010).

Previous studies in songbirds have focused on developmental song learning, while the possible role of early experience on the capacity to coordinate calls has not been studied as thoroughly, despite evidence for developmental changes in call interactions (Chow, Mitchell, Miller, & Miller, 2015; Mulard et al., 2010; Takahashi, Fenley, Teramoto, Narayanan, & Jeremy, 2015). Here we examine the role of social experience in the development of regulated vocal plasticity using the vocal robot approach presented in the Chapter 2 (Benichov, et al., In Review). We investigated the effect of prolonged developmental social isolation in birds that were trained with operant song playbacks. This way, we provided the birds with the opportunity to imitate a song, but deprived them from any experience with exchanging calls. This allowed us to ask if and to what extent birds need to experience contact call interactions as juveniles in order to develop vocal plasticity that would allow them to coordinate their calls. Being able to quantify this capacity, and given the that song system appears to be involved in the coordination of calls, we compared the effect of social isolation during development to the effects we observed after lesioning the forebrain song system. This comparison is important for judging to what extent the song system has to be trained during development, in order to be functional in call synchronization in adulthood. Making this comparison in birds where the song system
matured and functioned normally with respect to song learning, allowed us to judge if and to what extent song learning and vocal coordination capacities are functionally segregated in the song system.

3.3 METHODS

3.3.1 Social Isolation During the Sensitive Period for Vocal Learning

Six male birds were reared by their mothers until 30 days old. Birds were then placed singly in sound attenuation chambers and were kept socially isolated. From day 40 to 90, juveniles were trained with operant (key peck triggered) song playbacks. All six birds copied either some or all of the song syllables presented to them, and developed a fairly normal song (Figure 3.1). Note, however, that the operant playbacks included only songs, and therefore the birds never heard calls and never had any experience with calls interaction after day 30. Birds were kept isolated, and were introduced to the vocal robot partner when adult (ages 100-178 post-hatch).

3.3.2 Assessing Call Coordination Capacity in Social Isolates

Methods here are identical to those presented in chapter 2, briefly: each bird was paired with our vocal robot system and was delivered Isochronous Calls (IC’s) at a rate of 1hz across a 10 minute session as in Benichov et al (in review). Probability distributions of response latency for ICs were calculated across a 1s IC cycle and were used to determine the jamming window for the subsequent session. Additionally, call response latencies - relative to the onset of IC cycles - were used to compute call timing precision and skewness (see Chapter 2.3.3).
In a second session with the vocal robot, each bird was presented with pairs of jamming calls, over 10 minutes. Adaptation to these jamming calls was assessed using catch trials: single robot calls that were occasionally produced during a session containing jamming pairs. We quantified anticipatory change in call timing based on the percent of calls within the jamming window of those catch trials compared to the same window during the IC session.

3.4 RESULTS

3.4.1 Effects of social isolation on call response timing

In six birds that were socially isolated from day 30-90 post-hatch, we tested call responses to 1Hz isochronous calls (ICs) produced by the vocal robot. As shown in Figure 3.1, the distribution of response timing in the social isolates is much broader than that of birds raised with social experience. This effect can be clearly seen in the pooled probability distributions across each group of six birds. To assess these results quantitatively, we compared response latency precision and skewness (see Chapter 2) across social isolates and socialized birds (Figure 3.2). The mean response precision for six socialized birds was $4.46 \pm 0.52$ compared to $1.92 \pm 0.49$ for six isolated birds (t-test, **$p<0.01$). These groups also differed in the skewness of their responses within an IC cycle (Socialized: $1.31 \pm 0.12$; Isolates: $0.10 \pm 0.19$; ***$p<0.001$).
Fig. 3.1: Song learning and call responses in a developmentally isolated adult male.

Left: Sonogram of a model song playback (top) and a social isolate male’s song (bottom). Right: The same social isolate male’s call responses to robot ICs.
Fig. 3.2: Social isolates’ call response timing compared to socially reared birds.

Raster plots of responses to isochronous calls in each of the isolates on the left. Right: Six socialized birds, matched by responsiveness.
3.4.2 The effect of social isolation on call timing precision is similar to that of blocking the forebrain song system

The low precision in call timing that we observed in socially isolated birds is reminiscent of that which we observed in socialized birds whose song system output was blocked by RA lesions or inactivation with TTX (see Chapter 2.3.7). We therefore tested if the answer precision of the socially isolated birds differed from those of the RA lesioned (or blocked) birds. We found that the two effects are indistinguishable (Figure 3.3A). The precision for RA blocked birds was 1.32 ± 0.33, compared to 1.92 ± 0.49 in the socially isolated birds (n=6; t-test, ns, p=0.34). Similar results were obtained with skewness: (RA blocked: 0.44 ± 0.15, socially isolated: 0.10 ± 0.19, n=6, ns, p=0.20).

A possible explanation to this effect is that the socially isolated birds are not used to social interactions and are simply less engaged than colony raised birds. To test it we compared the rates of responses to robot calls across socially isolated birds and those of socialized birds, and in birds with RA lesions (Figure 3.3B, Isolates: 22.60 ± 3.86 calls per minute, n=6 vs. Socialized 20.37 ± 3.77, n=6, ns, p=0.69; RA-Lesioned: 26.80 ± 5.20, n=5, ns, p=0.52). As show, we observed no apparent (or statistically significant) differences across groups. Therefore, social isolates, like socialized birds without a functioning song system, appear to be less capable of controlling the precise timing of their responses.
Fig. 3.3: Features of isolate call timing resemble birds with RA lesions.

(A) Precision and skewness of responses to ICs for socialized birds, socialized birds with RA lesions or inactivation, and social isolate birds. (B) Call response rates to robot ICs across Socialized, RA-lesioned, and Isolate adults.

3.4.3 Effects of developmental isolation on adult jamming avoidance

Finally, we tested for effects on the jamming avoidance response by exposing the social isolates to pairs of jamming calls, with intervals that match the expected call back latency of each bird. Although three of the birds showed some adaptation in their call timing, across birds, the mean percent of calls in the jamming window for ICs did not differ significantly from that of jamming catch trials (17.83 ± 3.37%, during baseline vs. 13.00 ± 5.52% during catch trials; n=6, paired t-test, ns, p=0.17), suggesting that adaptation to jamming calls is deficient in these birds (Figure 3.4).
Fig. 3.4: Social isolates exhibit impaired jamming avoidance.

Jamming avoidance capacity of social isolates does not differ from that of socialized birds with RA lesions.

As with response precision and skewness for ICs, Isolate jamming avoidance (difference in percent of calling in the jamming window for ICs and jamming catch trials) was indistinguishable from that of birds with RA lesions (Figure 3.4; Isolates: 4.83 ± 3.01%, n=6; RA Lesioned: -1.20 ± 0.58%, n=5; t-test, ns, p=0.11). Additionally, isolate birds exhibited significantly less jamming avoidance than socialized birds before RA lesions (Socialized birds: 22.60 ± 3.79%, n=5, t-test, **p<0.01).
As shown in Figure 3.5, the response probability distributions for isolates indicate little adaptation to the jamming calls. These results suggest that social interactions during development are necessary for the song system to function in regulating coordinated calling behavior.

**Fig. 3.5: Response distributions across social conditions.**
Response probability distributions for IC’s and Jamming catch trials (green) plotted relative to a normalized jamming window onset.
3.5 DISCUSSION

Male zebra finches that were socially isolated during development had significantly lower response time precision and more uniform response latency distributions across IC cycles than birds with social experience during development. Furthermore, isolates behave like birds that lack forebrain control of call timing in terms of precision, skewness, and adaptation to jamming calls. It is worth noting that isolate response rates across a 10-minute session were in a range similar to that of socialized birds. This, along with the lack of apparent jamming avoidance and similarity to RA-blocked birds may suggest that disordered response times for ICs are more likely to reflect an inability to socially regulate call timing rather than an unwillingness to conform to an imposed call pattern.

Finally, the strong deficiency in call coordination capacity that we observed in social isolates occurred despite the fact that those birds developed fairly normal songs, including birds that preformed perfect imitation of the operant playbacks they heard. Together, these results indicate that although the song system is involved in both song learning and in calls coordination, the development of call coordination and song learning must be functionally segregated in the song system. This effect will require further studies to determine if this segregation is only functional or also anatomical.

Future studies should test for such an effect in females, who lack developmental song learning. This could determine if a vocal developmental sensitive period exists outside the context of song learning. For further characterization of the role of early social experience, future studies should test juvenile birds at various stages of development and periods of isolation to determine the boundaries of the critical period.
The evidence presented here raises a deep and readily testable question of whether the call coordination capacity can be recovered through “social rehabilitation” in developmentally isolated adults. If so, we can conclude that social experience is important for vocal flexibility but that a developmental sensitive period may not be so “critical” for learned regulation of vocal timing. If belated social experience has little effect on these deficits in call coordination it may reflect anatomically underdeveloped forebrain control mechanisms that are not easily modified in the adult brain. This would be consistent with the observed similarity between isolates and RA lesioned birds. In either case, we stand to gain purchase on the extent to which a basic aspect of vocal learning, namely socially modulated vocal timing, is tied to a more general developmental program.
Chapter 4: Vocal coordination beyond the dyad

4.1 SUMMARY

The capacity of zebra finches to adjust their call timing while interacting with the vocal robot raises the question if, and to what extent, does a bird take into account the behavior of a third party while interacting with a partner bird. To examine this question we built wireless audio transmitters, which allowed us to separate sound sources while two birds are simultaneously interacting with the vocal robot and with each other. Here we present a qualitative investigation, where we selected two birds with very similar answer latencies when interacting with the robot one-to-one. Placing the two birds together with the robot, we found that, within minutes, they managed to avoid jamming with each other as they responded to robot calls by changing the latencies of their answer calls. We then programmed the robot to jam the answer calls of one bird, and found that both birds adjusted the timing of their calls, to generate an uninterrupted sequence of group calling. These qualitative results suggest that birds are capable of adjusting the timing of their calls with respect to more than a single partner bird.

4.2 BACKGROUND AND RATIONALE

Until now, we have focused on call interactions between pairs of birds or between individual birds and the vocal robot. Within these dyads, vocal timing adaptation can occur in anticipation of a single partner’s calls (Benichov et al., under review). It is important to consider, however, that zebra finches normally live in large colonial groups and communicate within a complex acoustic environment (Elie et al., 2011; Waas, Colgan, & Boag, 2005; Zann, 1996). Although group vocal activity may seem to be comprised of unstructured chattering, call source
separation using bird-mounted transmitters has revealed that specificity of responses and tight call coordination can exist in the apparent cacophony (Anisimov et al., 2014; Elie et al., 2010; Ter Maat et al., 2014; Vyssotski et al., 2013). It remains unclear how pairs of birds go about organizing their interactions in a challenging acoustic context and whether coordination dynamics change in the presence of environmental interference.

Here it is important to make a distinction between the effects of informational masking (disruptive overlapping signals from conspecifics) vs. energetic masking (i.e. non communicative acoustic noise) in driving adaptive vocal changes. In the case of the environmental (e.g. anthropogenic) noise, vocal signaling can be adapted in terms of pitch and amplitude to avoid masking (Brumm & Zolinger, 2013; Brumm, 2006; Burghardt et al., 2008; Jeffrey Cynx et al., 1998; Jeffrey Cynx & Gell, 2004; Katti & Warren, 2004; Slabbekoorn, Peet, & Grier, 2003; Tumer & Brainard, 2007). In the case of interfering informational masking, signal expectation and contextual cues may further modulate vocal timing (Brumm, 2006; Naguib & Mennill, 2010; Searcy & Beecher, 2009; Yang, Ma, & Slabbekoorn, 2014). Anticipation of a partner, therefore, may provide a partial solution to problem of signal detection in noise, beyond reliance on the strategy to “out-signal” the competition.

In this broader context, the temporal changes in calling that we observed at the level of pairs may also promote social coordination of higher order, at the scale of the colony. As previously discussed, shifts in call timing reflect predictions for avoiding interference with a partner. In addition, the underlying ability to predict an incoming signal may also aid in recognizing the calls of particular individuals based on their timing in the presence of competing calls and thereby maintain specificity of the communication channel. Indeed expectation and
contextual cues have been shown to provide robust perceptual benefits in human auditory stream segregation as well as the perception of masked speech (Benichov, Cox, Tun, & Wingfield, 2012; McGowan, 2015; Winkler, Denham, Mill, Bohm, & Bendixen, 2012).

The importance of a partner’s temporal patterns has been noted in relation to antiphonal vocalizations of songbirds and mammals (Fortune et al., 2011; Ghazanfar et al., 2002). Evidence from antiphonal calling in marmosets suggests that the timing of their vocal interactions exhibits coupled-oscillator dynamics (Takahashi et al., 2013). The use of timing as a mechanism for specifying a communication channel between particular signalers and receivers in a shared medium is also intentionally employed in the engineering of computer and distributed telecommunications networks (e.g. “heartbeat” algorithms and time-division multiplexing (Andrews, 1991; Scott, 1998)). Interestingly, coupled timing mechanisms also appear to be at work in human speech processing, where low-level segmentation of the auditory stream involves realignment of intrinsic cortical theta rhythms to match the ~4-8hz oscillations of a specific incoming speech signal envelope, both without and within noise (Giraud & Poeppel, 2012; Luo & Poeppel, 2007; Peelle & Davis, 2012; Zion Golumbic et al., 2013). In each of these cases, interacting nodes or individuals possess local “clocks” that can be synchronized with those of others they communicate with. These coupled oscillators can thereby use their own representation of time to “predict” and more easily extract the signal of a partner when it is interleaved in time with competing signals (Figure 4.1).
If the above strategy is employed in zebra finch groups, interacting pairs should possess a shared representation of time and expectations that would then help them maintain specificity of the communication channel and actively avoid interference. An ability to cooperatively adapt vocal timing in the presence of a third party would be evidence of a shared representation of timing. To address the questions of whether and how pairs of zebra finches adjust the timing of their calls when a third calling individual is present, we combine the call source separation afforded by wireless recording systems with the empirical control of the vocal robot and assess the temporal effects on call interactions.

4.3 METHODS

4.3.1 Wireless audio transmitters for call source separation

We constructed several wireless audio transmitters at the Max Planck Institute for Ornithology, with the aid of their designers Hannes Sagunsky and Andries Ter Maat (Ter Maat, 2014). These ~0.6 gram devices (including V10 1.4V battery) can be mounted via elastic leg loops to continuously record the vocal activity of individual freely moving birds. The large
amplitude differences between vocalizations detected from the bird wearing the transmitter compared to those of birds nearby, allow for more reliable source separation than is possible with a single field microphone or by comparing signals across a stationary array of microphones. Birds were given 10 days to acclimate to the devices before experiments.

The non-overlapping frequency modulated signals from transmitters were captured by separate AOR5000 radio receivers (AOR, Ltd., Japan). To compensate for transmitter frequency drift, the peak of each FM carrier signal was tracked by a DSOX2004A oscilloscope (Keysight Technologies, Santa Rosa, CA) and each receiver was automatically tuned to the optimal frequency by DSOAsio V2.0 tracking software (Max Planck Institute for Ornithology). These signals, along with audio from a nearby C-2 cardioid condenser microphone (Behringer, Germany), were routed into a XENYX Q502USB audio mixer (Behringer, Germany) and recorded to time-synched files using Sound Analysis Pro (Figure 4.3).
Fig. 4.2: Source separation with miniature audio transmitters.

(A) Miniature wireless audio transmitters (see Ter Maat, 2014). (B) Transmitter detail. (C) Audio waveforms of a bird-robot interaction. Signals from the bird-mounted transmitter (blue) and a nearby cardioid condenser microphone in black show that the transmitter specifically senses calls from the bird (blue shaded) and not the robot calls (gray shaded).

4.3.2 Intrusive Robot Calls

To test how two interacting birds respond to independent call patterns produced by a third bird, we presented a pair of adult female birds — each wearing a wireless transmitter - with calls from the vocal robot (Figure 4.3). Each bird in the pair was first individually presented with 10 30-second blocks of isochronous robot calls, presented at a rate of 1Hz (ICs, Chapter 2). Birds were then housed together and again presented with ICs. Time windows of maximum jamming probability were calculated and in a subsequent session, jamming robot calls were presented (section 2.3.4). With this system, we asked if each bird’s call response times, relative to robot calls and to the other bird’s calls, are affected by the temporal pattern of robot calls.
Recording the vocalizations of individual freely-moving zebra finches during call interactions. The wireless transmitter method (Ter Maat, 2014) employs contact microphones that allow for high specificity of audio signal streams. Individual acoustic signals are transmitted over different FM channels. Vocal Robot delivers ICs and Jamming call patterns to the pair.

Fig. 4.3: Recording from a pair of birds interacting with the vocal robot.
4.4 RESULTS

4.4.1 Birds avoid jamming with each other as they respond to robot calls.

After each bird was allowed to acclimate to the transmitter for 10 days, they were individually paired with the vocal robot and presented with isochronous calls at a rate of 1 per second (ICs, see Chapter 2). When alone, the two females had similar responses to Robot ICs, with the highest proportion of responses occurring with a latency of ~235ms (Figure 4.4, top panel). Because the birds individually responded to the robot with similar latencies, the question became whether they would adjust these latencies to avoid jamming with each other when presented with ICs while together. Figure 4.4 (middle panel) shows that Female 1 (red) responds more quickly to robot calls in the presence of Female 2 than she did when alone. Conversely, Female 2 (blue) delayed her responses. As a result, both birds effectively avoided jamming with one another while answering the robot. The outcome was a tendency to produce a sequenced order of calling (e.g. Robot ➔ Female 1 ➔ Female 2). The ordered cascade of responses that resulted from this manipulation are consistent with preliminary observations of small groups of calling zebra finches (Vyssotski et al., 2013).

4.4.2 Jamming robot calls are incorporated into a sequence of group calling

To determine if Female 2 (blue) shifted simply because she preferred to respond to Female 1, we further challenge the pair by programmed the robot to produce a second jamming call within each cycle, at an interval most likely to interfere with Female 2’s calls (Chapter 2.3.4). In this context, Female 2 (blue) predictively avoided the robot’s jamming calls (yellow) primarily by further delaying her responses within each cycle, as measured during catch trials containing
a single robot call (Figure 4.4, bottom panel). Interestingly, Female 1 reverted back to her slower response distribution when no longer at risk of being jammed by Female 2. This asymmetry of adaptation resulted in the mutual accommodation of the jamming calls.

Fig. 4.4: Bird responses to the vocal robot alone and in a pair.
Top: Distributions of responses to robot ICs (grey) when Female 1 (red) and Female 2 (blue) are alone with the robot. Red and blue shaded intervals represent windows of maximum likelihood of jamming for Female 1 and 2 respectively. Middle: Call distributions for Females 1 and 2 when both birds are presented with robot ICs while paired together. Bottom: Response distributions from jamming catch trials when the pair is presented with Robot jamming calls (yellow) targeting Female 2.

Across paired conditions, there appear to be trends toward complementary, and roughly periodic, rising and falling of the response distributions of the two birds. The distribution of minor peaks may hint at a more complex alternating sequence. The exceptions seem be when both
birds reduce calling to accommodate the robot. In sum, when in the presence of jamming robot calls, this pair of female birds effectively sequenced their call order such that they appeared to alternate their responses to the robot calls (e.g. Robot ➔ Female 1 ➔ Robot ➔ Female 2). It is worth noting that although these birds had been housed together previously, they were able to not only coordinate with each other in the presence of robot calls, but were also able to incorporate robot calls into their call order within a single ten minute session.

4.5 DISCUSSION

We found that birds in a pair are capable of cooperatively adjusting the timing of their calls to avoid overlapping with one another as they respond to a vocal robot. In this case, one bird shortened its response latencies whereas the other bird tended to delay its responses. Furthermore, when one of the birds’ responses were targeted by jamming robot calls, both birds made timing adjustments to incorporate the robot into a larger patterned sequence of calls. These results, although only qualitative at this stage, may be indicative of a shared representation of vocal event timing across birds, characteristic of coupled oscillators. This may also explain why timing adaptation can occur within minutes, whereas changes in spectral features such as pitch can require days of repeated exposure to aversive reinforcement (Tumer & Brainard, 2007). The reliance on a shared temporal pattern may simultaneously allow individuals to maintain engagement while also providing a predictive model for driving changes in call timing to avoid interference, with the partner or others within a group.

Evidence that birds can adapt vocal timing not only in anticipation of single partner but also within a group of multiple callers raises the question of whether order of calling represents position in a social hierarchy. The answer is outside the scope of the current study. However,
anticipation of conspecifics and control over the timing of responses in order to yield or concede temporal bandwidth could conceivable underlie both hierarchical and affiliative pair calling. In either case, given the relatively simple and spectrally inflexible acoustics of short calls, plasticity in vocal timing as well as the ability to detect temporal patterns (Van Der Aa et al., 2015) is likely to play an important role in group communication. We propose that vocal plasticity in call timing may serve to maintain the structure of vocal interactions in the presence of environmental or conspecific interference. Furthermore, the ability to anticipate a partner or temporal regularities in the acoustic environment could represent an adaptive strategy that bypasses the need for an amplification arms race in order to achieve effective communication. The mechanisms that would underlie a predictive shared representation of time that can drive changes in calling, therefore, may be shared with - or have provided an evolutionary basis for – more complex forms of auditory motor synchronization and vocal learning.
Appendix

A1 ADDITIONAL METHODS DEVELOPED

A1.1 Responsive Robotic Vocal Interface

In addition to delivering leading call patterns, the vocal robot can engage in vocal interactions by responding to a bird adaptively using real-time acoustic analysis of the bird’s calls (e.g. matching the pitch and rate of the bird’s calls). The robot can interact recursively with the bird to generate sustained call bouts (Figure A1-A2)

Fig. A1: Vocal robot design and operation in responsive mode.

A) A bird calls (red sonogram) and the system processes the audio signal, detects pitch and harmonic structure, and triggers a robot call selected from a library of recorded sounds. The system can dynamically adjust the properties of the robot calls (e.g. delay, pitch, and amplitude) based on the bird’s intercall interval and the acoustic features of triggering calls. (B) 3D-printed plastic model of adult male zebra finch, perched on a servo motor that is driven by the amplitude of the delivered robot calls, synchronizing movement with sound. (C) Interactive call bout (bird calls, red; robot calls, blue).

A1.2 Robot responds to one bird in a pair

To test whether a pair of birds can collectively learn to adapt to contingent calls from a third party, we can record two birds interacting while the robot selectively responds to the calls of one of the birds. By programming the robot to respond to the targeted bird with the same
latency as the non-targeted bird (jamming window determined prior to the experiment), we can specifically ask if the non-targeted bird predictively avoids jamming with the robot when it responds to the targeted bird with the same stereotyped latency. (Figure A2).

**Fig. A2: Vocal robot selectively response to one bird within a pair.**

Audio signal from a targeted bird (red) is fed into the AVI vocal robot software to trigger Robot calls (green) that programmatically respond to that bird in a manner that would interfere with the non-targeted bird’s responses (blue).

**A1.3 Call modulation in response to predictable environmental noise**

To determine if birds exhibit predictive temporal modulation of call interactions in the presence of energetic masking (as they appear to do with social/informational masking from robot calls) we can present freely interacting pairs with white noise with periodic amplitude modulation. We can ask if birds prefer to call in the low amplitude gaps in the white noise. We
can assess predictive plasticity by phase shifting the temporal pattern of the noise and measuring the effect on call timing. (Figure A3).

**Fig. A3:** Spectrogram of a periodic white noise pattern with a phase shift.
**A1.4 3D accelerometer & Heart rate**

In addition to assessing interactions through vocal behavior alone, we have used a Neurologger device to continuously monitor the physiological correlates of call coordination. The Neurologger is a 1 gram chip that can be mounted to the back of the bird and locally write data to flash storage (Scriba et al., 2012; Alexei L Vyssotski et al., 2006) (Figure A4). This design permits birds to move freely within the sound-attenuated chamber, allowing them to exhibit a wide range of social behavior. The Neurologger records continuously from three sensors: an accelerometer, EKG leads, and an infrared (IR) detector. The 3D-accelerometer provides data to reconstruct movement trajectories, calculate proximity to the robot, and assess energy expenditure. EKG measures changes in heart rate (Figure A4C, A6). Lastly, the IR sensor will detect brief pulses of IR light delivered at key events in the experiment. IR pulses provide timestamps with which to precisely align events across multiple devices as well to the onsets of robot calls delivered to the bird (Figure A6).

![Fig. A4: Recording EKG with a Neurologger.](Image)

**(A)** Richard Hahnloser and Alexei Vyssotski have developed Neurologger devices, and they agreed to share this technology with us (Vyssotski, 2006). **(C)** Detailed EKG trace from a freely-moving bird.
Fig. A5: Recording EKG across an interactive call bout with a vocal-robot.

(A) *Top:* Using a bird-mounted device we recorded EKG to measure changes in instantaneous heart rate during vocal interactions with the robot. *Bottom:* Call bout duets between bird and robot over 1 minute. (B) Preliminary results demonstrate that the patterns of variability in heart rate correspond to temporal features of interactive call bouts with the vocal robot.
References


