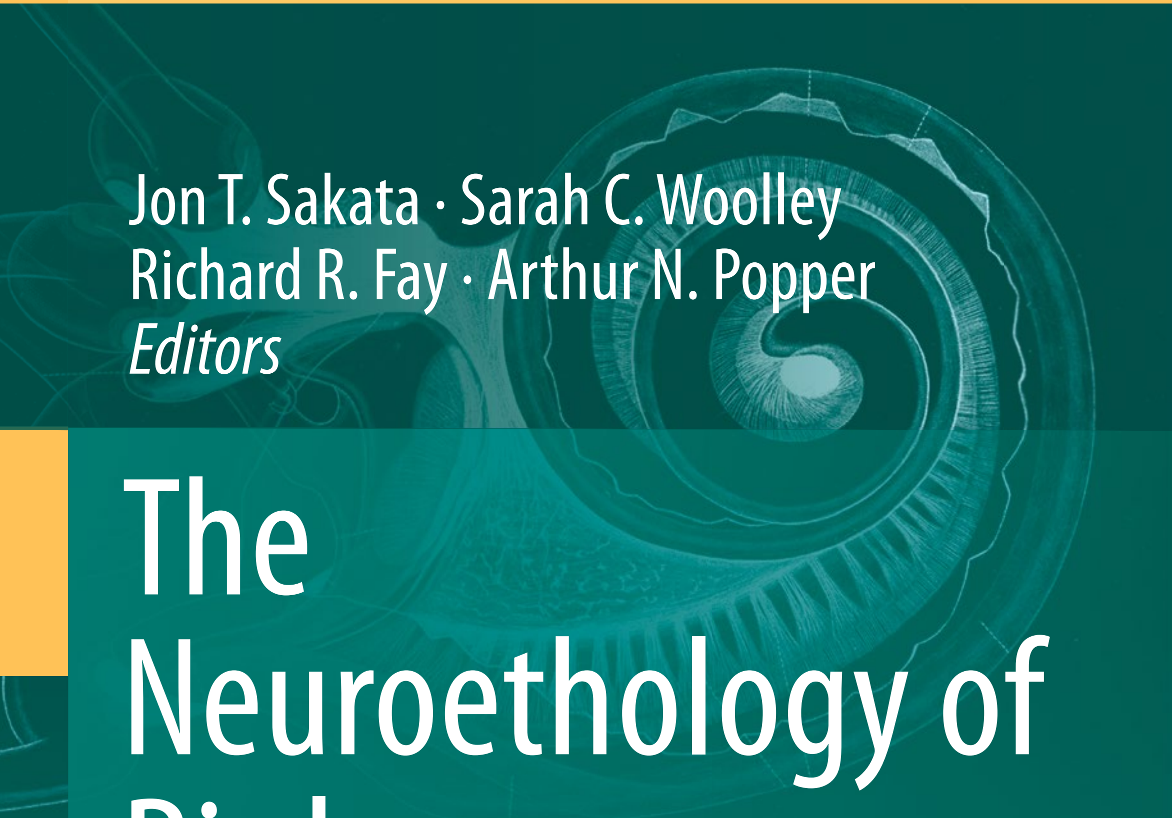
Springer Handbook of Auditory Research

Jon T. Sakata · Sarah C. Woolley
Richard R. Fay · Arthur N. Popper
Editors

The Neuroethology of Birdsong



ASA PRESS

Springer

Springer Handbook of Auditory Research

Volume 71

More information about this series at http://www.springer.com/series/2506

The ASA Press

ASA Press, which represents a collaboration between the Acoustical Society of America and Springer Nature, is dedicated to encouraging the publication of important new books as well as the distribution of classic titles in acoustics. These titles, published under a dual ASA Press/Springer imprint, are intended to reflect the full range of research in acoustics. ASA Press titles can include all types of books that Springer publishes, and may appear in any appropriate Springer book series.

ASA
PRESS

Jon T. Sakata • Sarah C. Woolley
Richard R. Fay • Arthur N. Popper
Editors

The Neuroethology of Birdsong

ASA PRESS

Springer

Editors
Jon T. Sakata
Department of Biology and Centre for
Research in Brain, Language, and Music
McGill University
Montreal, QC, Canada

Sarah C. Woolley
Department of Biology and Centre for
Research in Brain, Language, and Music
McGill University
Montreal, QC, Canada

Richard R. Fay
Department of Psychology
Loyola University Chicago
Chicago, IL, USA

Arthur N. Popper
Department of Biology
University of Maryland
College Park, MD, USA

ISSN 0947-2657 ISSN 2197-1897 (electronic)
Springer Handbook of Auditory Research
ISBN 978-3-030-34682-9 ISBN 978-3-030-34683-6 (eBook)
https://doi.org/10.1007/978-3-030-34683-6

© Springer Nature Switzerland AG 2020
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

This volume is dedicated to Allison Doupe, an inspiration and mentor to scientists all over the world. She was instrumental in spurring a generation of neuroethologists to understand the neural basis of avian communication and to reveal parallels between human and avian communication.

The Acoustical Society of America

On 27 December 1928 a group of scientists and engineers met at Bell Telephone Laboratories in New York City to discuss organizing a society dedicated to the field of acoustics. Plans developed rapidly, and the Acoustical Society of America (ASA) held its first meeting on 10–11 May 1929 with a charter membership of about 450. Today, ASA has a worldwide membership of 7000.

The scope of this new society incorporated a broad range of technical areas that continues to be reflected in ASA's present-day endeavors. Today, ASA serves the interests of its members and the acoustics community in all branches of acoustics, both theoretical and applied. To achieve this goal, ASA has established technical committees charged with keeping abreast of the developments and needs of membership in specialized fields, as well as identifying new ones as they develop.

The Technical Committees include acoustical oceanography, animal bioacoustics, architectural acoustics, biomedical acoustics, engineering acoustics, musical acoustics, noise, physical acoustics, psychological and physiological acoustics, signal processing in acoustics, speech communication, structural acoustics and vibration, and underwater acoustics. This diversity is one of the Society's unique and strongest assets since it so strongly fosters and encourages cross-disciplinary learning, collaboration, and interactions.

ASA publications and meetings incorporate the diversity of these Technical Committees. In particular, publications play a major role in the Society. *The Journal of the Acoustical Society of America* (JASA) includes contributed papers and patent reviews. *JASA Express Letters* (JASA-EL) and *Proceedings of Meetings on Acoustics* (POMA) are online, open-access publications, offering rapid publication. *Acoustics Today*, published quarterly, is a popular open-access magazine. Other key features of ASA's publishing program include books, reprints of classic acoustics texts, and videos. ASA's biannual meetings offer opportunities for attendees to share information, with strong support throughout the career continuum, from students to retirees. Meetings incorporate many opportunities for professional and social interactions, and attendees find the personal contacts a rewarding experience. These experiences result in building a robust network of fellow scientists and engineers, many of whom become lifelong friends and colleagues.

From the Society's inception, members recognized the importance of developing acoustical standards with a focus on terminology, measurement procedures, and criteria for determining the effects of noise and vibration. The ASA Standards Program serves as the Secretariat for four American National Standards Institute Committees and provides administrative support for several international standards committees.

Throughout its history to present day, ASA's strength resides in attracting the interest and commitment of scholars devoted to promoting the knowledge and practical applications of acoustics. The unselfish activity of these individuals in the development of the Society is largely responsible for ASA's growth and present stature.

Series Preface

Springer Handbook of Auditory Research

The following preface is the one that we published back in 1992. As anyone reading the original preface, or the many users of the series, will note, we have far exceeded our original expectation of eight volumes. Indeed, with books published to date and those in the pipeline, we are now set for over 75 volumes in SHAR, and we are still open to new and exciting ideas for additional books.

We are very proud that there seems to be consensus, at least among our friends and colleagues, that SHAR has become an important and influential part of the auditory literature. While we have worked hard to develop and maintain the quality and value of SHAR, the real value of the books is very much because of the numerous authors who have given their time to write outstanding chapters and to our many coeditors who have provided the intellectual leadership to the individual volumes. We have worked with a remarkable and wonderful group of people, many of whom have become great personal friends of both of us. We also continue to work with a spectacular group of editors at Springer. Indeed, several of our past editors have moved on in the publishing world to become senior executives. To our delight, this includes the current president of Springer USA, Dr. William Curtis.

But the truth is that the series would and could not be possible without the support of our families, and we want to take this opportunity to dedicate all of the SHAR books, past and future, to them. Our wives, Catherine Fay and Helen Popper, and our children, Michelle Popper Levit, Melissa Popper Levinsohn, Christian Fay, and Amanda Fay Sierra, have been immensely patient as we developed and worked on this series. We thank them and state, without doubt, that this series could not have happened without them. We also dedicate the future of SHAR to our next generation of (potential) auditory researchers – our grandchildren – Ethan and Sophie Levinsohn; Emma Levit, Nathaniel, Evan, and Stella Fay; and Sebastian Sierra-Fay.

Preface 1992

The Springer Handbook of Auditory Research presents a series of comprehensive and synthetic reviews of the fundamental topics in modern auditory research. The volumes are aimed at all individuals with interests in hearing research including advanced graduate students, post-doctoral researchers, and clinical investigators. The volumes are intended to introduce new investigators to important aspects of hearing science and to help established investigators to better understand the fundamental theories and data in fields of hearing that they may not normally follow closely.

Each volume presents a particular topic comprehensively, and each serves as a synthetic overview and guide to the literature. As such, the chapters present neither exhaustive data reviews nor original research that has not yet appeared in peer-reviewed journals. The volumes focus on topics that have developed a solid data and conceptual foundation rather than on those for which a literature is only beginning to develop. New research areas will be covered on a timely basis in the series as they begin to mature.

Each volume in the series consists of a few substantial chapters on a particular topic. In some cases, the topics will be ones of traditional interest for which there is a substantial body of data and theory, such as auditory neuroanatomy (Vol. 1) and neurophysiology (Vol. 2). Other volumes in the series deal with topics that have begun to mature more recently, such as development, plasticity, and computational models of neural processing. In many cases, the series editors are joined by a co-editor having special expertise in the topic of the volume.

Richard R. Fay, Chicago, IL, USA
Arthur N. Popper, College Park, MD, USA

SHAR logo by Mark B. Weinberg, Potomac, Maryland, used with permission

Volume Preface

This volume considers vocal communication in songbirds at various levels of analysis. Songbirds have been central species in the study of vocal communication for decades, in part because of their elegant, melodic, and complex vocal signals, aesthetic beauty, and evolutionary diversity. Moreover, they are one of the few vertebrates that, like humans, must learn how to produce their vocal communication signals. In addition, the learning, production, and control of song are regulated by discrete, specialized neural circuits that resemble motor and sensorimotor circuits in other vertebrates, including humans. As a consequence, songbirds provide a unique and powerful model system to analyze neural mechanisms underlying vocal learning, production, and perception and to generate and then test mechanistic models of speech acquisition in humans.

Chapter 1, by Jon T. Sakata and Sarah C. Woolley, provides an overview of birdsong research and of the volume. In Chap. 2, Jon T. Sakata and Yoko Yazaki-Sugiyama summarize a circuit perspective on the neural mechanisms underlying song learning. This is followed by Chap. 3 in which Karagh Murphy, Koedi S. Lawley, Perry Smith, and Jonathan F. Prather provide an overview of the neural circuits that regulate song control, focusing on processes required to maintain accurate vocalizations in adulthood. Chapter 4, by Arthur Leblois and David J. Perkel, is a comprehensive examination of the neural circuitry within the avian basal ganglia nucleus Area X and the degree to which the function, organization, and molecular and cellular composition of Area X parallel that of mammalian basal ganglia circuitry. In Chap. 5, Sarah C. Woolley and Sarah M. N. Woolley review the organization of the auditory system in songbirds against the backdrop of evolutionary, social, and comparative neuroscience. Neuromodulation of neural activity in sensory and sensorimotor circuits affects various aspects of vocal communication, and in Chap. 6, Luke Remage-Healey provides a conceptual overview of the role of steroids in the modulation of auditory processing. In Chap. 7, Julie Elie and Frédéric Theunissen present an in-depth examination of the forebrain, midbrain, and hindbrain circuits that control production and perception of all songbird vocalizations. The organization and function of neural circuits is critically shaped by gene expression, and in Chap. 8, Sarah E. London reviews genomic and genetic contributions to vocal communication in songbirds. Finally, in Chap. 9, Jeffrey Podos and HaCheol Sung discuss the constraints on, and functional consequences of, vocal performance in the context of the evolution of birdsong.

Together, these chapters provide a synthesis of many facets of avian communication, including the learning, control, and perception of birdsong. We hope that this volume, like the person to whom this volume is dedicated, motivates current and future generations of neuroscientists, psychologists, and ethologists to explore this fascinating behavior.

Jon T. Sakata, Montreal, QC, Canada
Sarah C. Woolley, Montreal, QC, Canada
Richard R. Fay, Chicago, IL, USA
Arthur N. Popper, College Park, MD, USA

Contents

Contributors

Julie E. Elie Department of Bioengineering, University of California, Berkeley, CA, USA

Koedi S. Lawley Neuroscience Program, Department of Zoology and Physiology, University of Wyoming, Laramie, WY, USA

Arthur Leblois Institut des Maladies Neurodégénératives, CNRS, UMR 5293, Université de Bordeaux, Bordeaux, France

Sarah E. London Department of Psychology, Institute for Mind and Biology, Grossman Institute for Neuroscience, Quantitative Biology and Human Behavior, University of Chicago, Chicago, IL, USA

Karagh Murphy Neuroscience Program, Department of Zoology and Physiology, University of Wyoming, Laramie, WY, USA

David J. Perkel Departments of Biology and Otolaryngology, University of Washington, Seattle, WA, USA

Jeffrey Podos Department of Biology, University of Massachusetts, Amherst, MA, USA

Jonathan F. Prather Neuroscience Program, Department of Zoology and Physiology, University of Wyoming, Laramie, WY, USA

Luke Remage-Healey Department of Psychological and Brain Sciences, University of Massachusetts Amherst, Amherst, MA, USA

Jon T. Sakata Department of Biology and Centre for Research in Brain, Language, and Music, McGill University, Montreal, QC, Canada

Center for Studies in Behavioral Neurobiology, Concordia University, Montreal, QC, Canada

Perry Smith Neuroscience Program, Department of Zoology and Physiology, University of Wyoming, Laramie, WY, USA

Ha-Cheol Sung Department of Biological Sciences, Chonnam National University, Gwangju, South Korea

Frédéric E. Theunissen Departments of Psychology, Integrative Biology, and Helen Wills Neuroscience Institute, University of California, Berkeley, CA, USA

Sarah C. Woolley Department of Biology and Centre for Research in Brain, Language, and Music, McGill University, Montreal, QC, Canada

Sarah M. N. Woolley Zuckerman Institute and Department of Psychology, Columbia University, Jerome L. Greene Science Center, New York, NY, USA

Yoko Yazaki-Sugiyama Okinawa Institute of Science and Technology (OIST) Graduate University, Okinawa, Japan

International Research Center for Neurointelligence (IRCN), The University of Tokyo, Tokyo, Japan

Chapter 1
Scaling the Levels of Birdsong Analysis

Jon T. Sakata and Sarah C. Woolley

Abstract Over the past decades, research into the vocal communication system of songbirds has flourished. Research has spanned many levels of analysis, from broad evolutionary analyses of song production and acquisition to groundbreaking studies of neural contributions to song control and learning. Furthermore, there is increasing appreciation that song learning and control in songbirds share profound parallels with speech acquisition and control in humans, and that the neural circuits underlying birdsong and speech are highly analogous as well as homologous. For those reasons, songbirds have become the preeminent model system for generating and testing mechanistic models of speech acquisition and control. These aspects of and approaches to birdsong are the topics of this volume of the Springer Handbooks on Auditory Research (SHAR). This chapter serves to provide brief historical and conceptual backgrounds to birdsong research.

Keywords Communication · Comparative approach · Neural circuits · Neuroethology · Social behavior · Song learning · Songbird · Tinbergen · Vocal learning · Zebra finch

---

J. T. Sakata (✉)
Department of Biology and Centre for Research in Brain, Language, and Music, McGill University, Montreal, QC, Canada

Center for Studies in Behavioral Neurobiology, Concordia University, Montreal, QC, Canada
e-mail: jon.sakata@mcgill.ca

S. C. Woolley
Department of Biology and Centre for Research in Brain, Language, and Music, McGill University, Montreal, QC, Canada
e-mail: sarah.woolley@mcgill.ca

J. T. Sakata et al. (eds.), *The Neuroethology of Birdsong*, Springer Handbook of Auditory Research 71, https://doi.org/10.1007/978-3-030-34683-6_1

1.1 Introduction

In *The Study of Instinct*, Niko Tinbergen (1951) outlined an analysis framework that has deeply shaped and guided studies of animal behavior. He argued that four levels of analysis are critical for a comprehensive examination of behavior: adaptive significance (i.e., function), phylogeny (i.e., evolutionary history), ontogeny (i.e., development), and mechanism (i.e., physiology). Briefly, investigations into *adaptive significance* deal with how behaviors affect the survival and reproduction of the organism; studies of *phylogeny* situate behaviors into their evolutionary context by examining behaviors across species; analyses of the *ontogeny* of behavior aim to demonstrate how focal behaviors develop over the life of the organism and how experiences during development affect the display of the behavior; and research into *mechanisms* reveals the biological substrates that regulate the display and plasticity of the behavior.

Investigations of a broad range of behaviors (e.g., communicative, territorial, defensive, sexual, migratory, and predatory behaviors) have benefited tremendously from the conceptualization of this framework. While studies at each of these distinct levels can independently yield valuable insights, the integration of findings from different levels of analysis can generate novel interpretations, hypotheses, and synergies (Thierry 2005; MacDougall-Shackleton 2011).

Communication is one of the most extensively studied social behaviors across species. Each year, hundreds of papers dealing with the evolution, development, or control of animal communication are published. Organisms can utilize signals for communication that span all sensory modalities. One of the best studied of these systems is acoustic communication, which is used by a wide range of species from insects to humans (Greenberg et al. 2004; Suthers et al. 2016). Given the diversity of species that rely on acoustic signals for transmitting information and the importance of vocal communication for social interactions in humans, it is important to understand how and why various species acquire and control their acoustic signals (Bradbury and Vehrencamp 2011).

Songbirds use elaborate vocal signals in communication (*birdsong*), and they have become one of the dominant models for the study of acoustic communication as they provide rich opportunity for integration across levels of analysis (Bradbury and Vehrencamp 2011; Sakata and Vehrencamp 2012). The vocal signals (i.e., songs) of male songbirds are important for attracting females and defending territories and, therefore, are under sexual selection pressure to provide honest information about the singer's quality and intentions (Catchpole and Slater 2008; Nowicki and Searcy 2014). Females also produce a number of vocal signals that are used during social interactions, including song in a number of species (Odom et al. 2014; Riebel et al. 2019). In addition, laboratory studies of a handful of songbird species have provided extensive insight into the developmental factors that shape birdsong as well as the neurobiological underpinnings of birdsong (see Sect. 1.6).

Songbirds have been studied for decades, and most studies in the twentieth century were focused on evolutionary and ecological aspects of birdsong. For example,

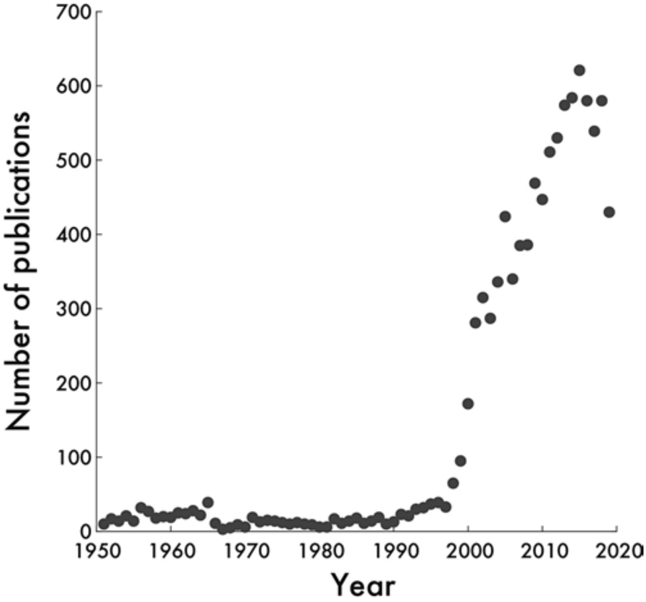


Fig. 1.1 The number of publications per year found using the search engine PubMed when searching for "birdsong" or "songbird" (1950–2018)

birdsong and other avian vocalizations (such as calls) have been central in the study of signal design (i.e., evolutionary arguments about the form and structure of acoustic signals), referential signaling, individual recognition, multimodal integration, information processing, and behavioral evolution (Catchpole and Slater 2008; Bradbury and Vehrencamp 2011). However, the discovery of neuroanatomically discrete and highly specialized neural circuits that are important for vocal learning, production, and control (e.g., Arnold et al. 1976; Nottebohm et al. 1976) led to an increase in studies on the development, plasticity, and physiological control of birdsong (see also Sect. 1.5). Since those ground-breaking discoveries, the prominence of songbirds (especially the zebra finch, *Taeniopygia guttata*) as models for understanding the neural and physiological bases of behavior has continued to grow. Indeed, the number of publications in biomedical, physiological, mechanistic, and related journals that include the word "birdsong" or "songbird" has increased dramatically since the beginning of the twenty-first century (Fig. 1.1). Related to this point, it is now not uncommon to see papers about the neural underpinnings of birdsong in broad and impactful journals.

1.2 Songbirds as Models for Human Speech

A central reason for the proliferation of mechanistic research on birdsong likely stems from the fact that songbirds are one of the few groups of species that, like humans, need to learn the structure of their vocal communication signals (*vocal production learning*). For most species that depend on acoustic signals for communication, the structure of their acoustic signals is *innate*, meaning that effective, species-typical communication signals can be produced by individuals regardless of their developmental history. Such species are referred to as *vocal nonlearners*. This is not to say that the usage or comprehension of acoustic signals does not involve a learning component in these species but just that the *production* of such signals does not require extensive learning (Seyfarth and Cheney 2010).

However, in a handful of species like humans and songbirds, individuals must learn how to produce acoustic communication signals through exposure to species-typical vocalizations (*vocal learners*). Speech and language development are shaped by exposure to particular sounds, and the accurate development and production of speech sounds require vocal practice and auditory feedback (i.e., *sensorimotor learning*). While humans can acquire new languages throughout their lives, an individual's ability to learn a new language or to accurately imitate accents diminishes with age, suggesting the existence of *sensitive periods* for speech and language acquisition (Friedmann and Rusou 2015; Werker and Hensch 2015). Because speech and language are central to the social life of humans and because deficits in speech and language acquisition can have profound emotional and psychological impacts (e.g., Beitchman et al. 2001; Conti-Ramsden et al. 2013), studying neurobiological processes in animals that learn their communication signals in a manner similar to humans provides a powerful and important opportunity to generate and experimentally test models of speech acquisition and production.

Like humans, songbirds require particular developmental experiences to acquire their communication signals. As early as the eighteenth century, publications provided evidence that certain species of birds learn their vocalizations during development. For example, Barrington (1773) noted how cross-fostering songbirds with a different species or raising songbirds in acoustic isolation led to the production of songs that deviated from species-typical song. In addition, William Thorpe's (1958) classic study of song dialects in chaffinches (*Fringilla coelebs*) and Peter Marler's (1970, 1997) seminal studies of sparrows paved the way for the contemporary era of birdsong research (for a more comprehensive historical perspective, see Roberts and Mooney 2013). These and other studies highlighted that songs are learned during restricted periods in development, that juvenile songbirds engage in a protracted period of vocal practice and development (*babbling*), and that auditory feedback is critical for the accurate development of song. Since those groundbreaking publications, many researchers have provided important insights into various aspects of song acquisition, including the factors that influence the fidelity of song learning, the diversity of learning strategies, and the neural mechanisms that underlying song learning.

Because of the numerous and deep behavioral parallels of communication across songbirds and humans, birdsong has become an important model system to understand mechanisms underlying speech acquisition. Speech refers to the articulatory, respiratory, and prosodic components of language, whereas language itself broadly refers to the semantic aspects of human communication that deals with the communication of complex conceptual issues (Fitch et al. 2010; Helekar 2013). Songbird researchers are careful to underscore the parallels of birdsong with speech and not with language; this is because no research to date has demonstrated that birdsong possesses the breadth and complexity of the semantics of language. This is certainly not meant to indicate that birdsong is not rich with communicative *meaning*, rather that the scope of meaning in birdsong is highly unlikely to parallel that of human language (Berwick et al. 2011). Conversely, both birdsong and speech require the coordination of muscular and respiratory mechanisms to generate vocal communication signals; as such, understanding how songbirds learn to produce their song is analogous to understanding how human infants learn to produce speech sounds (Doupe and Kuhl 1999).

Another reason why songbirds are considered powerful models to reveal mechanisms of speech is that songbirds possess discrete and specialized neural circuitry for song that are analogous and, for some brain regions, homologous to the neural circuits for speech in humans (Arriaga et al. 2012; Pfenning et al. 2014). By and large, these brain areas have not been found in vocal nonlearning birds (but see Liu et al. 2013). Furthermore, these circuits are accessible for the monitoring and manipulation of neural activity associated with the learning and control of vocalizations. Consequently, songbirds provide a powerful system to examine the function of brain circuitry, especially in the context of vocal communication.

Broadly speaking, birdsong is regulated by two canonical brain pathways (Fig. 1.2). The *vocal motor pathway* (VMP) connects forebrain areas such as HVC (used as a proper name) and the robust nucleus of the arcopallium (RA) with hind-

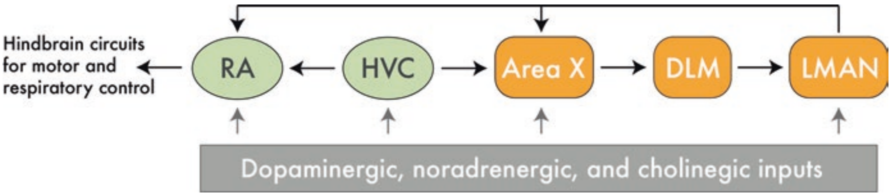


Fig. 1.2 A schematic outlining the canonical song system. The song system consists of the vocal motor pathway (VMP; *green ovals*) and the anterior forebrain pathway (AFP; *orange rectangles*). The VMP consists of forebrain areas like *HVC* (used as a proper name) and the robust nucleus of the arcopallium (*RA*) and is critical for song production. The AFP is a basal ganglia-thalamic-cortical circuit that is homologous to that found in mammals and is important for song plasticity and control. The AFP consists of *Area X* (the vocal portion of the avian basal ganglia), the medial portion of the dorsolateral thalamic nucleus (*DLM*), and the lateral magnocellular nucleus of the anterior nidopallium (*LMAN*). Most nuclei in the canonical song system receive neuromodulatory inputs from dopaminergic, noradrenergic, or cholinergic neurons (*grey rectangle*)

brain areas that contain vocal motor and respiratory neurons (see Elie and Theunissen, Chap. 7 for a detailed examination of circuitry). The nucleus HVC is functionally analogous to the premotor cortex, while RA is analogous to the vocal motor region of the primary motor cortex in mammals (Fujimoto et al. 2011; Hara et al. 2012). The *anterior forebrain pathway* (AFP) is a basal ganglia-thalamic-cortical circuit that includes the basal ganglia nucleus Area X, the medial portion of the dorsolateral thalamic nucleus (DLM), and the lateral magnocellular nucleus of the anterior nidopallium (LMAN). Area X is a specialized region of the avian basal ganglia, a homologue of the mammalian basal ganglia (Doupe et al. 2005; see Leblois and Perkel, Chap. 4), and LMAN is functionally analogous to parts of the mammalian frontal cortex (Reiner et al. 2004; Pfenning et al. 2014).

Neurons throughout the canonical song circuitry are sensorimotor in nature, displaying both auditory responses to song and motor activity during singing. Auditory information enters these two pathways through input from ascending auditory pathways into the nucleus interfacialis of the nidopallium (NIf) and HVC, and it is generally thought that HVC provides auditory information to the rest of the song system (e.g., RA, Area X, and LMAN). These nuclei historically have been the focus of research aimed at understanding the neural substrates of song learning and production, although there is growing interest in neuromodulatory circuits that impinge on these canonical circuits and in forebrain circuits that have parallel organization to the canonical song circuit (Bottjer and Altenau 2010; see Sakata and Yazaki-Sugiyama, Chap. 2).

Parrots and hummingbirds are also vocal learners, and similar circuitry for song learning, control, and processing is found in these species. For example, both hummingbirds and parrots have a caudal vocal motor pathway (like the VMP) that consists of sensorimotor neurons that are critical for song production (e.g., Jarvis and Mello 2000; Jarvis et al. 2000) and a rostral pathway that resembles basal ganglia-thalamic-cortical loops in mammals (Hara et al. 2012; Pfenning et al. 2014). While a handful of studies have described the functional neuroanatomy of song in these species, further examinations are important to elucidate conserved aspects of the neural control of song and the evolution of vocal learning (e.g., Feenders et al. 2008; Hara et al. 2012).

1.3 Defining Song Learning

As indicated in Sect. 1.2, one of the exceptional qualities of songbirds is that they learn some of the vocalizations used for social communication. This section will provide a more detailed examination of song learning by rigorously defining song learning and by presenting a more detailed overview of vocal learning, including a discussion of vocal learning processes in species other than songbirds.

Broadly speaking, vocal learners are species that are not born with the ability to produce their species-typical vocalizations but instead must learn how to produce these communicative sounds. In most cases, individuals acquire the ability to produce their species-typical vocal signals by interacting with and imitating the vocalizations of other individuals. Correspondingly, individuals that are not exposed to relevant vocalizations or not allowed to interact with conspecifics do not develop species-typical vocalizations. For example, chaffinches (*Fringilla coelebs*) that are deprived of exposure to adult conspecific songs throughout development produce aberrant, species-atypical songs as adults (Thorpe 1958). Indeed, developmental manipulations of auditory and social experiences have been central in determining whether a species should be considered as a vocal learner. Such controlled manipulations of developmental experiences are difficult to conduct across a wide range of species, and other types of vocal changes have been used as evidence for vocal learning in natural populations of animals (see Sect. 1.3.4).

1.3.1 Requirements of Song Learning: Sensory and Sensorimotor Learning

Fundamentally, song learning involves acquiring a sensory representation of a *target song* (*sensory learning*) and learning how to generate the motor commands to produce that target song (*sensorimotor learning*) (Mooney 2009; Brainard and Doupe 2013). Sensory learning is central to sensorimotor learning because sensorimotor learning involves comparing the current performance of the target song to the sensory representation of the target song (Doya and Sejnowski 1995; Fee and Goldberg 2011). Moreover, sensorimotor learning is critically dependent on auditory feedback to make the comparison between the song that was performed and the target song (see Sakata and Yazaki-Sugiyama, Chap. 2; Murphy, Lawley, Smith, and Prather, Chap. 3).

1.3.2 Oscine Versus Suboscines

By a number of estimates, there are approximately 10,000 species of birds. Birds are broadly segregated into passerine (perching birds) and nonpasserine birds (e.g., chickens, ducks, vultures, eagles, ostriches). Passerines are further classified into oscines and suboscines. Oscines are what we typically refer to when discussing songbirds (e.g., sparrows, robins, finches, corvids, cowbirds, lyrebirds), whereas suboscines refer to passerines that, generally speaking, do not engage in vocal learning (e.g., manakins, flycatchers, antbirds, pittas).

While this nomenclature implies categorical distinctions between oscines (vocal learners) and suboscines (vocal nonlearners), the variation in vocal learning and plasticity may be more continuous. Petkov and Jarvis (2012) provide a detailed overview of vocal learning in vertebrates and emphasize that there is no binary distinction between vocal learners and vocal nonlearners (also see Arriaga and Jarvis

2013). The authors underscore variation in the extent and nature of vocal learning and plasticity across birds and other vertebrates. They emphasize that some suboscines demonstrate rudimentary forms of vocal learning, whereas some oscines demonstrate limited forms of learning (i.e., there is a large genetic component to their adult vocalizations). Moreover, their review highlights the dearth of species in which vocal learning has been experimentally tested, and they urge more expansive experimental testing. Because complex traits like song learning can be lost at evolutionary transitions within the oscine clade, one should not necessarily assume that all oscines are vocal learners.

1.3.3 Types of Vocal Plasticity

Vocal learning typically refers to the developmental process of learning how to produce and imitate vocal signals but is only one form of vocal plasticity. In fact, other forms of vocal plasticity exist that relate in many ways, phenotypically and mechanistically, to developmental vocal learning. Two that have been particularly well-studied are plasticity in response to changes in environmental conditions and plasticity in response to more specific and targeted experimental manipulations. Determining the similarities between forms of plasticity can provide critical insight into their underlying mechanisms, especially for species in which it is difficult to manipulate developmental experiences experimentally.

A number of laboratory and field studies have investigated how birds (and other vertebrates) adaptively change their vocalizations in response to environmental conditions (Kight and Swaddle 2011). For example, researchers have revealed that birds and other vertebrates living in environments with different types and levels of background noise—both anthropogenic (urban) and natural—produce vocalizations with different acoustic structures (Slabbekoorn and Ripmeester 2008; Tyack and Janik 2013). Such changes can be driven by selection (i.e., only individuals whose signals are salient and effective under particular conditions successfully reproduce) or by plasticity (i.e., individuals change their vocalizations in response to noise such that the vocalizations remain salient and effective). The latter phenomenon is of particular interest because the neural mechanisms regulating vocal plasticity in response to changing environmental conditions could be similar to those mediating developmental song learning (see Sakata and Yazaki-Sugiyama, Chap. 2).

In addition to broad changes to song induced by exposure to environmental noise, songbirds show rapid changes to discrete acoustic features in more targeted experimental paradigms. In particular, computer software to accurately, consistently, and rapidly target acoustic features of an individual's song have been used to experimentally alter the acoustic structure, sequencing, and timing of vocal elements in songbirds (reviewed in Sakata and Yazaki-Sugiyama, Chap. 2) and to drive similar changes to vocal patterns in other species (Egnor et al. 2007; Roy et al. 2011). For example, computerized systems have been used to deliver stimuli (e.g., a burst of white noise) when adult songbirds produce a targeted *syllable* (acoustically distinct vocal element) at a fundamental frequency (FF) below a certain threshold. In response to this reinforcement paradigm, birds will adaptively shift the FF of the targeted syllable up to avoid white noise playbacks (e.g., Tumer and Brainard 2007; Andalman and Fee 2009). Such contingent feedback paradigms are thought to effectively drive adaptive song plasticity because birds monitor their song performance in real time (Cynx and von Rad 2001; Sakata and Brainard 2006) and because super-positioning white noise on top of the sound of the bird's own song interferes with the processing of auditory feedback. In addition, playbacks of loud bursts of white noise are also likely to be aversive to birds, leading to behavioral adaptations that minimize exposure to aversive stimuli.

Researchers have coupled these experimental paradigms in adult songbirds with neurophysiological recordings (e.g., Sakata and Brainard 2008; Keller and Hahnloser 2009) or manipulations of neural circuitry (e.g., Andalman and Fee 2009; Mandelblat-Cerf et al. 2014) to reveal the mechanisms underlying feedback processing and vocal plasticity. Such studies have highlighted the role of specialized brain areas in songbirds (see Sect. 1.2 in this chapter and other chapters in this volume), including LMAN (Andalman and Fee 2009; Charlesworth et al. 2012) and HVC (Roberts et al. 2017), in the adaptive shifts in song structure. Research has also demonstrated that brain areas outside the canonical neural circuitry for birdsong, including the ventral tegmental area (VTA) (Gadagkar et al. 2016), are critical for adaptive song plasticity (Hisey et al. 2018; Xiao et al. 2018). Such studies are also important for the evaluation of the degree to which the neural mechanisms regulating developmental song learning overlap with mechanisms underlying song plasticity in response to contingent feedback in adult songbirds. For example, the sensorimotor nucleus LMAN is critical for the developmental plasticity of learned song and for the adaptive modifications of song in response to reinforcement paradigms (reviewed in Brainard and Doupe 2013). Some recent experiments have explicitly assessed how the same neural circuitry mediates both types of plasticity (e.g., dopaminergic neurons in the VTA: Hisey et al. 2018).

The comparative approach can also be used to provide insight into mechanisms underlying various types of vocal plasticity. Most laboratory experiments on vocal plasticity have been conducted on vocal learners, namely songbirds such as the zebra finch and Bengalese finch (*Lonchura striata* var. *domestica*). In contrast, variation in song in response to ecological noise has been studied in both vocal learners and vocal nonlearners. Of particular interest are findings that some species of vocal nonlearners (suboscines) demonstrate vocal plasticity in response to environmental conditions (Francis et al. 2011; Gentry et al. 2018; but see Ríos-Chelén et al. 2018). Given that suboscines have not evolved the forebrain circuitry for song learning, such vocal changes are likely mediated by neural circuits outside the specialized forebrain circuit for song (e.g., the VTA). In addition, these findings in vocal nonlearners suggest the existence of a rudimentary and evolutionarily conserved neural circuitry for vocal plasticity, and the identification of such circuits could help neuroethologists gain insight into the evolution of song learning (see London, Chap. 8). A related approach to reveal mechanisms underlying vocal plasticity could be to examine and compare how non-songbirds and songbirds adaptively alter the acoustic structure of unlearned vocalizations (e.g., innate calls); such studies could provide important insights into the evolution of mechanisms for vocal plasticity (see Elie and Theunissen, Chap. 7).

1.3.4 Evidence for and Against Vocal Learning in Other Species

Vocal learning has been observed in only a handful of vertebrate groups: humans, three clades of birds (songbirds, hummingbirds, and parrots), and various species of bats, cetaceans, pinnipeds, and elephants (reviewed in Petkov and Jarvis 2012; Janik 2014). Much of the work on mammals has focused on the ability for vocal mimicry and vocal modifications in response to the social environment, rather than the influence of developmental experiences, as support for vocal learning. For example, a number of bat species change the acoustic structure of specific types of calls (e.g., screech or isolation calls), depending on their social group (reviewed in Knörnschild 2014); some dolphin species change their vocal signature whistle as a function of group composition (reviewed in Janik 2014); and some elephants can imitate speech sounds of humans (reviewed in Stoeger and Manger 2014). Aside from vocal-learning birds, bats offer significant promise in revealing mechanisms of vocal learning, and laboratories have begun to manipulate developmental experiences to unravel the nature of developmental vocal learning and plasticity (e.g., Prat et al. 2015, 2017). Coupling experimental paradigms of vocal plasticity with neurophysiological recordings, gene expression assays, and neural manipulations will lead to important insights into mechanisms of mammalian vocalizations and allow for a more comparative investigation into conserved neural circuitry for vocal learning (Hage and Nieder 2016; Lattenkamp and Vernes 2018).

In contrast to the evidence for vocal learning and plasticity in these groups of vertebrates, experimental evidence for vocal learning and plasticity in other vertebrates, such as rodents and nonhuman primates, has been quite limited historically. After the discovery that mice produce sequences of vocalizations that are individual-specific (Holy and Guo 2005), a number of investigators sought to document the extent to which the structure of rodent vocalizations was learned. Despite the number and diversity of studies aimed at revealing a learning component to rodent vocalizations, very few studies have documented a significant role for auditory and social experiences (or of auditory feedback) on vocal development in rodents (reviewed in Portfors and Perkel 2014; but see Arriaga and Jarvis 2013). For example, in contrast to vocal learners that critically depend on auditory feedback for vocal development, congenitally deaf mice produce vocalizations that are indistinguishable from conspecifics with intact hearing. In addition, while there exist strain differences in rodent vocalizations, cross-fostering pups of one strain with adults from a different (foster) strain does not lead to pups learning to produce the vocalizations of the foster strain (e.g., Kikusui et al. 2011).

Manipulations of developmental experiences are hallmarks of experiments of vocal learning, and there have been an increasing number of controlled studies documenting a learning component to nonhuman primate vocalizations. For example, several studies have found a significant role of parental and social reinforcement in vocal development in common marmosets *(Callithrix jacchus*) (e.g., Takahashi et al. 2015; Gultekin and Hage 2017). Of particular interest is the finding that experimental social reinforcement can promote the developmental maturation of phee calls (Takahashi et al. 2017). In this study, researchers played back a parental call to serve as a positive reinforcement signal, either 100% or 10% of the time, when developing marmosets produced a mature version of a phee call. Marmosets that received contingent parental reinforcement on 100% of the trials in which they produced a mature phee call developed the call more rapidly than marmosets that received contingent reinforcement on only 10% of the trials. From a mechanistic perspective, these studies of vocalizations provide a basis for examinations into the neural circuits that could mediate vocal plasticity in primates (e.g., VTA). Together with related experimental studies on vocal development in nonhuman primates, this study provides an important complement to studies describing vocal variation in response to social and environmental changes in natural populations of nonhuman primates (e.g., Egnor and Hauser 2004; Watson et al. 2015).

1.3.5 Relationship to Visual Imprinting

Many of the early pioneers of ethological and neuroethological approaches to birdsong outlined parallels between the sensory learning of song and filial and sexual imprinting in birds (Horn 2004; Bolhuis and Gahr 2006). Indeed, both forms of imprinting and song learning (at least in close-ended learners) are restricted to critical periods in development (see Sect. 1.3.2), are influenced by innate biases and constraints, are shaped by activity in rostral forebrain areas, and are accompanied by changes in neural morphology, transmission, and responsiveness. For example, NMDA receptor expression in rostral forebrain areas is important for both imprinting and song learning, and lesions within these rostral forebrain areas prevented developing birds from remembering visual stimuli (in the case of imprinting) or imitating learned sounds (in the case of song learning) (Horn 2004; Bolhuis and Gahr 2006).

Because filial and sexual imprinting are more widespread across avian taxa, this form of plasticity could represent a precursor to song learning. While the primary sensory modalities involved in imprinting (visual) and song learning (auditory) differ, both forms of sensory learning may rely on similar neural substrates. One possibility is that areas in the midbrain and hindbrain, such as the VTA and locus coeruleus (LC) that send neuromodulatory projections throughout sensory processing areas in the brain, could be important for the long-term memorization of sensory inputs during development (see Sakata and Yazaki-Sugiyama, Chap. 2;

Remage-Healey, Chap. 6). While there has been an unfortunate decline in the amount of research into mechanisms of imprinting, a promising line of inquiry would be to use modern and powerful methodologies (e.g., multi-channel neurophysiological recordings, optogenetics, fiber photometry, etc.) to assess the degree to which song learning and imprinting recruit similar or disparate mechanisms. Given that early visual experiences can have lasting effects on the processing of visual signals in humans (e.g., Anzures et al. 2013; Simion and Giorgio 2015), further understanding of the mechanisms of visual imprinting in birds could have translational impacts on our understanding of human social behavior.

1.4 Diversity of Song-Learning Strategies

Given that there are more than 5000 species of songbirds, it is not surprising that there is diversity in song-learning strategies across these species (Beecher and Brenowitz 2005; Clayton et al. 2009). In their review, Beecher and Brenowitz (2005) extensively describe the various ways in which songbirds learn song throughout their lives (see also Catchpole and Slater 2008; Bradbury and Vehrencamp 2011) and highlight the importance of adopting a comparative perspective (see Petkov and Jarvis 2012; Murphy et al. 2017). They emphasize that song-learning strategies should be viewed in light of phylogenetic history because species could possess similar learning strategies and similar neural mechanisms for song learning due to similar adaptive pressures (convergence) or shared ancestry.

Variation in song-learning strategies can be broadly categorized as variation in the timing of song learning, variation in the extent of song learning, and variation in the dependence on social interactions for song learning. The following sections describe these types of variation in song learning.

1.4.1 Variation in the Timing of Song Learning

Broadly speaking, evolutionary variation in the timing of song learning encompasses variations in the duration of time in which birds can learn new songs (i.e., the existence, duration, and timing of critical periods for song learning) as well as variations in the relative timing of sensory and sensorimotor periods of song learning.

1.4.1.1 Open-Ended Versus Close-Ended Learning and Critical Periods for Song Learning

The most commonly studied songbirds are *close-ended learners*: species in which song learning is restricted to a critical period in development. Because learning is restricted to a critical period, quantifying and manipulating vocal learning can be relatively easy in such species. However, a number of songbirds, like nightingales (*Luscinia megarhynchos*) and European starlings (*Sturnus vulgaris*), as well as another clade of vocal learners, parrots (Pepperberg 2010), are *open-ended learners*. Like humans, these species are able to acquire new vocalizations throughout their lives.

Understanding the neurobiological mechanisms underlying open-ended versus close-ended learning has been a central question in birdsong research (Murphy et al. 2017; Yazaki-Sugiyama 2019). Such lines of inquiry are important not just from a neuroethological perspective but also for translational reasons. While humans can acquire new communicative sounds throughout their lives, this capacity to acquire new languages and speech styles degrades over time, and discovering how to extend or re-open periods for vocal learning would enhance language learning in adults (Sakata and Yazaki-Sugiyama, Chap. 2; London, Chap. 8). Evolutionary variations in the expression of *molecular brakes* (molecules that restrict plasticity; Takesian and Hensch 2013) covary with some evolutionary variations in adult song learning (Cornez et al. 2017). While this line of inquiry is promising, additional studies are required. Nevertheless, the ability to compare open-ended and close-ended learners offers a powerful opportunity to reveal neurobiological factors that enable and constrain plasticity.

Among close-ended learners, the developmental periods that shape song output can vary dramatically from species to species. The crux of song learning occurs within 1–3 months after hatching in the most studied songbirds, including zebra finches and Bengalese finches. However, song learning primarily occurs long after hatching for some species. For example, interspecific brood parasites (i.e., birds that lay their eggs in the nests of another host species), such as the brown-headed cowbird (*Molothrus ater*), seem to acquire much of their song months after they are born, during their first winter. Conversely, recent studies have documented that, for some species, sensory experiences *in ovo* can affect the production and perception of conspecific signals. For instance, pre-hatching exposure to adult vocalizations affects song learning in superb fairywrens (*Malurus cyaneus*), a species that produces precocial offspring (i.e., hatch in a relatively mature state) (Colombelli-Négrel et al. 2012; Colombelli-Négrel Diane et al. 2014). Further studies of precocial birds that undergo much of their development *in ovo* could reveal further effects of experiences before hatching.

Interestingly, dramatic variation in the timing of song learning can be found among individuals within a species, depending on when individuals are born. Song learning occurs in the summer for marsh wrens (*Cistothorus palustris*) that hatch early in the spring but occurs in the following spring for individuals that hatch late in the breeding season (Kroodsma and Pickert 1980). This seasonal effect is mediated by photoperiod or auditory experience and is evidence of plasticity in the timing of critical periods for song learning (see Sakata and Yazaki-Sugiyama, Chap. 2; London, Chap. 8).

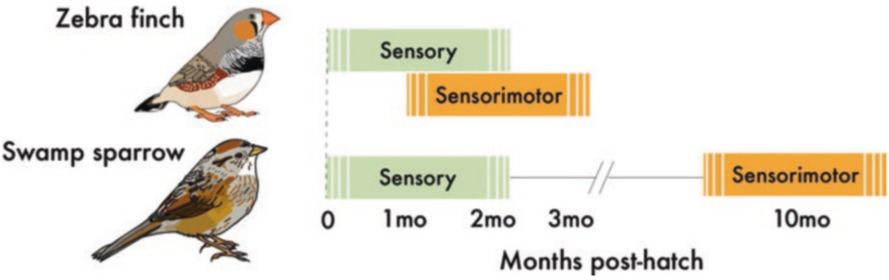


Fig. 1.3 Species variation in the timing of critical periods for song learning. Song learning consists of a *sensory period* (during which the sensory representation of song is learned) and a *sensorimotor period* (during which the motor commands to generate the song are learned). In species like the zebra finch (*Taeniopygia guttata)*, the sensory and sensorimotor periods overlap. In contrast, in species like the swamp sparrow (*Melospiza georgiana*), the sensory and sensorimotor periods are separated by a number of months

1.4.1.2 Overlap of Sensory and Sensorimotor Periods

As indicated in Sect. 1.3.1, song learning generally entails the acquisition of the sensory representation of the song(s) to imitate (*song template*) as well as learning how to generate the motor commands to produce vocalizations that match the song template (see Murphy, Lawley, Smith, and Prather, Chap. 3). The sensory and sensorimotor learning of song can occur at separate or overlapping periods in development (Fig. 1.3) (Brainard and Doupe 2002). For example, in zebra finches and Bengalese finches, the latter part of the sensory period overlaps with the initial phases of the sensorimotor period. In contrast, the sensorimotor period of song learning can occur months after the end of the sensory period for song learning in emberizine sparrows (Marler and Peters 1981; Brainard and Doupe 2002). For example, song practice (i.e., sensorimotor learning) can begin up to eight months after swamp sparrows (*Melospiza georgiana*) have learned the sounds of the songs to imitate (Fig. 1.3) (Marler and Peters 1981).

1.4.1.3 Variation in Reliance on Sensory Learning

Despite the general rule that song learning requires both sensory and sensorimotor learning, vocal development and learning in some songbird species can deviate from that framework. For example, in brood parasites such as brown-headed cowbirds, song learning could proceed without sensory learning. In this species, juveniles produce a range of variable and unstructured vocalizations during their first over-wintering period, and adult females perform social displays in response to hearing elements in a juvenile's song that they prefer (West et al. 1981). Juvenile cowbirds attend to these female social displays and selectively retain the vocalizations that elicit more social displays from females (West and King 1988). While little is known about the extent and nature of sensory learning in such species (but see Louder et al. 2019), social feedback in response to variable vocalizations can be sufficient to develop species-typical songs.

1.4.2 Variation in the Extent of Song Learning

Evolutionary variation in the extent of song learning can be broadly categorized as variation in the number of sounds that birds learn to produce (i.e., size of the vocal repertoire) and the range of sounds that can be incorporated into a song repertoire.

1.4.2.1 Variation in Song Repertoires

Birds have been classified by the number of song types they can produce as adults. Songbirds like the zebra finch learn to produce only one stereotyped sequence of vocal elements (i.e., one song) that they sing throughout their lives, whereas songbirds such as wood thrushes (*Hylocichla mustelina*) and superb lyrebirds (*Menura novaehollandiae*) learn to produce hundreds of song types. The neural mechanisms that allow for learning large repertoires have been of interest both from an evolutionary perspective (Byers and Kroodsma 2009; Soma and Garamszegi 2011) and from a translational perspective (i.e., as a way to enhance vocal learning).

Complementary to the song repertoire is the notion of the *syllable repertoire*. For example, two individuals from different songbird species or two individuals of the same species could each learn to produce only one song type, but the individuals could vary in the number of syllables in that song type. Species variations in syllable repertoire size and the neural correlates of these variations were extensively examined in the 1990s and 2000s (Devoogd 2004; Garamszegi and Eens 2004), though fewer studies have been devoted to this comparative analysis since then (e.g., Moore et al. 2011).

1.4.2.2 Variation in the Breadth of Acoustic Elements that Can be Learned

Generally speaking, songbirds are most adept at learning the sounds that are typical of their own species. In fact, a number of studies have demonstrated that relatively little exposure to conspecific vocalizations is required for juvenile songbirds to learn conspecific song (Catchpole and Slater 2008; Deshpande et al. 2014). For example, juvenile zebra finches demonstrated significant learning following a single day of housing with an adult tutor (Roberts et al. 2010; Chen et al. 2016), and song sparrows (*Melospiza melodia*) can learn song from as little as thirty exposures to a song type during development (Peters et al. 1992). Relatedly, when songbirds are tutored with conspecific and heterospecific songs, they produce more accurate imitations of conspecific vocalizations (Woolley 2012; see Sakata and Yazaki-Sugiyama, Chap. 2).

Although there is striking similarity across songbird species in their impressive abilities and strong preferences to learn conspecific song (Moore and Woolley 2019), there exists significant variation in the degree to which individual songbird species can learn to produce the songs of heterospecifics. Some songbird species are rigid and limited in their ability to learn heterospecific song, while other species can readily learn to produce heterospecific song. For those able to learn heterospecific song, the degree and quality of heterospecific song copying can vary, depending on the acoustic similarities between conspecific and heterospecific song as well as other factors (e.g., social interactions). For example, some species can only copy elements of acoustically similar songs; others, such as superb lyrebirds and other vocal mimics, can not only produce the songs of other bird species but also imitate anthropogenic sounds (Dalziell et al. 2015; Goller and Shizuka 2018). How central and peripheral factors interact to regulate the breadth of sounds that individual species can imitate is largely unknown.

An influential study by Marler and Peters (1988) highlighted the need to consider the temporal organization of song in the acquisition of conspecific and heterospecific syllables. They found that the extent to which juvenile song sparrows learned the syllables of a related species, the swamp sparrow, depended on how the heterospecific syllables were temporally arranged. Song sparrows and swamp sparrows produce songs that consist of acoustically distinct syllables that also are temporally arranged in distinct ways. Swamp sparrows perform trilled songs wherein the song consists of a single segment in which a single syllable is repeated; this contrasts with song sparrow song, which consists of multiple segments of distinct syllables. Juvenile song sparrows do not readily learn the swamp sparrow syllables when the syllables are played back as a trill with a single segment, but they imitated the acoustic structure of swamp sparrow syllables to a greater degree if the swamp sparrow syllables were organized and sequenced into multiple segments (i.e., organized into a pattern typical of song sparrows). Therefore, as illustrated by those studies, a number of experimental approaches are required to fully reveal the extent to which songbirds can learn heterospecific song.

1.4.3 Variation in the Dependence of Song Learning on Social Interactions

Across species, variations in song learning strategies also depend on the extent to which social interactions are important for song learning (Snowdon and Hauserberger 1997; Soma 2011). Beecher and Brenowitz (2005) briefly consider the importance of social factors, alluding to findings that some species (but not others) were unable to learn from passive playbacks of song. This suggested that there is species variation in the importance of social interactions for song learning.

Since the Beecher and Brenowitz (2005) review, a number of studies have further highlighted the importance of social factors for song learning and control (Beecher 2017; see Sakata and Yazaki-Sugiyama, Chap. 2). For example, social interactions shape the fidelity of song learning in a number of avian species, including zebra finches (e.g., Eales 1989; Chen et al. 2016), white-crowned sparrows (e.g., Baptista and Petrinovich 1986), European starlings (Chaiken et al. 1993), and parrots (Farabaugh et al. 1994). Overall, juvenile birds that are allowed to visually and acoustically interact with adult tutors produce more accurate imitations of their tutor's song. Relatedly, the degree to which birds incorporate song elements from a heterospecific tutor depends on whether the bird is socially or nonsocially tutored.

Revealing species variation in the importance of social interactions for song learning and elucidating the neural circuits that allow for social interactions to gate vocal learning are important not only for understanding birdsong but also for understanding speech and language. Deficits in speech and language acquisition tend to be associated with dysfunctions in social behavior (Kuhl 2007; Eigsti et al. 2011); hence, understanding the neural circuits that process social information in order to influence vocal learning could provide insight into the etiology of communicative disorders (Kuhl 2010).

Social interactions also affect the production of learned song in another way. Territorial species, such as white-crowned sparrows and song sparrows, learn a number of song types. However, which song types they eventually produce depends on the songs of their neighbors. In particular, males will select song types from their own repertoire that best match their neighbors' song types (reviewed in Marler 1997). The mechanisms underlying this form of vocal plasticity remains unknown but could be mechanistically related to how social-feedback signals from adult female cowbirds shape the songs of developing cowbirds.

1.5 Allison Doupe as a Seminal Figure in the Study of the Neural Basis of Birdsong

Birdsong has become established not only as a powerful model system to reveal the neural circuits underlying evolutionarily important behaviors but also as a translational model system to reveal potential mechanisms underlying the acquisition and control of speech. While there are a number of findings and publications that have contributed to this appreciation of songbirds (e.g., the discovery of neurogenesis in adult songbirds, the development of cutting-edge neuroscience technologies), the use of songbirds for translational studies were, in our opinion, strongly influenced by an impactful review paper by Allison Doupe and Patricia Kuhl titled "Birdsong and Speech: Common Themes and Mechanisms" (Doupe and Kuhl 1999). In this publication, which is among the most cited papers in birdsong (over 1500 citations at the time of this publication), the authors outline the many parallels between birdsong and speech acquisition, including the role of auditory feedback, innate predispositions in learning, the effect of vocal learning on perception, sensitive periods in vocal learning, and functional neural circuitry. These topics continue to be the pillars of research programs in laboratories studying birdsong and speech around the world; indeed, these topics permeate through many of the chapters in this book.

After reading the review by Doupe and Kuhl, aspiring researchers (like ourselves) gained a profound appreciation for the impressive similarities between birdsong and speech acquisition. Moreover, that review dramatically increased the visibility of birdsong to researchers outside of neuroethology and behavioral ecology and inspired generations of neuroscientists, clinicians, and cellular and molecular biologists to view songbirds as translational models. While correlative, a figure depicting the number of "birdsong" and "songbird" publications (listed on PubMed) over time underscores the potential importance of the Doupe and Kuhl review (Fig. 1.1). From 1950 to 2000, the number of publications including the word "birdsong" or "songbird" was relatively low and steady, with a median of about 14 publications per year. However, the number of such publications skyrocketed after the year 2000, with a median of about 447 publications per year from 2000 to 2018. Regardless of whether the review by Doupe and Kuhl sparked the rise in songbird research, it has had undeniable impact on the field.

Beyond the Doupe and Kuhl (1999) review, Allison Doupe was a pioneer who helped uncover the importance of the avian forebrain-basal ganglia circuit in song learning (e.g., Brainard and Doupe 2000; Solis and Doupe 2000). In addition, her lab published impactful papers on the role of social context in modulating neural activity and behavior (e.g., Hessler and Doupe 1999; Kao et al. 2008), the influence of learning and maturation on the "tuning" of sensorimotor neurons (e.g., Solis and Doupe 1999; Solis and Doupe 2000), the role of specific neural circuits on the practice and performance of song (e.g., Kao et al. 2005; Woolley et al. 2014), and the complex tuning properties of auditory neurons (e.g., Theunissen et al. 2000; Nagel and Doupe 2006). Because of the profound impact she has had on the field and on the editors themselves, this SHAR volume is dedicated to Allison's memory.

1.6 Overview of Chapters and Conclusion

As highlighted in Fig. 1.1, research on communication in songbirds has seen an exponential growth since the year 2000. In addition to the increase in the sheer quantity of research, there also have been substantial advances in the tools and approaches used to interrogate the song circuit and, consequently, in the detailed understanding of the neural mechanisms underlying song learning, production, and perception. This SHAR volume assembles a collection of reviews on central aspects of avian communication with each chapter weaving together historical perspectives and modern advances. Most chapters provide conceptual overviews of the mechanistic aspects of birdsong, ranging from the neural circuits underlying song learning and performance to the genomic architecture underlying the production, perception, and acquisition of avian communication. Further, most authors integrate findings from multiple species of songbirds to highlight shared and divergent mechanisms underlying birdsong. They also consider questions for further investigations and technological advances that can help spur such research. Given the importance of examining behaviors from multiple levels of analysis, this volume concludes with a review of vocal performance from functional and evolutionary perspectives.

Because songbirds are among the few vertebrates that, like humans, must learn how to produce their vocal communication signals, songbirds provide a unique and powerful model system to analyze neural mechanisms underlying vocal learning. In Chap. 2, Sakata and Yazaki-Sugiyama provide a circuit perspective on the neural mechanisms underlying song learning. Core principles of their chapter deal with the extent of overlap in the neural substrates for sensory versus sensorimotor learning of birdsong and the regulation of the timing of critical periods for sensory and sensorimotor learning. In addition, they discuss the mechanisms underlying the contribution of social interactions and biological (innate) predispositions to song learning.

Birdsong is central in the social life of songbirds; consequently, it is critical for songbirds to produce their vocalizations with a high degree of consistency throughout their adult life. In Chap. 3, Murphy, Lawley, Smith, and Prather provide an overview of the neural circuits that regulate song control. Their review concentrates on the processes required to maintain accurate vocalizations in adulthood, in particular, the roles that auditory feedback and song template comparisons play. Because the processes of song control and maintenance depend on the acquisition of a sensory representation of song during development, their review integrates discussions about the neural circuitry for vocal learning.

A central node in the neural circuitry for birdsong is Area X, a part of the avian basal ganglia that is specialized for acoustic communication (Doupe et al. 2005). In Chap. 4, Leblois and Perkel provide a comprehensive examination of the neural circuitry in Area X and the degree to which the function, organization, and molecular and cellular composition of Area X parallel those of mammalian basal ganglia circuitry. They provide extensive discussions of the various cell types in Area X and how these individual populations of neurons function in the context of song learning and control. Moreover, they consider the function of Area X in motor learning, execution, and plasticity in light of a number of models of basal ganglia function and discuss how insights gained from Area X may contribute to a unified understanding of the basal ganglia in vertebrates.

Advances in molecular and neurophysiological methods have allowed songbird researchers to gain deep insights into the organization of the auditory system of songbirds and other birds and to establish that the avian auditory system is homologous to the mammalian auditory system in many ways. In Chap. 5, Woolley and Woolley provide an overview of the organization of the auditory system in songbirds and the extent of the homologies in the auditory circuitry across songbirds and mammals against the backdrop of evolutionary and comparative neuroscience. Their review summarizes the heterogeneity and the hierarchy of song processing across nodes in the auditory circuit as well as the role of those circuits in pivotal behaviors, ranging from song learning to mate choice.

Neuromodulation of neural activity in sensory and sensorimotor circuits affects various aspects of vocal communication. In Chap. 6, Remage-Healey provides a conceptual overview of the role of steroids in the modulation of auditory processing. The chapter reviews the history of research into steroid function and the recent changes in perspectives on steroid function with a central focus on brain-derived estrogens. The chapter also highlights the complexity of steroid modulation of auditory processing and emphasizes how the nature of this neuromodulation depends on sex, age, and brain hemisphere.

Songbirds depend not only on songs but also on calls for social communication (as in nonsongbirds). Despite the importance of calls for intraspecific and interspecific communication, relatively little is known about the neural underpinnings of call production and processing. In Chap. 7, Elie and Theunissen provide a thorough examination of the forebrain, midbrain, and hindbrain circuits that control vocal production and perception. The central goals of their review were to evaluate the extent of overlap in the neural regulation of song versus calls and to assess the degree to which insights about song production and perception can help us understand mechanisms of call production and perception (and vice versa). Their review is instrumental in contextualizing research on birdsong in light of the broader communicative repertoires of birds.

The organization and function of neural circuits is critically shaped by gene expression. In Chap. 8, London reviews genomic and genetic contributions to vocal communication in songbirds. The chapter progresses from describing the *Central Dogma of Molecular Biology* to reviewing the contemporary understanding of the various ways in which the genome contributes to the organization and function of the nervous system. The chapter highlights the myriad ways in which the genome is regulated and involved in the processing, acquisition, and production of song and calls.

In discussing mechanisms of avian communication from neurophysiological, neurochemical, and genomic perspectives, it is important to situate these mechanisms in an evolutionary context to consider how neural mechanisms that shape vocal performances have long-term consequences for organisms. In Chap. 9, Podos and Sung review the functional consequences of vocal performance. The authors provide a comparative overview of consequences and constraints on vocal performance with a unifying perspective on the evolution of birdsong. The chapter serves as a bridge between levels of analysis and as a roadmap for how researchers can (and should) integrate evolutionary considerations into their mechanistic investigations and take advantage of mechanistic insights when conceptualizing evolutionary forces and trajectories.

Together, these chapters provide a synthesis of the mechanisms and functions of many facets of avian communication with a central focus on the learning, control, and perception of birdsong. In addition, these chapters serve as a road map for future investigators to provide deeper and broader insights into the neuroethology of birdsong.

Compliance With Ethics Requirements Jon T. Sakata declares that he has no conflict of interest. Sarah C. Woolley declares that she has no conflict of interest.

References

Andalman AS, Fee MS (2009) A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. Proc Natl Acad Sci U S A 106:12518–12523. https://doi.org/10.1073/pnas.0903214106

Anzures G, Quinn PC, Pascalis O, Slater AM, Tanaka JW, Lee K (2013) Developmental origins of the other-race effect. Curr Dir Psychol Sci 22:173–178. https://doi.org/10.1177/0963721412474459

Arnold AP, Nottebohm F, Pfaff DW (1976) Hormone concentrating cells in vocal control and other areas of the brain of the zebra finch (*Poephila guttata*). J Comp Neurol 165:487–511. https://doi.org/10.1002/cne.901650406

Arriaga G, Jarvis ED (2013) Mouse vocal communication system: are ultrasounds learned or innate? Brain Lang 124:96–116. https://doi.org/10.1016/j.bandl.2012.10.002

Arriaga G, Zhou EP, Jarvis ED (2012) Of mice, birds, and men: the mouse ultrasonic song system has some features similar to humans and song-learning birds. PLoS One 7:e46610. https://doi.org/10.1371/journal.pone.0046610

Baptista LF, Petrinovich L (1986) Song development in the white-crowned sparrow: social factors and sex differences. Anim Behav 34:1359–1371. https://doi.org/10.1016/S0003-3472(86)80207-X

Barrington D (1773) XXXI. Experiments and observations on the singing of birds, by the Hon. Daines Barrington, vice Pres. R. S. in a letter to Mathew Maty, M. D. sec. R. S. Phil Trans R Soc London 63:249–291. https://doi.org/10.1098/rstl.1773.0031

Beecher MD (2017) Birdsong learning as a social process. Anim Behav 124:233–246. https://doi.org/10.1016/j.anbehav.2016.09.001

Beecher MD, Brenowitz EA (2005) Functional aspects of song learning in songbirds. Trends Ecol Evol 20:143–149. https://doi.org/10.1016/j.tree.2005.01.004

Beitchman JH, Wilson B, Johnson CJ, Atkinson L, Young A, Adlaf E, Escobar M, Douglas L (2001) Fourteen-year follow-up of speech/language-impaired and control children: psychiatric outcome. J Am Acad Child Psy 40:75–82. https://doi.org/10.1097/00004583-200101000-00019

Berwick RC, Okanoya K, Beckers GJ, Bolhuis JJ (2011) Songs to syntax: the linguistics of birdsong. Trends Cogn Sci 15:113–121. https://doi.org/10.1016/j.tics.2011.01.002

Bolhuis JJ, Gahr M (2006) Neural mechanisms of birdsong memory. Nat Rev Neurosci 7:347–357. https://doi.org/10.1038/nrn1904

Bottjer SW, Altenau B (2010) Parallel pathways for vocal learning in basal ganglia of songbirds. Nat Neurosci 13:153–155. https://doi.org/10.1038/nn.2472

Bradbury JW, Vehrencamp SL (2011) Principles of animal communication, 2nd edn. Sinauer Associates, Sunderland

Brainard MS, Doupe AJ (2000) Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. Nature 404:762–766. https://www.nature.com/articles/35008083

Brainard MS, Doupe AJ (2002) What songbirds teach us about learning. Nature 417:351–358. https://doi.org/10.1038/417351a

Brainard MS, Doupe AJ (2013) Translating birdsong: songbirds as a model for basic and applied medical research. Ann Rev Neurosci 36:489–517. https://doi.org/10.1146/annurev-neuro-060909-152826

Byers BE, Kroodsma DE (2009) Female mate choice and songbird song repertoires. Anim Behav 77:13–22. https://doi.org/10.1016/j.anbehav.2008.10.003

Catchpole CK, Slater PJB (2008) Bird song: biological themes and variations. Cambridge University Press

Chaiken M, Böhner J, Marler P (1993) Song acquisition in European starlings, *Sturnus vulgaris*: a comparison of the songs of live-tutored, tape-tutored, untutored, and wild-caught males. Anim Behav 46:1079–1090. https://doi.org/10.1006/anbe.1993.1298

Charlesworth JD, Warren TL, Brainard MS (2012) Covert skill learning in a cortical-basal ganglia circuit. Nature 486:251–255. https://doi.org/10.1038/nature11078

Chen Y, Matheson LE, Sakata JT (2016) Mechanisms underlying the social enhancement of vocal learning in songbirds. Proc Natl Acad Sci U S A 113:6641–6646. https://doi.org/10.1073/pnas.1522306113

Clayton DF, Balakrishnan CN, London SE (2009) Integrating genomes, brain and behavior in the study of songbirds. Curr Biol 19:R865–R873. https://doi.org/10.1016/j.cub.2009.07.006

Colombelli-Négrel D, Hauber ME, Robertson J, Sulloway FJ, Hoi H, Griggio M, Kleindorfer S (2012) Embryonic learning of vocal passwords in superb fairy-wrens reveals intruder cuckoo nestlings. Curr Biol 22:2155–2160. https://doi.org/10.1016/j.cub.2012.09.025

Colombelli-Négrel D, Hauber ME, Kleindorfer S (2014) Prenatal learning in an Australian songbird: habituation and individual discrimination in superb fairy-wren embryos. Proc Royal Soc B 281:20141154. https://doi.org/10.1098/rspb.2014.1154

Conti-Ramsden G, Mok PLH, Pickles A, Durkin K (2013) Adolescents with a history of specific language impairment (SLI): strengths and difficulties in social, emotional and behavioral functioning. Res Dev Disabil 34:4161–4169. https://doi.org/10.1016/j.ridd.2013.08.043

Cornez G, Madison FN, Van der Linden A, Cornil C, Yoder KM, Ball GF, Balthazart J (2017) Perineuronal nets and vocal plasticity in songbirds: a proposed mechanism to explain the difference between closed-ended and open-ended learning. Dev Neurobiol 77:975–994. https://doi.org/10.1002/dneu.22485

Cynx J, von Rad U (2001) Immediate and transitory effects of delayed auditory feedback on bird song production. Anim Behav 62:305–312. https://doi.org/10.1006/anbe.2001.1744

Dalziell AH, Welbergen JA, Igic B, Magrath RD (2015) Avian vocal mimicry: a unified conceptual framework. Biol Rev 90:643–668. https://doi.org/10.1111/brv.12129

Deshpande M, Pirlepesov F, Lints T (2014) Rapid encoding of an internal model for imitative learning. Proc Biol Sci B 281:20132630. https://doi.org/10.1098/rspb.2013.2630

Devoogd TJ (2004) Where is the bird? Annals N Y Acad Sci 1016:778–786. https://doi.org/10.1196/annals.1298.033

Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. Ann Rev Neurosci 22:567–631. https://doi.org/10.1146/annurev.neuro.22.1.567

Doupe AJ, Perkel DJ, Reiner A, Stern EA (2005) Birdbrains could teach basal ganglia research a new song. Trends Neurosci 28:353–363. https://doi.org/10.1016/j.tins.2005.05.005

Doya K, Sejnowski TJ (1995) A novel reinforcement model of birdsong vocalization learning. In: Advances in neural information processing. MIT Press, pp 101–108

Eales LA (1989) The influences of visual and vocal interaction on song learning in Zebra finches. Anim Behav 37:507–508. https://doi.org/10.1016/0003-3472(89)90097-3

Egnor SER, Hauser MD (2004) A paradox in the evolution of primate vocal learning. Trends Neurosci 27:649–654. https://doi.org/10.1016/j.tins.2004.08.009

Egnor SER, Wickelgren JG, Hauser MD (2007) Tracking silence: adjusting vocal production to avoid acoustic interference. J Comp Physiol A 193:477–483. https://doi.org/10.1007/s00359-006-0205-7

Eigsti I-M, de Marchena AB, Schuh JM, Kelley E (2011) Language acquisition in autism spectrum disorders: a developmental review. Res Autism Spect Dis 5:681–691. https://doi.org/10.1016/j.rasd.2010.09.001

Farabaugh SM, Linzenbold A, Dooling RJ (1994) Vocal plasticity in budgerigars (*Melopsittacus undulatus*): evidence for social factors in the learning of contact calls. J Comp Psychol 108:81–92. https://doi.org/10.1037/0735-7036.108.1.81

Fee MS, Goldberg JH (2011) A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. Neurosci 198:152–170. https://doi.org/10.1016/j.neuroscience.2011.09.069

Feenders G, Liedvogel M, Rivas M, Zapka M, Horita H, Hara E, Wada K, Mouritsen H, Jarvis ED (2008) Molecular mapping of movement-associated areas in the avian brain: a motor theory for vocal learning origin. PLoS One 3:e1768. https://doi.org/10.1371/journal.pone.0001768

Fitch WT, Huber L, Bugnyar T (2010) Social cognition and the evolution of language: constructing cognitive phylogenies. Neuron 65:795–814. https://doi.org/10.1016/j.neuron.2010.03.011

Francis CD, Ortega CP, Cruz A (2011) Vocal frequency change reflects different responses to anthropogenic noise in two suboscine tyrant flycatchers. Proc Royal Soc B 278:2025–2031. https://doi.org/10.1098/rspb.2010.1847

Friedmann N, Rusou D (2015) Critical period for first language: the crucial role of language input during the first year of life. Curr Opin Neurobiol 35:27–34. https://doi.org/10.1016/j.conb.2015.06.003

Fujimoto H, Hasegawa T, Watanabe D (2011) Neural coding of syntactic structure in learned vocalizations in the songbird. J Neurosci 31:10023–10033. https://doi.org/10.1523/JNEUROSCI.1606-11.2011

Gadagkar V, Puzerey PA, Chen R, Baird-Daniel E, Farhang AR, Goldberg JH (2016) Dopamine neurons encode performance error in singing birds. Science 354:1278–1282. https://doi.org/10.1126/science.aah6837

Garamszegi LZ, Eens M (2004) Brain space for a learned task: strong intraspecific evidence for neural correlates of singing behavior in songbirds. Brain Res Rev 44:187–193. https://doi.org/10.1016/j.brainresrev.2003.12.001

Gentry KE, McKenna MF, Luther DA (2018) Evidence of suboscine song plasticity in response to traffic noise fluctuations and temporary road closures. Bioacoustics 27:165–181. https://doi.org/10.1080/09524622.2017.1303645

Goller M, Shizuka D (2018) Evolutionary origins of vocal mimicry in songbirds. Evol Lett 2:417–426. https://doi.org/10.1002/evl3.62

Greenberg S, Ainsworth WA, Popper AN, Fay RR (eds) (2004) Speech processing in the auditory system. Springer, New York

Gultekin YB, Hage SR (2017) Limiting parental feedback disrupts vocal development in marmoset monkeys. Nat Commun 8:14046. https://doi.org/10.1038/ncomms14046

Hage SR, Nieder A (2016) Dual neural network model for the evolution of speech and language. Trends Neurosci 39:813–829. https://doi.org/10.1016/j.tins.2016.10.006

Hara E, Rivas MV, Ward JM, Okanoya K, Jarvis ED (2012) Convergent differential regulation of Parvalbumin in the brains of vocal learners. PLoS One 7:e29457. https://doi.org/10.1371/journal.pone.0029457

Helekar SA (ed) (2013) Animal models of speech and language disorders. Springer, New York

Hessler NA, Doupe AJ (1999) Social context modulates singing-related neural activity in the songbird forebrain. Nat Neurosci 2:209–211. https://doi.org/10.1038/6306

Hisey E, Kearney MG, Mooney R (2018) A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. Nat Neurosci 21:589–597. https://doi.org/10.1038/s41593-018-0092-6

Holy TE, Guo Z (2005) Ultrasonic songs of male mice. PLoS Biol 3:e386. https://doi.org/10.1371/journal.pbio.0030386

Horn G (2004) Pathways of the past: the imprint of memory. Nat Rev Neurosci 5:108–120. https://doi.org/10.1038/nrn1324

Janik VM (2014) Cetacean vocal learning and communication. Curr Opin Neurobiol 28:60–65. https://doi.org/10.1016/j.conb.2014.06.010

Jarvis ED, Mello CV (2000) Molecular mapping of brain areas involved in parrot vocal communication. J Comp Neurol 419:1–31. https://doi.org/10.1002/(SICI)1096-9861(20000327)419:1<1::AID-CNE1>3.0.CO;2-M

Jarvis ED, Ribeiro S, da Silva ML, Ventura D, Vielliard J, Mello CV (2000) Behaviourally driven gene expression reveals song nuclei in hummingbird brain. Nature 406:628–632. https://doi.org/10.1038/35020570

Kao MH, Doupe AJ, Brainard MS (2005) Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song. Nature 433:638–643. https://doi.org/10.1038/nature03127

Kao MH, Wright BD, Doupe AJ (2008) Neurons in a forebrain nucleus required for vocal plasticity rapidly switch between precise firing and variable bursting depending on social context. J Neurosci 28:13232–13247. https://doi.org/10.1523/JNEUROSCI.2250-08.2008

Keller GB, Hahnloser RHR (2009) Neural processing of auditory feedback during vocal practice in a songbird. Nature 457:187–190. https://doi.org/10.1038/nature07467

Kight CR, Swaddle JP (2011) How and why environmental noise impacts animals: an integrative, mechanistic review. Ecol Lett 14:1052–1061. https://doi.org/10.1111/j.1461-0248.2011.01664.x

Kikusui T, Nakanishi K, Nakagawa R, Nagasawa M, Mogi K, Okanoya K (2011) Cross fostering experiments suggest that mice songs are innate. PLoS One 6:e17721. https://doi.org/10.1371/journal.pone.0017721

Knörnschild M (2014) Vocal production learning in bats. Curr Opin Neurobiol 28:80–85. https://doi.org/10.1016/j.conb.2014.06.014

Kroodsma DE, Pickert R (1980) Environmentally dependent sensitive periods for avian vocal learning. Nature 288:477–479. https://doi.org/10.1038/288477a0

Kuhl PK (2007) Is speech learning 'gated' by the social brain? Dev Sci 10:110–120. https://doi.org/10.1111/j.1467-7687.2007.00572.x

Kuhl PK (2010) Brain mechanisms in early language acquisition. Neuron 67:713–727. https://doi.org/10.1016/j.neuron.2010.08.038

Lattenkamp EZ, Vernes SC (2018) Vocal learning: a language-relevant trait in need of a broad cross-species approach. Curr Opin Behav Sci 21:209–215. https://doi.org/10.1016/j.cobeha.2018.04.007

Liu W, Wada K, Jarvis ED, Nottebohm F (2013) Rudimentary substrates for vocal learning in a suboscine. Nat Commun 4. https://doi.org/10.1038/ncomms3082

Louder MI, Balakrishnan CN, Louder AA, Driver RJ, London SE, Hauber ME (2019) An acoustic password enhances auditory learning in juvenile brood parasitic cowbirds. Curr Biol 23: 4045–4051. https://doi.org/10.1016/j.cub.2019.09.046

MacDougall-Shackleton SA (2011) The levels of analysis revisited. Phil Trans Royal Soc B 366:2076–2085. https://doi.org/10.1098/rstb.2010.0363

Mandelblat-Cerf Y, Las L, Denisenko N, Fee MS (2014) A role for descending auditory cortical projections in songbird vocal learning. eLife 3. https://doi.org/10.7554/eLife.02152

Marler P (1970) A comparative approach to vocal learning: song development in white-crowned sparrows. J Comp Physiol Psych 71:1–25. https://doi.org/10.1037/h0029144

Marler P (1997) Three models of song learning: evidence from behavior. J Neurobiol 33:501–516. https://doi.org/10.1002/(SICI)1097-4695(19971105)33:5<501::AID-NEU2>3.0.CO;2-8

Marler P, Peters S (1981) Sparrows learn adult song and more from memory. Science 213:780–782. https://doi.org/10.1126/science.213.4509.780

Marler P, Peters S (1988) The role of song phonology and syntax in vocal learning preferences in the song sparrow, Melospiza melodia. Ethology 77:125–149. https://doi.org/10.1111/j.1439-0310.1988.tb00198.x

Mooney R (2009) Neurobiology of song learning. Curr Opin Neurobiol 19:654–660. https://doi.org/10.1016/j.conb.2009.10.004

Moore JM, Szekely T, Buki J, DeVoogd TJ (2011) Motor pathway convergence predicts syllable repertoire size in oscine birds. Proc Natl Acad Sci U S A 108:16440–16445. https://doi.org/10.1073/pnas.1102077108

Moore JM, Woolley SM (2019) Emergent tuning for learned vocalizations in auditory cortex. Nat Neurosci 22:1469–76. https://doi.org/10.1038/s41593-019-0458-4

Murphy K, James LS, Sakata JT, Prather JF (2017) Advantages of comparative studies in songbirds to understand the neural basis of sensorimotor integration. J Neurophysiol 118:800–816. https://doi.org/10.1152/jn.00623.2016

Nagel KI, Doupe AJ (2006) Temporal processing and adaptation in the songbird auditory forebrain. Neuron 51:845–859. https://doi.org/10.1016/j.neuron.2006.08.030

Nottebohm F, Stokes TM, Leonard CM (1976) Central control of song in the canary,Serinus canarius. J Comp Neurol 165:457–486. https://doi.org/10.1002/cne.901650405

Nowicki S, Searcy WA (2014) The evolution of vocal learning. Curr Opin Neurobiol 28:48–53. https://doi.org/10.1016/j.conb.2014.06.007

Odom KJ, Hall ML, Riebel K, Omland KE, Langmore NE (2014) Female song is widespread and ancestral in songbirds. Nat Commun 5:3379. https://doi.org/10.1038/ncomms4379

Pepperberg IM (2010) Vocal learning in Grey parrots: A brief review of perception, production, and cross-species comparisons. Brain Lang 115:81–91. https://doi.org/10.1016/j.bandl.2009.11.002

Peters S, Marler P, Nowicki S (1992) Song sparrows learn from limited exposure to song models. Condor 94:1016–1019. https://doi.org/10.2307/1369302

Petkov CI, Jarvis ED (2012) Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. Front Evol Neurosci 4. https://doi.org/10.3389/fnevo.2012.00012

Pfenning AR, Hara E, Whitney O, Rivas MV, Wang R, Roulhac PL, Howard JT, Wirthlin M, Lovell PV, Ganapathy G, Mountcastle J, Moseley MA, Thompson JW, Soderblom EJ, Iriki A, Kato M, Gilbert MTP, Zhang G, Bakken T, Bongaarts A, Bernard A, Lein E, Mello CV, Hartemink AJ, Jarvis ED (2014) Convergent transcriptional specializations in the brains of humans and song-learning birds. Science 346:1256846. https://doi.org/10.1126/science.1256846

Portfors CV, Perkel DJ (2014) The role of ultrasonic vocalizations in mouse communication. Curr Opin Neurobiol 28:115–120. https://doi.org/10.1016/j.conb.2014.07.002

Prat Y, Taub M, Yovel Y (2015) Vocal learning in a social mammal: demonstrated by isolation and playback experiments in bats. Sci Adv 1:e1500019. https://doi.org/10.1126/sciadv.1500019

Prat Y, Azoulay L, Dor R, Yovel Y (2017) Crowd vocal learning induces vocal dialects in bats: playback of conspecifics shapes fundamental frequency usage by pups. PLoS Biol 15:e2002556. https://doi.org/10.1371/journal.pbio.2002556

Reiner A, Perkel DJ, Bruce LL, Butler AB, Csillag A, Kuenzel W, Medina L, Paxinos G, Shimizu T, Striedter G, Wild M, Ball GF, Durand S, Güntürkün O, Lee DW, Mello CV, Powers A, White SA, Hough G, Kubikova L, Smulders TV, Wada K, Dugas-Ford J, Husband S, Yamamoto K, Yu J, Siang C, Jarvis ED (2004) Revised nomenclature for avian telencephalon and some related brainstem nuclei. J Comp Neurol 473:377–414. https://doi.org/10.1002/cne.20118

Riebel K, Odom KJ, Langmore NE, Hall ML (2019) New insights from female bird song: towards an integrated approach to studying male and female communication roles. Biol Lett 15:20190059. https://doi.org/10.1098/rsbl.2019.0059

Ríos-Chelén AA, Cuatianquiz-Lima C, Bautista A, Martínez-Gómez M (2018) No reliable evidence for immediate noise-induced song flexibility in a suboscine. Urban Ecosyst 21:15–25. https://doi.org/10.1007/s11252-017-0690-1

Roberts TF, Mooney R (2013) Motor circuits help encode auditory memories of vocal models used to guide vocal learning. Hear Res 303:48–57. https://doi.org/10.1016/j.heares.2013.01.009

Roberts TF, Tschida KA, Klein ME, Mooney R (2010) Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning. Nature 463:948–952. https://doi.org/10.1038/nature08759

Roberts TF, Hisey E, Tanaka M, Kearney MG, Chattree G, Yang CF, Shah NM, Mooney R (2017) Identification of a motor-to-auditory pathway important for vocal learning. Nat Neurosci 20:978–986. https://doi.org/10.1038/nn.4563

Roy S, Miller CT, Gottsch D, Wang X (2011) Vocal control by the common marmoset in the presence of interfering noise. J Exp Biol 214:3619–3629. https://doi.org/10.1242/jeb.056101

Sakata JT, Brainard MS (2006) Real-time contributions of auditory feedback to avian vocal motor control. J Neurosci 26:9619–9628. https://doi.org/10.1523/JNEUROSCI.2027-06.2006

Sakata JT, Brainard MS (2008) Online contributions of auditory feedback to neural activity in avian song control circuitry. J Neurosci 28:11378–11390. https://doi.org/10.1523/JNEUROSCI.3254-08.2008

Sakata JT, Vehrencamp SL (2012) Integrating perspectives on vocal performance and consistency. J Exp Biol 215:201–209. https://doi.org/10.1242/jeb.056911

Seyfarth RM, Cheney DL (2010) Production, usage, and comprehension in animal vocalizations. Brain Lang 115:92–100. https://doi.org/10.1016/j.bandl.2009.10.003

Simion F, Giorgio ED (2015) Face perception and processing in early infancy: inborn predispositions and developmental changes. Front Psychol 6. https://doi.org/10.3389/fpsyg.2015.00969

Slabbekoorn H, Ripmeester EAP (2008) Birdsong and anthropogenic noise: implications and applications for conservation. Mol Ecol 17:72–83. https://doi.org/10.1111/j.1365-294X.2007.03487.x

Snowdon CT, Hausberger M (1997) Social influences on vocal development (Cambridge University Press)

Solis MM, Doupe AJ (1999) Contributions of tutor and Bird's own song experience to neural selectivity in the songbird anterior forebrain. J Neurosci 19:4559–4584. https://doi.org/10.1523/JNEUROSCI.19-11-04559.1999

Solis MM, Doupe AJ (2000) Compromised neural selectivity for song in birds with impaired sensorimotor learning. Neuron 25:109–121. https://doi.org/10.1016/S0896-6273(00)80875-2

Soma MF (2011) Social factors in song learning: a review of Estrildid finch research. Ornithol Sci 10:89–100. https://doi.org/10.2326/osj.10.89

Soma M, Garamszegi LZ (2011) Rethinking birdsong evolution: meta-analysis of the relationship between song complexity and reproductive success. Behav Ecol 22:363–371. https://doi.org/10.1093/beheco/arq219

Stoeger AS, Manger P (2014) Vocal learning in elephants: neural bases and adaptive context. Curr Opin Neurobiol 28:101–107. https://doi.org/10.1016/j.conb.2014.07.001

Suthers RA, Fitch WT, Fay RR, Popper AN (eds) (2016) Vertebrate sound production and acoustic communication. Springer, New York

Takahashi DY, Fenley AR, Teramoto Y, Narayanan DZ, Borjon JI, Holmes P, Ghazanfar AA (2015) The developmental dynamics of marmoset monkey vocal production. Science 349:734–738. https://doi.org/10.1126/science.aab1058

Takahashi DY, Liao DA, Ghazanfar AA (2017) Vocal learning via social reinforcement by infant marmoset monkeys. Curr Biol 27:1844–1852.e6. https://doi.org/10.1016/j.cub.2017.05.004

Takesian AE, Hensch TK (2013) Balancing plasticity/stability across brain development. In: Prog Brain Res Elsevier, pp 3–34. https://doi.org/10.1016/B978-0-444-63327-9.00001-1

Theunissen FE, Sen K, Doupe AJ (2000) Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. J Neurosci 20:2315–2331. https://doi.org/10.1523/JNEUROSCI.20-06-02315.2000

Thierry B (2005) Integrating proximate and ultimate causation: just one more go. Anim Behav 89(4). https://www.jstor.org/stable/24110969

Thorpe WH (1958) The learning of song patterns by birds, with especial reference to the song of the chaffinch *Fringilla Coelebs*. Ibis 100:535–570. https://doi.org/10.1111/j.1474-919X.1958.tb07960.x

Tinbergen N (1951) The study of instinct. Clarendon Press/Oxford University Press, New York

Tumer EC, Brainard MS (2007) Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. Nature 450:1240–1244. https://doi.org/10.1038/nature06390

Tyack PL, Janik VM (2013) Effects of noise on acoustic signal production in marine mammals. In: Brumm H (ed) Animal communication and noise. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 251–271

Watson SK, Townsend SW, Schel AM, Wilke C, Wallace EK, Cheng L, West V, Slocombe KE (2015) Vocal learning in the functionally referential food grunts of chimpanzees. Curr Biol 25:495–499. https://doi.org/10.1016/j.cub.2014.12.032

Werker JF, Hensch TK (2015) Critical periods in speech perception: new directions. Ann Rev Psych 66:173–196. https://doi.org/10.1146/annurev-psych-010814-015104

West MJ, King AP (1988) Female visual displays affect the development of male song in the cowbird. Nature 334:244–246. https://doi.org/10.1038/334244a0

West MJ, King AP, Eastzer DH (1981) Validating the female bioassay of cowbird song: relating differences in song potency to mating success. Anim Behav 29:490–501. https://doi.org/10.1016/S0003-3472(81)80110-8

Woolley SMN (2012) Early experience shapes vocal neural coding and perception in songbirds. Dev Psychobiol 54:612–631. https://doi.org/10.1002/dev.21014

Woolley SC, Rajan R, Joshua M, Doupe AJ (2014) Emergence of context-dependent variability across a basal ganglia network. Neuron 82:208–223. https://doi.org/10.1016/j.neuron.2014.01.039

Xiao L, Chattree G, Oscos FG, Cao M, Wanat MJ, Roberts TF (2018) A basal ganglia circuit sufficient to guide birdsong learning. Neuron 98 e5:208–221. https://doi.org/10.1016/j.neuron.2018.02.020

Yazaki-Sugiyama Y (2019) Neuronal mechanisms regulating the critical period of sensory experience-dependent song learning. Neurosci Res 140:53–58. https://doi.org/10.1016/j.neures.2018.11.002

Chapter 2
Neural Circuits Underlying Vocal Learning in Songbirds

Jon T. Sakata and Yoko Yazaki-Sugiyama

Abstract Songbirds have become the most prominent animal model to reveal the neural mechanisms underlying vocal learning. Similar to other forms of sensorimotor learning, vocal learning in songbirds entails acquiring a sensory representation of communicative sounds (songs) and learning the motor commands to produce the memorized sounds. The process of song learning in songbirds is highly similar to the process of speech acquisition in humans; as such, there are extensive efforts to discover and characterize the neural circuitry for song learning in songbirds to gain insight into potential mechanisms underlying speech acquisition in humans. This chapter provides a conceptual review of the neural circuits regulating the sensory and sensorimotor learning of birdsong as well as the mechanisms underlying variation in the extent and fidelity of vocal learning (e.g., timing of song learning, social interactions, and biological predispositions). Findings from various songbird species are integrated to provide a comparative perspective on neural mechanisms. In addition, outstanding questions and new research directions for birdsong research are considered.

Keywords Auditory feedback · Birdsong · Critical period · Dopamine · HVC · Locus coeruleus · NCM · Norepinephrine · Sensory learning · Speech · Ventral tegmental area · Zebra finch

---

J. T. Sakata (✉)
Department of Biology and Centre for Research in Brain, Language, and Music,
McGill University, Montreal, QC, Canada
e-mail: jon.sakata@mcgill.ca

Y. Yazaki-Sugiyama
Okinawa Institute of Science and Technology (OIST) Graduate University, Okinawa, Japan

International Research Center for Neurointelligence (IRCN), The University of Tokyo,
Tokyo, Japan
e-mail: Yazaki-sugiyama@oist.jp

J. T. Sakata et al. (eds.), *The Neuroethology of Birdsong*, Springer Handbook
of Auditory Research 71, https://doi.org/10.1007/978-3-030-34683-6_2

2.1 Songbirds as a Model to Reveal Neural Mechanisms Underlying Vocal Learning

A wide range of animals, including humans, depend on vocal signals for communication (Bradbury and Vehrencamp 2011). For example, many vertebrate species use vocalizations to communicate information about internal state, intention, individual identity, reproductive fitness, and objects in the environment (e.g., food, predators). Given the importance of vocal communication to social behavior, it is critical to reveal fundamental and evolutionarily conserved mechanisms of vocal communication.

The vast majority of species that use vocal signals for communication are born with the ability to produce species-typical communication signals (Petkov and Jarvis 2012; Nowicki and Searcy 2014). However, some species, including songbirds and humans, must learn how to produce their vocalizations to effectively communicate (*vocal production learning* or *vocal learning*) (Catchpole and Slater 2008; Sakata and Woolley, Chap. 1). Indeed, decades of research have revealed that the process of vocal learning in songbirds shares many important features with the process of speech acquisition in humans (Fig. 2.1) (Doupe and Kuhl 1999; Mooney 2009). For example, both songbirds and humans memorize the vocalizations of adults and engage in vocal practice or *babbling* to refine their vocalizations into sound patterns that resemble the memorized vocalizations. Because hearing is necessary for memorizing communicative sounds and for providing auditory feedback about vocal performance, hearing is central throughout the process of vocal learning across songbirds and humans (Prather 2013; Roberts and Mooney 2013). Furthermore, songbirds possess discrete neural circuits for song learning that are analogous to neural circuits implicated in speech acquisition (Fig. 2.2) (Doupe et al. 2005; Brainard and Doupe 2013). Consequently, songbirds provide a powerful opportunity to reveal general mechanisms of vocal learning and to generate and test mechanistic models of speech acquisition (Mooney 2009; Kuhl 2010).

Fig. 2.1 Schematic of song learning in songbirds. Like the process of speech acquisition in humans, song learning in songbirds involves *sensory* and *sensorimotor* learning. In songbirds like the zebra finch (*Taeniopygia guttata*), under normal circumstances, sensory learning occurs during a critical period in development starting ~20–30 days post-hatch and ending ~60 days post-hatch. Sensorimotor learning is concentrated around 40–90 days post-hatch. Whereas the sensory and sensorimotor period overlap in songbirds like the zebra finch, these period are temporally separated in other species (e.g., swamp sparrows, *Melospiza georgiana*) (see Sakata and Woolley, Chap. 1)

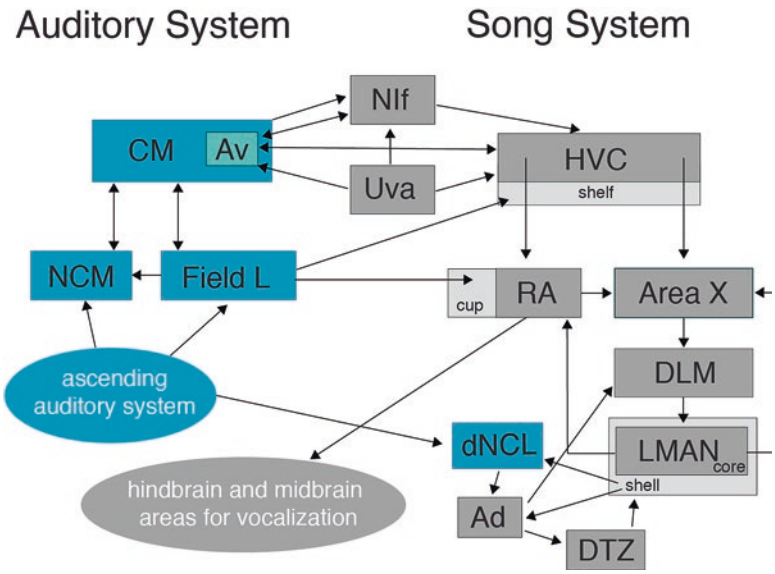


Fig. 2.2 Summary of some of the neural circuitry implicated in song learning. *Blue boxes* denote auditory processing areas, whereas *dark grey boxes* denote sensorimotor brain areas. *Light grey boxes* adjacent to or surrounding darker grey boxes denote cup or shell regions around particular brain areas. This diagram does not include some parts of the ascending auditory pathway (see Woolley and Woolley, Chap. 5; Elie and Theunissen, Chap. 7) as well as some other brain areas that provide inputs into these areas (e.g., cerebellum; midbrain and hindbrain catecholaminergic areas). *Ad*, dorsal arcopallium; *Area X* (vocal portion of the basal ganglia); *Av*, avalanche region of CM, shown in lighter blue; *CM*, caudal mesopallium (includes CMM and CLM); *DLM*, medial portion of the dorsolateral thalamic nucleus; *dNCL*, dorsal region of the caudolateral nidopallium; *DTZ*, dorsal thalamic zone; *Field L*; *HVC* (used as proper name); *LMAN*, lateral magnocellular nucleus of the anterior nidopallium; *NCM*, caudomedial nidopallium; *NIf*, nucleus interfacialis of the nidopallium; *RA*, robust nucleus of the arcopallium; *Uva*, nucleus uvaeformis

Fundamentally, song learning involves acquiring a sensory representation of a target song (*sensory learning*) and learning how to generate the motor commands to produce that target song (*sensorimotor learning*) (Troyer and Doupe 2000; Mooney 2009). Sensorimotor learning critically depends on sensory learning because individuals learn how to produce the target sounds by comparing current vocalizations to target sounds (see Sects. 2.2, 2.3) (Doya and Sejnowski 1995; Fee and Goldberg 2011). Sensory and sensorimotor learning can occur at distinct times and involve distinct processes (Fig. 2.1) (see Sakata and Woolley, Chap. 1); consequently, the extent to which neural circuits for sensory and sensorimotor learning are shared or distinct remains a central question in the neuroethology of birdsong.

2.1.1 Critical Periods for Song Learning

Experiences during critical periods (CPs) in development have lasting impacts on sensory processing, cognitive function, and behavior (Hensch 2004; Knudsen 2004). The concept of CPs has its origins in developmental biology and is based on findings that drug administration at specific time periods (but not others) causes specific malformations in developing embryos (Stockard 1921). For behavior, the concept of CPs stems from findings by researchers, such as Konrad Lorenz (e.g., Lorenz 1935), Patrick Bateson (e.g., Bateson 1966), and Gabriel Horn (reviewed in Horn 2004), who collectively detailed how developing animals *imprint* on individuals or other visual stimuli that they are exposed to during a particular developmental time window (Michel and Tyler 2005). Because sensory experiences shape neural representations of the world and guide future behavior, how experiences during CPs affect neural circuit development has been widely investigated (Knudsen 2004; Takesian and Hensch 2013). Furthermore, because CPs represent heightened periods of neural and behavioral plasticity, understanding the mechanisms underlying the opening and closing of CPs is fundamental to understanding general mechanisms of learning, memory, and behavioral plasticity.

Individuals learn how to produce their vocalizations during CPs in development. Indeed, the existence of CPs for speech and language acquisition and perception has firmly been established (Werker and Hensch 2015). In general, vocal learning in humans can be characterized by an early CP of purely perceptual learning during which individuals memorize the sounds of the vocalizations of others (*sensory period*) and a period during which individuals monitor and gradually improve their own vocalizations to match the learned perceptual targets (*sensorimotor period*) (Hickok et al. 2011; Werker and Hensch 2015). Like speech acquisition in humans, song learning in songbirds (Fig. 2.1) is characterized by an early CP of sensory learning (the sensory period) followed by a period of sensorimotor learning (the sensorimotor period) (Brainard and Doupe 2013). Details of the timing and the durations of CPs for song learning have been revealed through behavioral studies, which have been conducted most extensively in zebra finches (*Taeniopygia guttata*) (Brainard and Doupe 2002; Gobes et al. 2017). Generally, these studies describe developmental timeframes in which sensory and sensorimotor experiences are important for song development (Gobes et al. 2017; see also Sakata and Woolley, Chap. 1 for discussion of species variation in CPs for song learning). For example, juvenile zebra finches are able to memorize the songs of conspecifics before they themselves begin to sing, indicating that the sensory period opens before the sensorimotor period. Further, exposing normally raised zebra finches to different tutors later than 60 days post-hatching (dph), a period when they are still developmentally honing their songs, does not lead to additional song learning, indicating that the sensory period ends before the sensorimotor period. While much remains to be explored, these studies lay the foundation for understanding how CPs regulate song learning.

2.1.2 Neural Circuitry for Song Learning, Production, and Perception

Decades of research on songbirds have revealed a complex network of brain areas that regulate song learning (Fig. 2.2). Fundamentally, song learning involves interactions among neurons within and across sensory (auditory), motor, sensorimotor, and reinforcement circuits. This section provides a brief overview of these circuits to contextualize subsequent sections that provide details of the contribution of specific neural populations to sensory and sensorimotor learning.

The ascending auditory pathway of songbirds is homologous to that observed in other vertebrates, including mammals (Woolley and Woolley, Chap. 5; Elie and Theunissen, Chap. 7). Information from the cochlea ascends to auditory processing areas in the pallium (*cortex*) via hindbrain, midbrain, and thalamic brain areas (Woolley 2017). Field L, in particular the L2 sub-region of Field L, is homologous and functionally analogous to the input layer of the mammalian primary auditory cortex (Wang et al. 2010; Woolley 2017). Field L neurons project out to higher-order auditory areas (e.g., the caudomedial mesopallium [CMM] and caudomedial nidopallium [NCM]; all abbreviations appear in Table 2.1 that are important for auditory processing, learning, and memory).

Importantly, neurons that process auditory information ultimately project to sensorimotor brain areas that integrate auditory information and produce motor commands for song. The canonical circuitry underlying song plasticity and control (*song system*) includes the nucleus interfacialis of the nidopallium (NIf), HVC (used as a proper name), the robust nucleus of the arcopallium (RA), the vocal basal ganglia nucleus Area X, the medial portion of the dorsolateral thalamic nucleus (DLM), and the lateral magnocellular nucleus of the anterior nidopallium (LMAN; Fig. 2.2) (see also Murphy, Lawley, Smith, and Prather, Chap. 3). The song system is parsed into the *vocal motor pathway* (VMP), which includes NIf, HVC, and RA (as well as hindbrain areas that contain vocal motor and respiratory neurons) and the *anterior forebrain pathway* (AFP), which is a basal ganglia-thalamic-cortical circuit that includes Area X, DLM, and LMAN. Neurons in the canonical song circuitry are sensorimotor in nature; for example, NIf receives ascending auditory inputs and is one of the first brain areas in the song system that generates motor commands for song (Lewandowski et al. 2013; Vyssotski et al. 2016). The nucleus HVC, which is analogous to the premotor cortex in mammals, receives auditory input from a number of sources, including the caudal mesopallium (CM) and NIf, and is critical for adult song production. It is generally thought that auditory information reaches the rest of the song system (e.g., RA, Area X, and LMAN) predominantly via HVC; in this respect, HVC is considered one of the most important areas in which to reveal mechanisms of auditory processing, sensory learning, and sensorimotor integration (Roberts and Mooney 2013; Murphy et al. 2017).

Table 2.1 Abbreviations

Ad	Dorsal arcopallium
AFP	Anterior forebrain pathway
AIV	Ventral portion of the intermediate arcopallium
Av	Avalanche, a ventral region in CLM
CLM	Caudolateral mesopallium
CM	Caudal mesopallium (CLM + CMM)
CMM	Caudomedial mesopallium
CP	Critical period
DA	Dopamine
DLM	Medial portion of the dorsolateral thalamic nucleus
dNCL	Dorsal region of the caudolateral nidopallium
Dph	Days post-hatch
DTZ	Dorsal thalamic zone
GABA	Gamma-aminobutyric acid
HVC	Used as proper name for vocal motor nucleus in the nidopallium
LC	Locus coeruleus
LMAN	Lateral magnocellular nucleus of anterior nidopallium
LMAN-core	Core region of LMAN
LMAN-shell	Shell region of LMAN
NCM	Caudomedial nidopallium
NE	Norepinephrine
NIf	Nucleus interfacialis of the nidopallium
PAG	Avian periaqueductal gray
PNN	Perineuronal net
PV	Parvalbumin
RA	Robust nucleus of the arcopallium
Uva	Nucleus uvaeformis
VMP	Vocal motor pathway
VP	Ventral pallidum
VTA	Ventral tegmental area
VTA_X	Neurons in the VTA that project to area X

Brain areas outside the canonical auditory and song circuits are also involved in the process of song learning. For example, in addition to the neurons in the *core region* of LMAN (LMAN-core; Fig. 2.2), neurons in the *shell region* of LMAN (LMAN-shell; Fig. 2.2) form a parallel circuit to the canonical circuitry in the AFP and have been implicated in song learning (Bottjer and Altenau 2010; Bottjer et al. 2010). Neurons in the deep cerebellar nucleus indirectly project to Area X and are important for the accurate imitation of song (Nicholson et al. 2018; Pidoux et al. 2018). Given that neuromodulators shape neural plasticity, the role of neuromodulatory inputs into brain areas implicated in sensory and sensorimotor learning has also received considerable attention (see Sect. 2.1). For example, the secondary auditory processing area NCM receives noradrenergic and dopaminergic input, respectively, from the locus coeruleus (LC) and ventral tegmental area (VTA) (Castelino and Schmidt 2010); sensorimotor neurons in HVC, RA, and Area X receive catecholaminergic inputs from the VTA, LC, and the avian periaqueductal gray (PAG) (Castelino and Schmidt 2010); and HVC and RA receive cholinergic inputs that modulate auditory responses to song from the nucleus basalis (Li and Sakaguchi 1997; Shea and Margoliash 2003). Taken together, while current discussions of the neural mechanisms underlying song learning still center around the canonical auditory and song circuits, it has become increasingly important to understand how brain areas outside the song system interact with the canonical song circuit to enable sensory and sensorimotor learning.

Detailed knowledge of the connectivity within and among sensory, sensorimotor, and neuromodulatory circuits sets the structural framework to reveal mechanisms underlying song learning in songbirds. Further studies employing neurophysiological recordings, analyses of gene expression, and targeted manipulations of brain activity build upon this framework and provide direct insight into the functional neuroanatomy underlying the sensory and sensorimotor learning of song.

2.2 Mechanisms Underlying the Sensory Learning of Song

Birdsong, like speech, is a learned behavior that is acquired through imitation of the communicative sounds performed by others (Doupe and Kuhl 1999; Mooney 2009). Similar to humans, developing songbirds rely on sensory input to acquire sensory models (*templates*) of the sounds that they will eventually produce as adults (Bolhuis and Gahr 2006; Brainard and Doupe 2013). Accurately memorizing the song of a tutor is the critical first step for song learning. Without an accurate representation of the tutor's song (target song), juvenile birds cannot learn to produce an accurate imitation of their tutor's song on their own, regardless of the amount of practice. Consequently, understanding the mechanisms underlying the sensory learning of song and the nature of the representation of the tutor's song in the songbird brain are paramount (Hahnloser and Kotowicz 2010; Bolhuis and Moorman 2015).

2.2.1 Neural Circuits for the Sensory Learning of Song

Most studies to date have investigated the contribution of auditory (e.g., NCM) and sensorimotor brain areas (e.g., LMAN, HVC) to the sensory learning of song. For example, many studies provide converging support for the notion that secondary auditory processing areas are critical for the sensory learning of song. Together, NCM and CMM are analogous to the secondary auditory cortex in mammals (Woolley and Woolley, Chap. 5), and genomic, neurophysiological, and pharmacological approaches demonstrate that neurons in these areas are essential for song learning and for encoding the representation of the tutor's song (Jin and Clayton

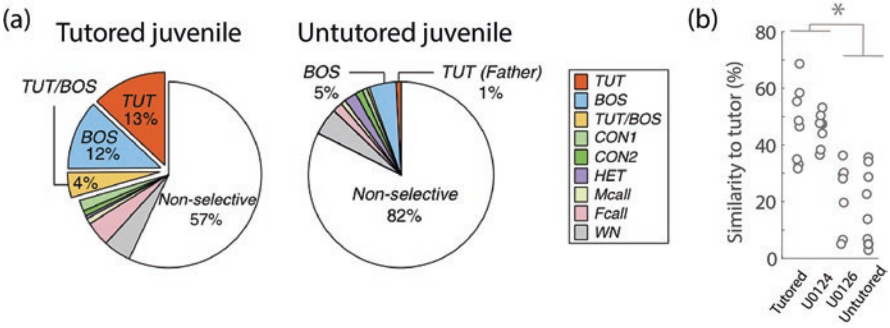


Fig. 2.3 The role of NCM (caudomedial nidopallium) neurons in the sensory learning and encoding of song. **(a)** A proportion of broad-spiking NCM neurons becomes tuned to the tutor's (*TUT*) song after song tutoring (*left*). In contrast, in untutored birds (*right*), very few broad-spiking neurons in NCM are tuned to the sound of ethologically relevant stimuli, including the song of a juvenile's father. Abbreviations: BOS, bird's own song; CON1, 2, conspecific songs; Fcall, female call; HET, heterospecific song; Mcall, male call; TUT, tutor song; WN, white noise. **(b)** Infusion of *U0126*, a chemical that inhibits MAPK/ERK kinases, into NCM during song tutoring prevents juveniles from producing accurate imitations of their tutor's song. *U0124* is a related compound that does not inhibits MAPK/ERK kinases (i.e., control infusion). (**a** adapted with permission from Yanagihara and Yazaki-Sugiyama 2016; **b** data replotted from London and Clayton 2008)

1997; London and Clayton 2008). Importantly, after tutoring, a small but stable percentage of neurons in NCM preferentially or selectively responded to the tutor's song in developing and adult songbirds (e.g., Phan et al. 2006; Yanagihara and Yazaki-Sugiyama 2016). In developing juveniles, approximately 15% of broad-spiking NCM neurons responded more to playbacks of their tutor's song than to playbacks of the songs of other conspecifics or of the bird's own developing song (Fig. 2.3a) (Yanagihara and Yazaki-Sugiyama 2016). This finding is particularly important because the evaluation of song performance requires neuronal populations that accurately encode the sound of the tutor's song (see Sect. 2.3 on sensorimotor learning). The proportion of broad-spiking neurons in NCM that are preferentially tuned to the tutor's song remains stable across development, suggesting that this population could continue to be important for song evaluation throughout development and, possibly, into adulthood.

The importance of NCM neurons to the sensory learning of song has been revealed by manipulations of NCM activity (see London, Chap. 8). For example, manipulation of cellular cascades in NCM during song tutoring disrupted song learning without affecting auditory discrimination abilities (London and Clayton 2008). Specifically, infusing U0196 (a compound that inhibits MAPK/ERK kinases) into NCM when juveniles were exposed to a tutor's song prevented those birds from producing songs as adults that resembled those of their tutor (Fig. 2.3b). In a related study, lesions of NCM significantly reduced a bird's preference for his tutor's song without affecting the ability to discriminate between other songs (Gobes and Bolhuis

2007). In addition to numerous other studies (reviewed in Bolhuis and Moorman 2015), those experiments provide compelling support for the importance of secondary auditory processing areas in the sensory learning of song.

While many studies to date have investigated the role of auditory circuits in the sensory learning of song, historically, the contributions of sensorimotor brain areas to the sensory learning of song have been highlighted (Doupe and Kuhl 1999; Roberts and Mooney 2013). Indeed, one of the first insights into the neural mechanisms underlying the sensory learning of song came from the finding that infusing NMDA antagonists into the sensorimotor nucleus LMAN (the output of an avian basal ganglia circuit devoted to song plasticity; Fig. 2.2) during song tutoring significantly impaired song imitation (Basham et al. 1996). Additional support for the role of LMAN in the sensory learning of song came from the finding that in juveniles that had memorized the song of their tutor and had begun to practice their songs (~45–60 dph), a number of neurons in LMAN were preferentially activated by the sound of the tutor's song relative to the sounds of other conspecific song and the bird's own song (Solis and Doupe 1999). Furthermore, experience with tutor song leads to a number of changes in LMAN, including electrophysiological changes (e.g., Livingston and Mooney 1997; Yazaki-Sugiyama and Mooney 2004) and changes to neurotransmitter systems (e.g., Singh et al. 2000; Heinrich et al. 2002).

While the core region of LMAN continues to be the central focus of mechanistic studies, some studies highlight a contribution of neurons in LMAN-shell to the sensory learning of song. The shell region of LMAN is part of a circuit involving the dorsal arcopallium (Ad), the dorsal thalamic zone (DTZ), and the dorsal region of the caudolateral nidopallium (dNCL) (Fig. 2.2). Roughly 30% of neurons in the LMAN-shell of 45-day-old juvenile zebra finches can be selectively activated by the sound of their tutor's song (Achiro and Bottjer 2013), and the plasticity in tuning that follows song tutoring suggests that this circuit could be important for song evaluation and learning. That manipulations within this circuit during song tutoring lead to deficits in song learning supports this contention (Bottjer and Altenau 2010; Achiro and Bottjer 2013).

The sensorimotor nucleus HVC also plays a significant role in the sensory learning of song (Roberts and Mooney 2013; Zhao et al. 2019). Neurophysiological recordings demonstrated that, like LMAN neurons, HVC neurons also preferentially responded to the tutor song during development (~45–60 dph) (Nick and Konishi, 2005; Warren et al., 2011; Vallentin et al., 2016), and this change in tuning can occur with minimal exposure to a tutor's song (Tanaka et al. 2018). The dendritic morphology of HVC neurons rapidly stabilizes during the sensory learning of song (Roberts et al. 2010); moreover, acute perturbations of HVC activity during song tutoring significantly impaired song learning (Roberts et al. 2012). For example, brief microstimulation of HVC neurons when juveniles hear a specific syllable in the tutor's song will impair the learning of the targeted syllable but not of other syllables in the tutor's song (Roberts et al. 2012). In addition, dopamine (DA) concentrations in HVC increased when juveniles heard a live tutor produce song, and manipulation of DA inputs into HVC (e.g., lesioning DA inputs, blocking DA receptors) during song tutoring significantly inhibited song learning (see also Sect. 2.4.1) (Tanaka et al. 2018).

In summary, neuronal populations in multiple brain circuits encode the sound of the tutor's song (i.e., template) and form a distributed representation of a sensory target. The individual contributions of each of these areas are beginning to be unraveled, but studies to date clearly indicate that auditory and sensorimotor circuitry act in concert to regulate the sensory learning of song.

2.2.2 Neural Mechanisms Underlying Timing of the Critical Period for Sensory Learning

In many species of songbirds, the sensory learning of song occurs during a restricted window in development (sensory period, described in Sect. 2.1.1). Prior examinations into the regulation of this CP have analyzed how developmental changes in song learning relate to developmental changes in dendritic arborization (e.g., Heinrich et al. 2005; Roberts et al. 2010), NMDA receptor subunit expression (e.g., Singh et al. 2000; Heinrich et al. 2002), NMDA receptor currents (e.g., Livingston and Mooney 1997; White et al. 1999), sex steroid hormone receptor expression (e.g., Gahr and Metzdorf 1997), and neuroanatomical connections (e.g., Iyengar and Bottjer 2002; Aronov et al. 2008). Subsequent research into the mechanisms of CP timing for song learning in songbirds was modeled after advances in the understanding of neural plasticity and tuning in the visual and auditory systems in mammals (de Villers-Sidani et al. 2007; Takesian and Hensch 2013). Those studies highlight how the balance of excitation and inhibition and the expression of molecules that restrict plasticity (*molecular brakes*) regulate CP timing.

In mammals, the maturation of inhibitory circuitry triggers the onset of the CP in sensory cortices like the primary visual cortex. For example, the emergence of a particular type of inhibitory cell—large basket parvalbumin-expressing (PV) interneurons—in primary sensory cortices closely corresponds to the onset of the CP for sensory plasticity (Takesian and Hensch 2013). Additionally, manipulations that impede or promote GABA function impede or promote sensory plasticity, respectively (e.g., Fagiolini and Hensch 2000; Southwell et al. 2010). Of particular importance with regard to the role of inhibitory circuitry in CP timing are findings that injections of diazepam, a GABA current accelerator, and transplantation of immature PV neurons into primary sensory areas restored visual plasticity after CP closure (Takesian and Hensch 2013; Tang et al. 2014).

Little is known about how inhibitory circuitry in sensory or sensorimotor circuits regulates the opening of the sensory period for birdsong. Just as in mammals, inhibitory PV neurons are found throughout sensory and sensorimotor brain areas in songbirds (Balmer et al. 2009; Cornez et al. 2015). Furthermore, PV neurons and other inhibitory neurons increase in abundance over the course of development in brain areas like the sensorimotor nucleus RA (Sakaguchi 1996; Cornez et al. 2018). While the pattern of PV neuron expression in some sensorimotor areas corresponds with the opening of the sensory period, PV neuron numbers in the primary auditory cortex and other auditory areas in zebra finches do not dramatically change around the opening of the sensory period (Cornez et al. 2018). This contrasts with the robust changes in PV neuron abundance in primary sensory cortices of mammals around the onset of the CP for sensory plasticity (Takesian and Hensch 2013). It is clear that additional studies of inhibitory neuron distribution and function are required to further understand mechanisms underlying the opening of the sensory period for song learning.

Studies in mammals indicate that the closing of the CP for sensory plasticity is generally determined by the emergence of molecular brakes for neural plasticity (Takesian and Hensch 2013). Examples of molecular brakes include the nogo receptor, the acetylcholine receptor linker LYNX1, and perineuronal nets (PNNs). Perineuronal nets are extracellular matrices that wrap around neurons (e.g., PV neurons) and are hypothesized to stabilize neuronal circuitry (Takesian and Hensch 2013; Sonntag et al. 2015). Indeed, more PV neurons are encased by PNNs in the adult nervous system than in the developing nervous system, and degradation of PNNs in primary sensory cortices can enhance sensory plasticity in adult rodents (Pizzorusso et al. 2006; Happel et al. 2014).

Perineuronal nets, in particular PNNs in sensorimotor brain areas, have been implicated in the closing of the CP for sensory learning in songbirds. As in mammals, PNNs are associated with PV neurons in songbirds, and the developmental time course of PNN expression supports a contribution of PNNs to the closure of CPs for sensory learning (Balmer et al. 2009; Cornez et al. 2018). For example, more PV neurons in sensorimotor areas like HVC, RA, and Area X are surrounded by PNNs in adult birds that sing mature song (~90 dph) than in juveniles that are just beginning to learn song (~30 dph; Balmer et al. 2009; Cornez et al. 2018). Further, preventing juvenile songbirds from hearing song delays the closure of the sensory period and the encasement of PV neurons in HVC and Area X by PNNs (Balmer et al. 2009). Interestingly, the extent of PNN formation in sensorimotor brain areas of the adult songbird brain is reduced in songbird species that can learn to produce new songs as adults, supporting the notion that PNNs serve as brakes for neural plasticity required for song learning (Cornez et al. 2015). Despite these findings, the functional consequences of PNN expression to song learning need to be explicitly examined.

However, studies of sensory plasticity in mammals and songbirds diverge in ways that suggest there may be different mechanisms that regulate the closing of the CP for sensory learning. While sensory experiences during CPs affect the tuning of sensory neurons in both mammals and songbirds (Takesian and Hensch 2013; Yanagihara and Yazaki-Sugiyama 2016), the relationship between PNN expression in sensory processing areas (e.g., Field L, NCM) and the closure of the CP for sensory learning is less conspicuous in songbirds than in mammals. For instance, the percentage of PV neurons in Field L or NCM that are encased by PNNs does not increase at the end of the sensory period for song learning in the zebra finch (Cornez et al. 2018). This suggests that the developmental expression of other molecular brakes in sensory processing areas needs to be studied, or that inhibitory circuitry and molecular brakes critically regulate the timing of the CP for song learning in sensorimotor circuits but not in sensory circuits.

2.3 Mechanisms Underlying the Sensorimotor Learning of Song

After juvenile songbirds memorize their tutor's song, they engage in a protracted period of sensorimotor learning (Fig. 2.1). This period is characterized by extensive vocal practice that is analogous to babbling in humans, when human infants begin their own attempts to match the sounds produced by others (Hahnloser and Kotowicz 2010; Murphy et al. 2017). A juvenile songbird's imitation of the memorized vocalizations is typically quite poor at the outset of practice. However, through feedback-dependent, trial-and-error learning, juveniles gradually become more proficient and eventually learn to precisely and accurately imitate those models (Tchernichovski et al. 2001; Okubo et al. 2015). Comparisons of motor actions and the resulting sensory feedback (*sensorimotor integration*) are essential for the development of accurate imitations of communicative sounds (Brainard and Doupe 2000); indeed, both songbirds and humans who lose their hearing at a very young age (i.e., at a time when they have been exposed to communicative sounds but have had very little time to rehearse their own performance of those sounds) are not capable of mastering the intricate details of communicative sounds (Westermann and Miranda 2004; Murphy et al. 2017). Because the acquisition of birdsong and speech entail similar types of sensorimotor learning, and because animal models are necessary for generating and testing mechanistic models of sensorimotor learning, songbirds have become a central model to reveal how the nervous system links sensory and motor information for sensorimotor development and learning (Tchernichovski and Marcus 2014; Nick 2015).

2.3.1 Neural Circuits for the Sensorimotor Learning of Song

Sensorimotor integration involves the comparison of actual sensory feedback to internal representations of the sensory consequences of motor commands (*sensory targets*) (Hickok et al. 2011; Murphy et al. 2017). There has been extensive speculation about and examination of how the brain compares sensory feedback, in particular auditory feedback, to sensory targets in order to compute *performance errors* for vocal production and to adaptively modify the vocal motor program to improve performance (i.e., minimize performance errors) (Troyer and Doupe 2000; Fee and Goldberg 2011). Given those speculations, research on the neural circuitry for sensorimotor learning has focused on identifying where and how *error signals* or *reinforcement signals* are computed and translated into adaptive motor plasticity. Most studies have focused on feedback or reinforcement signals that are internally generated (Sutton and Barto 1998), though a few studies have speculated on brain circuits that translate reinforcement signals produced by conspecifics (West and King 1988; Carouso-Peck and Goldstein 2019) into neural signals for sensorimotor plasticity (Hamilton et al. 1997; Benichov et al. 2016).

Because basal ganglia-cortical circuits are known to be important for sensorimotor plasticity in mammals, how neurons in the AFP contribute to the sensorimotor learning of song has been extensively examined (Brainard and Doupe 2013; reviewed in Leblois and Perkel, Chap. 4). For example, classic studies indicated that lesions of the basal ganglia nucleus Area X or the forebrain sensorimotor area LMAN in juvenile songbirds prevent the normal development of song (Bottjer et al. 1984; Scharff and Nottebohm 1991). Furthermore, neurons in Area X and LMAN gradually shift from being selectively responsive to the tutor's song to becoming tuned to the sound of the bird's own song as juveniles continue to practice their song (Doupe and Solis 1997; Achiro and Bottjer 2013). Similarly, neurons in LMAN-shell selectively respond to the tutor's song in juvenile songbirds and, moreover, change their singing-related firing depending on vocal performance (Achiro and Bottjer 2013; Achiro et al. 2017). Collectively, these and other studies highlighted the AFP as a major circuit underlying the sensorimotor learning of song and motivated investigations into brain areas that could modulate AFP function.

The VTA is an evolutionarily conserved brain area underlying behavioral plasticity, including sensorimotor learning (Fields et al. 2007; Morales and Margolis 2017), that sends DA projections into sensorimotor areas like Area X. Changes in VTA activity are correlated with changes in DA concentration in Area X (Yanagihara and Hessler 2006; Sasaki et al. 2006). Moreover, lesioning the DA neurons in the VTA that project to Area X prevented the development of accurate imitations of tutor song in juvenile zebra finches (Hisey et al., 2018). Therefore, it is important to understand how information about song performance could reach the VTA. Neurons in the VTA receive inputs from the ventral intermediate arcopallium (AIV) and dorsal arcopallium (Ad). Neurons in the AIV process auditory feedback during singing, and lesions of AIV or Ad neurons reduced the accuracy of developmental song learning (Bottjer and Altenau 2010; Mandelblat-Cerf et al. 2014). Neurons in Area X also indirectly project back to the VTA (via the ventral pallidum, VP) (Gale et al. 2008), indicating that the AFP itself can provide information about vocal performance to the VTA (Gale and Perkel 2010; Chen et al. 2019).

In addition to the AFP, neural circuits involving the sensorimotor areas HVC and NIf are important for developmental song learning. For example, a brain area within the secondary auditory region CM called Avalanche (Av; Fig. 2.2) receives input from the auditory thalamic nucleus uvaeformis (Uva) and is reciprocally connected to both HVC and NIf (Akutagawa and Konishi 2010). Importantly, genetic ablation of HVC neurons that project to Av precluded the sensorimotor learning of tutored songs across development (Roberts et al. 2017). In addition, a population of HVC neurons in adult swamp sparrows (*Melospiza georgiana*) encodes both the tutor's song and the bird's own song, thereby possibly serving as a bridge between the sensory representation of the target song and the bird's current performance (see Sect. 2.2.1 for discussion of the role of HVC in sensory learning) (Moseley et al. 2017). More of these *bridge cells* were observed in individuals who learned to produce more accurate imitations of their tutor songs, suggesting a contribution of this population to sensorimotor integration and song development.

In addition to studying the effects of neural manipulations on song development in juvenile songbirds, substantial insight into the neural circuits for sensorimotor plasticity has been gained by examining vocal plasticity in adult songbirds. Although the crux of sensorimotor learning occurs during development, plasticity in spectral and temporal features of song can be observed in adulthood (Nordeen and Nordeen 1992; Brainard and Doupe 2000). Like developing songbirds, adult songbirds require auditory feedback to maintain their stereotyped song (e.g., Leonardo and Konishi 1999; Sakata and Brainard 2006); the nervous system continues to monitor vocal performance, even after mastering song, to maintain a high level of performance (Konishi 2004; Hahnloser and Kotowicz 2010). Furthermore, adult songbirds can adaptively modify the acoustic structure of their song depending on reinforcement contingencies and sensory feedback (e.g., Sober and Brainard 2009; Hoffmann and Sober 2014). Because adult song maintenance and plasticity involve similar computations to developmental song learning, the neural mechanisms important for adult song maintenance and plasticity are also likely to be involved in developmental song learning (Hisey et al. 2018). Finally, experimentally driving sensorimotor plasticity is more tractable in adult songbirds than in developing birds because adults produce much more stereotyped vocalizations compared to juvenile songbirds, making it more feasible to experimentally target vocalizations for contingent reinforcement in adult songbirds.

A number of experimental paradigms are designed to identify neural populations for sensorimotor integration and learning in adult songbirds. For example, one paradigm randomly delivers localized perturbations of auditory feedback at targeted portions of song to simulate vocal errors. Pairing neurophysiological recordings with these feedback manipulations allows one to discover neural populations that monitor auditory feedback. Such experiments have revealed that neurons in the sensorimotor nucleus HVC in Bengalese finches (*Lonchura striata* var. *domestica*) (Sakata and Brainard 2008), and neurons in auditory cortical and thalamic areas in zebra finches (Keller and Hahnloser 2009; Lei and Mooney 2010) are sensitive to auditory feedback. In addition, VTA neurons that project to the basal ganglia nucleus Area X (VTA$_X$ neurons) demonstrate acute decreases in activity in response to unpredictable perturbations of feedback at targeted locations in song (errors) (Gadagkar et al. 2016). Interestingly, VTA$_X$ neurons demonstrate increases in activity on unpredictable control trials in which feedback was left unperturbed, a condition that could simulate correct vocal performance (Gadagkar et al. 2016). Such feedback-dependent changes in VTA activity are consistent with negative and positive "prediction error" signals found in mammals

(Hikosaka et al. 2008; Schultz 2017). These studies identify brain areas that encode information about auditory feedback and the quality of vocal performance and then propagate this information to neural populations that promote vocal motor plasticity.

A related paradigm has been used to identify brain circuits important for feedback-dependent song plasticity. This paradigm relies on similar computerized methods to detect targeted syllables in a bird's song but, instead of perturbing auditory feedback in a randomized manner, feedback perturbations are contingent upon a bird's vocal performance. Such contingent feedback perturbations are meant to simulate vocal motor errors and to provide reinforcement signals that can be used to guide adaptive vocal plasticity. Critical to these studies is the observation that birds will try to "escape" feedback perturbations by changing their song performance in adaptive ways. For example, white noise bursts can be played back to a finch when he produces a targeted syllable at a pitch lower than a specified threshold, and this reinforcement schedule would direct the nervous system to adapt song performance to produce the targeted syllable at a pitch higher than the threshold in order to avoid white noise playback. Tumer and Brainard (2007) were the first to use such a computerized system to drive adaptive song plasticity (Fig. 2.4a), and a number of subsequent studies using those methods have provided further insight into sensorimotor plasticity in adult Bengalese finches (e.g., Charlesworth et al. 2011; Warren et al. 2012) and zebra finches (e.g., Andalman and Fee 2009; Ali et al. 2013).

By coupling these experimental approaches with neural manipulations, researchers have begun to unravel the neural circuitry for sensorimotor plasticity. Consistent with its role in developmental song plasticity, LMAN is particularly important for adaptive song plasticity in adult songbirds (Brainard and Doupe 2013). For example, LMAN neurons promoted adaptive shifts in pitch in response to this reinforcement paradigm with contingent feedback perturbations (Andalman and Fee 2009; Warren et al. 2011) and allowed for the expression of such learning (Charlesworth et al. 2012). These data are consistent with other findings that LMAN neurons can bias the pitch of syllables on a moment-by-moment basis (Kao et al. 2005; Andalman and Fee 2009). Neurons in the basal ganglia nucleus Area X also contribute to some forms of sensorimotor plasticity in response to targeted perturbations (Ali et al. 2013). Interestingly, although LMAN and Area X contribute to feedback-dependent plasticity, studies to date have not documented auditory feedback signals in the AFP of songbirds like the zebra finch (e.g., Leonardo 2004; Kozhevnikov et al. 2007).

As mentioned previously, VTA_X neurons encode information about song performance that can be used to drive sensorimotor plasticity (Gadagkar et al. 2016), and a number of studies have revealed the functional significance of this population to vocal plasticity. In a first study, selective lesions of VTA_X neurons prevented some forms of feedback-dependent and reinforcement-dependent plasticity in adult Bengalese finches (Hoffmann et al. 2016). Subsequent studies supported and elegantly expanded on this initial finding and demonstrated that VTA_X neurons are important for adult song plasticity (Hisey et al. 2018; Xiao et al. 2018). These studies optogenetically stimulated or inhibited VTA_X neurons in a manner that was

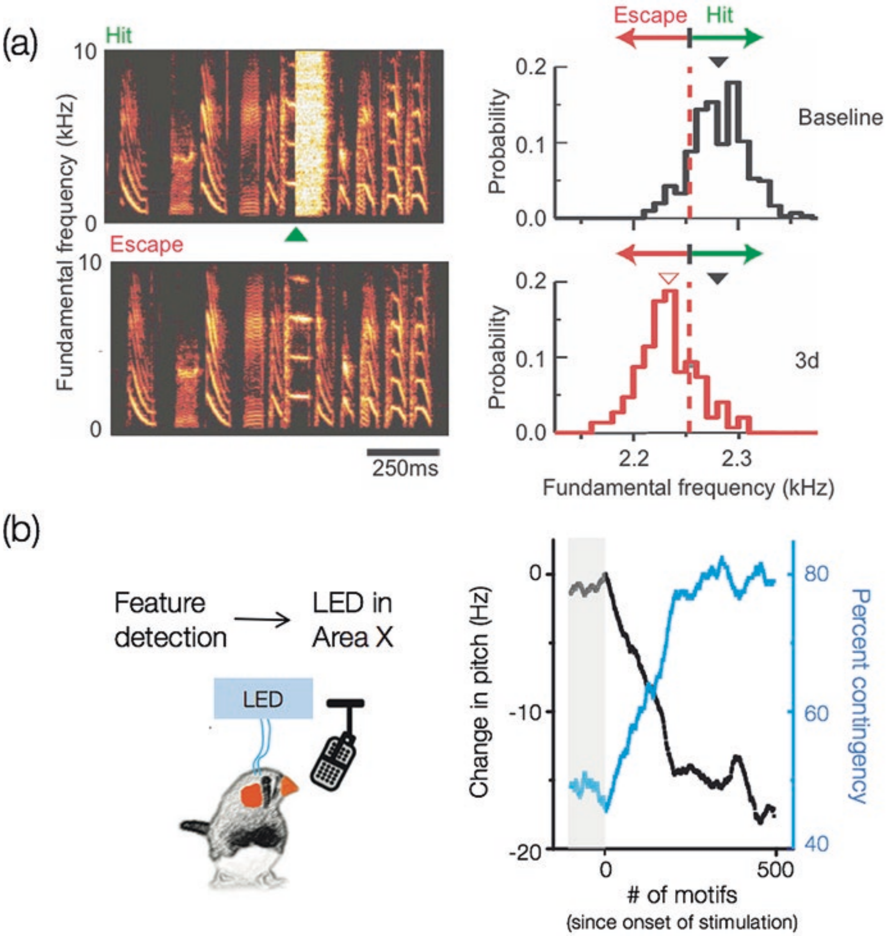


Fig. 2.4 Experimental approaches to study adult song plasticity in songbirds. **(a)** Operant approaches can be used to shift the fundamental frequency (FF) of syllables in the songs of an adult Bengalese finch (*Lonchura striata* var. *domestica*) (e.g., Tumer and Brainard 2007). *Right*: For this experiment, white noise (WN) was played back to a bird when the FF of the target syllable was above a threshold (*dashed line* in histograms; *Hit*); when the bird produces the target syllable at a FF below the threshold, white noise was avoided (*Escape*). *Left*: Spectrograms of a rendition of song with WN playback (*Hit*, top) and a rendition of song without WN playback (*Escape,* bottom). After three days (*3d*) of targeted WN feedback, the bird shifted the FF of the target syllable down such that it escaped WN feedback on the majority of trials (*lower right*). *Green triangle* below spectrogram indicates time of syllable detection; and *triangles* above histograms denote mean of distributions. **(b)** Optogenetic techniques have been used to identify populations of neurons that can drive adult song plasticity (e.g., Xiao et al. 2018). For example, in this experiment, dopaminergic neurons from the ventral tegmental area that project to the basal ganglia nucleus Area X were optically stimulated when the adult zebra finch produced a syllable with a FF below a particular threshold. Over time, the bird reduced the pitch of the syllable (*black line in right panel*) and increased the proportion of trials in which optogenetic stimulation of this population was triggered (percent contingency, *blue line in right panel*). (**a** adapted with permission from Tumer and Brainard 2007; **b** adapted with permission from Xiao et al. 2018)

contingent upon vocal performance and drove predictable and robust changes in song structure. For example, stimulating VTA_X neurons when the bird produces a syllable with a pitch above a specified threshold will lead to gradual increases in the pitch of that syllable (Fig. 2.4b) (Hisey et al. 2018; Xiao et al. 2018). Conversely, inhibiting VTA_X neurons when the bird produces a syllable with a pitch above a specified threshold will lead to gradual decreases in the pitch of that syllable (Xiao et al. 2018). These observations are consistent with studies of VTA_X contributions to developmental song learning (Hisey et al. 2018) and underscore the importance of dopaminergic circuitry to performance-based vocal plasticity in developing and adult songbirds.

2.3.2 Neural Mechanisms Underlying the Timing of the Critical Period for Sensorimotor Learning

The sensorimotor period is characterized by extensive vocal motor practice and plasticity during which an individual's song becomes more stereotyped and more similar to the song memorized during the sensory period (Fig. 2.1). Generally speaking, the sensorimotor period is thought to begin at the onset of vocal practice and end when song *crystallizes* or stabilizes (which tends to occur around the time of sexual maturation). However, relative to our understanding of the timing of the sensory period, the precise timing of the sensorimotor period for vocal learning is more elusive. In fact, the extent to which there is a CP for sensorimotor learning remains unclear in species like the zebra finch, since various forms of vocal motor plasticity have been observed in adults with crystallized song (Funabiki and Konishi 2003; Zevin et al. 2004). In this respect, the sensorimotor period is best considered as a period of heightened vocal motor plasticity and not as a restricted period for vocal motor plasticity.

Despite the difficulty in defining the timing of CPs for sensorimotor song learning, molecular mechanisms that promote or restrict sensorimotor plasticity could be similar to those regulating sensory plasticity (i.e., balance of excitation and inhibition, emergence of molecular brakes). Indeed, the percentage of PV neurons in HVC, RA, and Area X that are encased by PNNs is positively correlated with the extent of vocal maturity (Balmer et al. 2009; Cornez et al. 2018). Manipulations that shift the timing of vocal practice affect the encapsulation of PV neurons by PNNs (Balmer et al. 2009). Furthermore, the development of inhibitory circuitry in HVC is correlated with the maturation and imitation of song (Vallentin et al. 2016). Despite these insights, much remains to be explored regarding neural mechanisms shaping the timing and extent of sensorimotor plasticity (see Sect. 2.5).

2.4 Mechanisms Underlying Variation in Song Learning

The previous sections broadly introduced the neural circuits that regulate the sensory and sensorimotor learning of birdsong. In this section, two factors that modulate the extent and nature of vocal learning are discussed as well as the neural circuits that mediate such variations.

2.4.1 Mechanisms Underlying Social Influences on Vocal Learning

Social interactions between adults and developing individuals profoundly affect the rate, fidelity, and trajectory of vocal learning in a variety of species, including songbirds and humans (Snowdon and Hauserberg 1997; Kuhl 2007). For example, whereas passive exposure to speech or birdsong leads to minimal vocal learning in developing humans and songbirds, social and vocal interactions with adults lead to robust learning of communicative sounds (Kuhl et al. 2003; Chen et al. 2016). Indeed, circuits that encode social behavior have been proposed to gate the acquisition of speech in humans (Kuhl 2007). Despite the importance of social influences on vocal learning, relatively little is known about the neural mechanisms underlying the social enhancement and modulation of vocal learning.

Kuhl (2007) hypothesized that social interactions enhance the strength and fidelity of speech and language learning in humans by modulating attentional, sensory, and/or sensorimotor mechanisms. For example, visual and acoustic information provided by adults during interactions with developing individuals and reciprocal vocal interactions between adults and infants are hypothesized to enhance attention to and acquisition of speech sounds (West and King 1988; Goldstein et al. 2003). Consistent with this idea, social tutoring in songbirds increases behavioral indices of attention to song in juvenile songbirds, which in turn correlates with variation in song learning (Houx and Cate 1998; Chen et al. 2016).

While social interactions throughout development can influence the trajectory of song learning (West and King 1988; Carouso-Peck and Goldstein 2019), even social interactions limited to epochs of sensory learning enhance song learning (e.g., Chen et al. 2016). Consequently, it is likely that neurons that integrate sensory information with information about arousal, attention, and valence mediate social influences on vocal learning. Midbrain and hindbrain neurons that synthesize and release DA or norepinephrine (NE) are likely to serve as this integratory hub (O'Connell and Hofmann 2011; Sara and Bouret 2012). For example, across a wide range of species, DA neurons in the VTA and PAG and NE neurons in the LC receive sensory information from various sources and social information from hypothalamic and limbic forebrain areas. In addition, these DA and NE neurons regulate attention and arousal, and project to sensory and sensorimotor areas important for vocal learning (Castelino and Schmidt 2010). Consistent with a role of catecholaminergic neurons in the social modulation of song, social tutoring (but not passive tutoring) leads to a significant increase in the expression of the immediate early gene, EGR-1, in DA and NE neurons in the VTA and LC, respectively (Chen et al. 2016). Furthermore, while the effects of NE in juveniles is unknown, NE acting in the auditory area NCM enhances auditory responses to song and promotes the formation of song memories in adult zebra finches (Velho et al. 2012; Ikeda et al. 2015). These data suggest that DA and NE populations may be important for the sensory processing and learning of song (see also Nordeen et al. 2009; Tanaka et al. 2018).

While little is known about the extent to which catecholaminergic inputs into auditory processing areas mediate social influences on sensory learning, DA projections from the PAG to the sensorimotor area HVC are important for the social enhancement of song learning (Tanaka et al. 2018). In a series of elegant experiments, Tanaka et al. (2018) found that PAG neurons are differentially activated when juvenile zebra finches are socially tutored and that more DA from PAG neurons is released into HVC when birds are socially tutored. In addition, they reported that blocking DA receptors in HVC inhibits song learning in response to social tutoring. Those experiments provide compelling support for the role of PAG neurons in the sensory learning of song. However, given that the sensory learning of song seems to involve a distributed network of brain areas (see Sect. 2.2), it remains possible that catecholamines acting in other areas also contribute to the social modulation of sensory learning.

In addition to influencing the sensory learning of song, social interactions also affect the sensorimotor development of learned vocalizations. Cowbirds provide a striking example of the role of feedback signals in the sensorimotor development of song. Adult female cowbirds produce “wing strokes,” a copulation solicitation display, in response to hearing juvenile vocalizations that they prefer (West et al. 1981). Juveniles accordingly shape their vocal repertoire in response to those cues such that they retain acoustic elements that are preferred by females and prune away elements that are not preferred by females (West and King 1988). Similarly, social cues by adult females also affect song learning in juvenile male zebra finches (Carouso-Peck and Goldstein 2019), and social feedback signals shape the development of human infant vocalizations (Goldstein and Schwade 2008; Lytle and Kuhl 2017). Given that VTA neurons are sensitive to social stimuli and that VTA activity shapes vocal motor plasticity (see Sect. 2.3.1), it is plausible that DA neurons in the VTA or PAG that project to sensorimotor brain areas (e.g., Area X, HVC) encode social reinforcement signals and mediate this type of plasticity.

2.4.2 Mechanisms Underlying Biological Predispositions in Learning

Decades of research have revealed that juvenile songbirds are not blank slates but are born with specific biological predispositions for vocal learning (Marler 1997; Catchpole and Slater 2008). One of the most striking and consistent findings is that juvenile songbirds are biased to learn conspecific song over heterospecific song. Evidence for this bias comes from a number of discoveries: (1) juveniles incorporate more conspecific elements into their adult song when they are tutored with both conspecific and heterospecific songs, (2) conspecific song learning is very rapid and requires minimal song exposure, and (3) juvenile songbirds display more attention and arousal-like behaviors in response to conspecific song than heterospecific song (Marler 1997; Wheatcroft and Qvarnström 2015). In addition, particular types of species-typical song syllables are observed more frequently across populations, and when song-naïve juvenile zebra finches are tutored with conspecific syllables that are frequently versus infrequently observed across populations, they preferentially learn to produce syllables that are frequently observed in a population (ter Haar et al. 2014). Taken together, these experiments reveal biological predispositions to learn to produce particular acoustic elements.

Juvenile songbirds also transform and adapt species-atypical song inputs to make them more species typical (i.e., *normalization* of song). For example, zebra finches raised without exposure to song throughout development (*isolates*) produce songs with a number of species-atypical features, and juveniles that are tutored by isolates do not produce an exact match of the isolate tutor's song; instead, they modify the isolate's song in ways to make the song more species-typical (Fehér et al. 2009). Isolates produce songs with atypically long syllables, and zebra finches tutored by isolates produce truncated versions of the isolate tutor's syllables, causing syllable durations to be more consistent with species-typical durations. Intriguingly, when such tutoring is repeated across generations, a song emerges that resembles that of normally tutored zebra finches (Fehér et al. 2009).

Species not only vary in the acoustic structure of their syllables but also in the temporal organization of their songs (*syllable timing* and *sequencing*), and a number of studies have revealed how innate processes modulate the learning of temporal aspects of song. For instance, while song sparrows do not readily learn swamp sparrow syllables when tutored with normal swamp sparrow song, song sparrows were more likely to imitate swamp sparrow syllables if those syllables were presented in temporal patterns that resembled species-typical song sparrow songs (Marler 1997). Consistent with this, genetic (interfamily) variation in song timing is observed in Bengalese finches, and juveniles will more accurately imitate the tempo of a tutor's song if it is presented at a tempo that corresponds to the tempo of their biological father (Mets and Brainard 2018; Mets and Brainard 2019).

Additionally, a number of studies have found that birds will transform the timing and sequencing of learned song syllables in accordance with genetic predispositions. For example, zebra finches and Bengalese finches produce songs that differ in tempo, and zebra finches tutored with Bengalese finch song normalize the timing of Bengalese finch syllables into a more zebra finch-like pattern (e.g., Clayton 1989; Araki et al. 2016). The songs of canaries (*Serinus canaria*) are characterized by a series of phrases that each contain a repetition of a single syllable (i.e., syllable repetition represents a species-typical syntax), and when canaries are tutored with a random sequence of nonrepeating syllables (a species-atypical pattern), they reorganize these learned syllables into a series of phrases of repeated syllables (Gardner et al. 2005). Similarly, the songs of white-crowned sparrows (*Zonotrichia leucophrys*) are characterized by a stereotyped sequence of syllables that is remarkably consistent across individuals and populations, and juvenile white-crowned sparrows that are tutored with species-atypical sequences learn to produce sequences with features that are typical of their species (Plamondon et al. 2010).

Biological predispositions in sequence learning have also been documented in species without rigid species-typical syntax. While zebra finch songs are characterized by a single, stereotyped sequence of syllables (*motif*), there do not seem to be constraints on how syllables are sequenced within the motif (e.g., Lipkind et al. 2013). However, surveys of wild and laboratory populations of zebra finches have revealed that some patterns of syllable sequences are more prevalent than others, suggesting the possibility of biological predispositions in vocal sequence learning (Zann 1993; Lachlan et al. 2016). In order to reveal biological predispositions in sequence learning, naïve juvenile zebra finches were tutored with synthesized songs that consisted of five species-typical syllables that were arranged in every possible five-syllable sequence variant ($n = 120$ variants; Fig. 2.5) (James and Sakata 2017). Importantly, each bird in this study heard the 120 sequence variants in equal proportions and in randomized orders. Despite being tutored with randomized sequences in which each of the five syllables was equally likely to occupy a particular position in the motif, tutored birds preferentially produced particular syllables at distinct positions in their motif. For example, there was a significant bias for birds to produce syllables that resembled distance calls at the end of their motif and to produce short, high-pitched syllables in the middle of their motifs. Interestingly, birds tutored with randomized sequences positioned syllables into patterns that are commonly observed in wild populations of zebra finches, indicating that these birds normalized their song output when tutored with randomized sequences of song syllables (James and Sakata 2017).

Such learning biases could reflect perceptual (auditory) and/or motor biases and, to date, evidence for both mechanisms exists. While behavioral responses of developing birds support the notion of perceptual biases (Nelson 2000; ter Haar et al. 2014), there are only a handful of studies directly investigating innate auditory biases or *filters* (but see Moore and Woolley 2019). For example, some neurons in and around HVC will respond preferentially to conspecific song in song-naive juvenile zebra finches (Adret et al. 2012). Auditory neurons in zebra finches are also preferentially tuned to the temporal organization of conspecific song. As alluded to previously, the songs of zebra finches contain shorter silent gaps between syllables than the songs of Bengalese finches, and a population of neurons in the L3 region of the avian auditory cortex (Field L) of zebra finches were more responsive to sounds

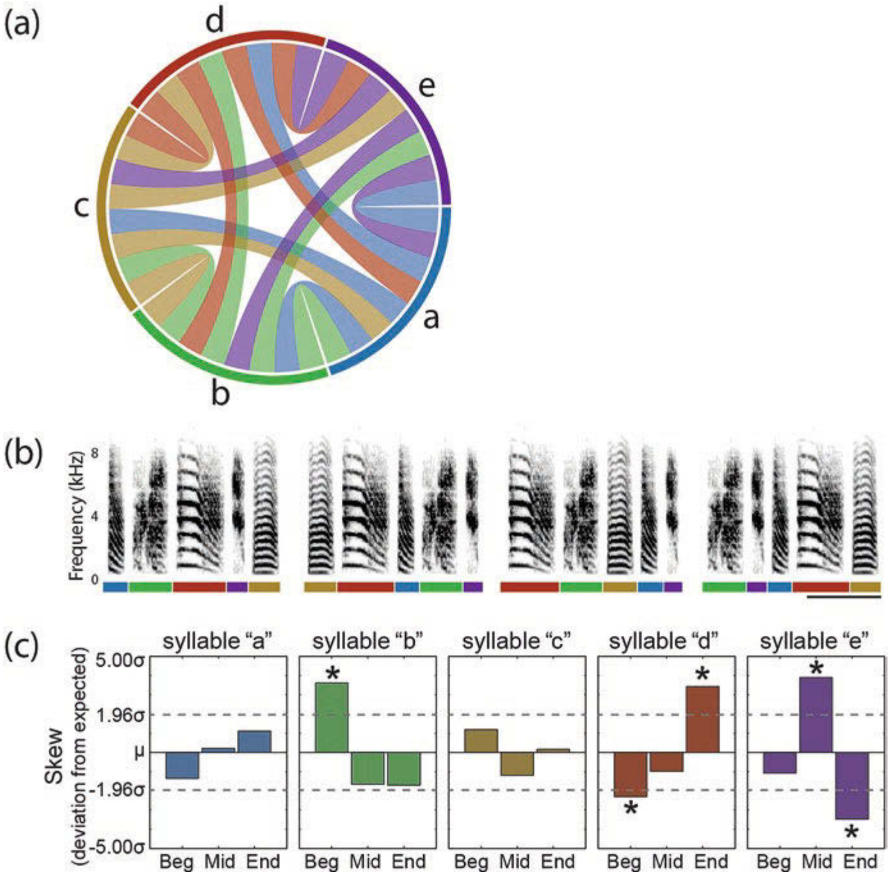


Fig. 2.5 Biological predispositions for vocal sequence learning in the zebra finch (*Taeniopygia guttata*). **(a)** Tutoring naïve zebra finches with randomized sequences of syllables reveals predispositions to learn to produce particular sequences of syllables. The *circle plot* summarizes transitions between syllables within the 120 different motif variants of five canonical zebra finch syllables (*"a"* – *"e"*), with the width of the line being proportional to the transition probabilities between syllables [*colors* correspond to those used for syllables in (b) and (c)]. Key aspects of this stimulus set are that transition probabilities between all syllables are equal and that motifs are equally likely to start and end with any of the five syllables. **(b)** Example of a song stimulus played back during song tutoring. The stimulus consists of four motifs, each composed of the same syllables (color coded as in [c]) but in a different order. **(c)** The likelihood of birds producing a particular syllable at particular positions in the motif (beginning, *Beg*; middle, *Mid*; end, *End*) was not random. For example, syllable "b" was more likely than chance to be produced at the beginning of the motif, syllable "e" was more likely than chance to be produced in the middle of the motif, and syllable "d" was more likely than chance to be produced at the end of the motif; asterisk indicates $p < 0.05$. (adapted with permission from James and Sakata 2017)

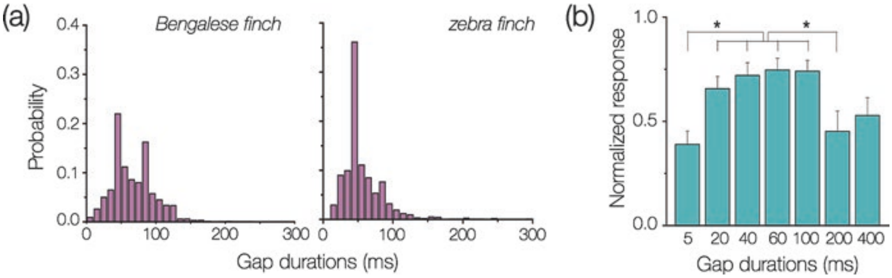


Fig. 2.6 Neurons in Field L3 are tuned to species-typical temporal aspects of song. **(a)** Gap durations within song differ between Bengalese finches (*Lonchura striata* var. *domestica*) and zebra finches (*Taeniopygia guttata*). In particular, one is more likely to observe relatively long gaps (e.g., >100 ms) in Bengalese finch song than in zebra finch song. **(b)** Neurons in Field L3 of zebra finches are more active in response to stimuli that contain species-typical gap durations (<100 ms) than species-atypical gap durations (>100 ms); asterisk indicates $p < 0.05$. (adapted with permission from Araki et al. 2016)

that were separated by species-typical gaps than species-atypical gaps (Fig. 2.6) (Araki et al. 2016). This preference was observed even in juvenile zebra finches that had not been exposed to conspecific song during the CP for song learning, suggesting an innate tuning to the temporal pattern of syllables (Araki et al. 2016). Together, these studies suggest that innate perceptual filters could contribute to species biases in song learning and to the normalization of song in response to aberrant song inputs (but see Whaling et al. 1997).

In addition to perceptual biases, a number of studies support a contribution of motor biases to biological predispositions in song learning and production (see Podos and Sung, Chap. 9). Because isolates develop their songs without memorizing the song of an adult conspecific, the songs of isolates have been proposed to reflect motor biases in song production (Marler 1997). Indeed, a number of species differences that are observed in tutored birds are also observed among isolates of those species (Marler 1997), and the vocal differences complement morphological differences in the organization of the vocal periphery (Riede and Goller 2014; Düring and Elemans 2016). Also, interfamily variation in song tempo can be observed in the vocalizations of birds *before* tutoring (Sato et al. 2016), supporting the notion that innate biases in syllable timing could shape song learning (Mets and Brainard 2018).

The process of song normalization in birds tutored with species-atypical songs could also support the role of motor biases in song learning and production. Canaries that were tutored with species-atypical syntax (Gardner et al. 2005) and chaffinches (*Fringilla coelebs*), song sparrows (*Melospiza melodia*), and swamp sparrows that were tutored with both heterospecific and conspecific songs (Thorpe 2008; Marler 1997) produced imitations of atypical or heterospecific songs during development. However, when the birds reached sexual maturity, they stopped producing species-atypical vocalizations but continued to produce species-typical sounds and patterns. Because testicular hormones surge at the time of sexual maturation, androgens are hypothesized to be important for the normalization of species-atypical vocalizations into species-typical patterns. Consistent with this role, administration of androgens to canaries that produced species-atypical vocal patterns resulted in the normalization of vocal patterns (Gardner et al. 2005; Alliende et al. 2010). Importantly, because the vocal motor changes occurred after the CP for sensory learning had closed and because muscles for vocal control and neurons in the song system are replete with androgen receptors, these androgen-induced song modifications may reflect changes to the vocal motor system. However, more direct manipulations of the vocal motor system are necessary to test this notion.

In summary, there exist numerous predispositions for the acquisition of spectral and temporal features of song, and it is likely that both sensory and motor biases contribute to such biological predispositions. Given the reciprocal influences of sensory and motor systems during development, disentangling these influences under normal physiological conditions is difficult. Nevertheless, because biological predispositions in learning are prevalent across taxa and across types of behaviors, revealing mechanisms of vocal-learning biases could provide insights into the general mechanisms of behavioral plasticity (e.g., Wang et al. 2019).

2.5 New Perspectives and Directions for the Study of Vocal Learning in Songbirds

This chapter provides a broad overview of general mechanisms underlying song learning in songbirds with an emphasis on mechanisms underlying sensory learning and sensorimotor learning. Given the scope and nature of this chapter, a number of other important aspects of song learning are not covered, including the lateralization of brain function (Bolhuis and Moorman 2015), the role of sleep in vocal consolidation (Margoliash and Schmidt 2010; Bolhuis and Moorman 2015), and the importance of variability to song plasticity (Brainard and Doupe 2013; Dhawale et al. 2017). However, it is useful to highlight some avenues of research that are considered particularly important, exciting, and fruitful.

The chapter extensively discusses the role of DA and NE in sensory and sensorimotor learning. Studies linking reinforcement and attentional circuits with sensory and sensorimotor circuits have been groundbreaking and insightful (e.g., Hisey et al. 2018; Xiao et al. 2018). To further our understanding of the neuroethology of birdsong, it will be important to reveal how these neuromodulatory populations contribute to species biases and constraints in song learning (see Sect. 2.4.2). For example, just as auditory neurons in Field L are tuned to temporal properties of species-typical song (Araki et al. 2016), it is possible that DA and NE neurons are innately tuned to species-typical features of song. In addition, it will be important to reveal how these neurons encode social stimuli that shape song learning (Chen et al. 2016; Tanaka et al. 2018).

Although the notion of CPs pervades discussions of song learning, much remains to be explored regarding the mechanisms of CP timing for song learning. For example, molecular brakes, such as nogo or LYNX1, are implicated in restricting visual plasticity, but their roles in closing the CP for sensory learning of birdsong has yet to be explored. In addition, while much is known about how the timing of the CP for sensory learning can be modified by experiences, little is known about how experiences regulate the timing of the sensorimotor period. Indeed, manipulations of sensory experiences can shape sensorimotor plasticity (Funabiki and Konishi 2003; Zevin 2004) and can be used to identify mechanisms that regulate developmental changes in sensorimotor plasticity (Tachibana et al. 2017; Hayase et al. 2018). As alluded to in Sect. 2.3, the extent to which the sensorimotor period should be considered a CP requires additional attention. While it is evident that songbirds like the zebra finch cannot learn to produce a new songs as adults, their songs change following sensory perturbations (e.g., deafening) and adaptively shift in response to targeted feedback and reinforcement; in other words, sensorimotor plasticity persists following the end of the developmental period for sensorimotor learning. Finally, most studies tend to envision common mechanisms regulating the opening and closing of the sensory and sensorimotor periods, but future studies should aim to dissect and differentiate mechanisms underlying developmental changes in sensory and sensorimotor learning (Yazaki-Sugiyama 2019).

A related question is how auditory learning for vocal learning mechanistically differs from auditory learning that is not involved in vocal learning (George et al. 1995; Jin and Clayton 1997). The most extensively studied songbirds are *close-ended learners* that learn their songs during a restricted period in development. While it is evident that both juvenile and adult songbirds engage in auditory learning (e.g., learning the songs of other males) and that this auditory learning is an integral part of social behavior in songbirds (Bradbury and Vehrencamp 2011; Beecher 2017), little is known about why auditory learning during development translates into changes in vocal production, whereas auditory learning in adulthood has minimal impact on a bird's song. In addition, while both males and females engage in auditory learning, auditory learning leads to vocal changes only in males in a number of songbird species, and relatively little is known about sex differences in the translation of auditory learning to vocal learning (but see London, Chap. 8). This finding is particularly interesting from an evolutionary perspective because the ancestral condition appears to be song production in both males and females (Odom et al. 2014). Understanding how sensory learning becomes dissociated from sensorimotor learning is important for understanding age-limited and sex-limited vocal learning. One possibility is that developmental changes to and sex differences in the functional connectivity and plasticity between auditory areas (like NCM) and sensorimotor areas (like HVC) could underlie this change. Understanding the nature of these neural interactions could also provide insight into mechanisms of speech acquisition, since the interaction between NCM and HVC are proposed to be reminiscent of the interactions between Wernicke's and Broca's areas in humans (e.g., Bolhuis and Moorman 2015).

Adopting a comparative approach can be powerful for revealing mechanisms of CP regulation. Most studies of neural mechanisms are focused on species that learn their songs during a restricted period in development. However, a number of songbirds, like humans, are able to acquire new vocalizations throughout their lives (Murphy et al. 2017). While understanding this persistence of vocal learning and plasticity has been emphasized for decades, little is known about such mechanisms. In support of this contention, species variation in the expression of molecular brakes covaries with species variation in adult song learning (Cornez et al. 2017). Continuing such lines of inquiry are important not just from a neuroethological perspective but also for translational reasons (e.g., enhancing language learning in adults).

Finally, an overarching question throughout this review is the extent to which sensory learning and sensorimotor learning rely on shared versus distinct circuits. Clearly, both sensory (auditory) and sensorimotor brain areas are important for the sensory learning of birdsong, and sensorimotor circuits are critical for the sensorimotor learning of birdsong. However, relatively little is known about the degree to which auditory processing areas like NCM influence the sensorimotor development of song. Some studies in adult songbirds indicate that NCM is important for some aspects of feedback-based and reinforcement-based song plasticity (Canopoli et al. 2014), but little is known about the contribution of NCM or other auditory areas to developmental improvements in song performance (but see Canopoli et al. 2016).

2.6 Chapter Summary

Songbirds are ideal species to reveal the neural mechanisms underlying vocal learning. Decades of research have revealed how sensory, sensorimotor, and reinforcement mechanisms contribute to sensory and sensorimotor aspects of song learning in songbirds. Given the numerous parallels in vocal development and plasticity in songbirds and humans and the accumulation of evidence for deep homologies in brain organization between songbirds and humans, songbirds can continue to provide important insights into potential mechanisms underlying speech acquisition in humans.

Acknowledgements We thank Mimi Kao for helpful comments on a previous version of the manuscrFipt.

Compliance With Ethics Requirements Jon T. Sakata declares that he has no conflict of interest. Yoko Yazaki-Sugiyama declares that she has no conflict of interest.

References

Achiro JM, Bottjer SW (2013) Neural representation of a target auditory memory in a cortico-basal ganglia pathway. J Neurosci 33:14475–14488. https://doi.org/10.1523/JNEUROSCI.0710-13.2013

Achiro JM, Shen J, Bottjer SW (2017) Neural activity in cortico-basal ganglia circuits of juvenile songbirds encodes performance during goal-directed learning. eLife 6:e26973. https://doi.org/10.7554/eLife.26973

Adret P, Meliza CD, Margoliash D (2012) Song tutoring in presinging zebra finch juveniles biases a small population of higher-order song-selective neurons toward the tutor song. J Neurophysiol 108:1977–1987. https://doi.org/10.1152/jn.00905.2011

Akutagawa E, Konishi M (2010) New brain pathways found in the vocal control system of a songbird. J Comp Neurol 518:3086–3100. https://doi.org/10.1002/cne.22383

Ali F, Otchy TM, Pehlevan C, Fantana AL, Burak Y, Ölveczky BP (2013) The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. Neuron 80:494–506. https://doi.org/10.1016/j.neuron.2013.07.049

Alliende JA, Méndez JM, Goller F, Mindlin GB (2010) Hormonal acceleration of song development illuminates motor control mechanism in canaries. Dev Neurobiol 70:943–960. https://doi.org/10.1002/dneu.20835

Andalman AS, Fee MS (2009) A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. PNAS USA 106:12518–12523. https://doi.org/10.1073/pnas.0903214106

Araki M, Bandi MM, Yazaki-Sugiyama Y (2016) Mind the gap: neural coding of species identity in birdsong prosody. Science 354:1282–1287. https://doi.org/10.1126/science.aah6799

Aronov D, Andalman AS, Fee MS (2008) A specialized forebrain circuit for vocal babbling in the juvenile songbird. Science 320:630–634. https://doi.org/10.1126/science.1155140

Balmer TS, Carels VM, Frisch JL, Nick TA (2009) Modulation of perineuronal nets and parvalbumin with developmental song learning. J Neurosci 29:12878–12885. https://doi.org/10.1523/JNEUROSCI.2974-09.2009

Basham ME, Nordeen EJ, Nordeen KW (1996) Blockade of nmda receptors in the anterior forebrain impairs sensory acquisition in the zebra finch (*Poephila guttata*). Neurobiol Learn Mem 66:295–304. https://doi.org/10.1006/nlme.1996.0071

Bateson PPG (1966) The characteristics and context of imprinting. Biol Rev 41:177–217. https://doi.org/10.1111/j.1469-185X.1966.tb01489.x

Beecher MD (2017) Birdsong learning as a social process. Anim Behav 124:233–246. https://doi.org/10.1016/j.anbehav.2016.09.001

Benichov JI, Benezra SE, Vallentin D, Globerson E, Long MA, Tchernichovski O (2016) The forebrain song system mediates predictive call timing in female and male zebra finches. Curr Biol 26:309–318. https://doi.org/10.1016/j.cub.2015.12.037

Bolhuis JJ, Gahr M (2006) Neural mechanisms of birdsong memory. Nat Rev Neurosci 7:347–357. https://doi.org/10.1038/nrn1904

Bolhuis JJ, Moorman S (2015) Birdsong memory and the brain: in search of the template. Neurosci Biobehav Rev 50:41–55. https://doi.org/10.1016/j.neubiorev.2014.11.019

Bottjer SW, Alderete TL, Chang D (2010) Conjunction of vocal production and perception regulates expression of the immediate early gene ZENK in a novel cortical region of songbirds. J Neurophysiol 103(4):1833–1842. https://doi.org/10.1152/jn.00869.2009

Bottjer SW, Altenau B (2010) Parallel pathways for vocal learning in basal ganglia of songbirds. Nature Neurosci 13:153–155. https://doi.org/10.1038/nn.2472

Bottjer SW, Miesner EA, Arnold AP (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. Science 224:901–903. https://doi.org/10.1126/science.6719123

Bradbury JW, Vehrencamp SL (2011) Principles of animal communication, 2nd edn. Sinauer Associates, Sunderland

Brainard MS, Doupe AJ (2000) Auditory feedback in learning and maintenance of vocal behaviour. Nat Rev Neurosci 1:31–40. https://doi.org/10.1038/35036205

Brainard MS, Doupe AJ (2002) What songbirds teach us about learning. Nature 417:351–358. https://doi.org/10.1038/417351a

Brainard MS, Doupe AJ (2013) Translating birdsong: songbirds as a model for basic and applied medical research. Annu Rev Neurosci 36:489–517. https://doi.org/10.1146/annurev-neuro-060909-15282

Canopoli A, Herbst JA, Hahnloser RHR (2014) A higher sensory brain region is involved in reversing reinforcement-induced vocal changes in a songbird. J Neurosci 34:7018–7026. https://doi.org/10.1523/JNEUROSCI.0266-14.2014

Canopoli A, Zai A, Hahnloser R (2016) Lesions of a higher auditory brain area during a sensorimotor period do not impair birdsong learning. Matters (Zürich). https://doi.org/10.19185/matters.201603000018

Carouso-Peck S, Goldstein MH (2019) Female social feedback reveals non-imitative mechanisms of vocal learning in zebra finches. Curr Biol 29:631–636. https://doi.org/10.1016/j.cub.2018.12.026

Castelino CB, Schmidt MF (2010) What birdsong can teach us about the central noradrenergic system. J Chem Neuroanat 39:96–111. https://doi.org/10.1016/j.jchemneu.2009.08.003

Catchpole CK, Slater PJB (2008) Bird song: biological themes and variations. Cambridge University Press

Charlesworth JD, Tumer EC, Warren TL, Brainard MS (2011) Learning the microstructure of successful behavior. Nature Neurosci 14:373–380. https://doi.org/10.1038/nn.2748

Charlesworth JD, Warren TL, Brainard MS (2012) Covert skill learning in a cortical-basal ganglia circuit. Nature 486:251–255. https://doi.org/10.1038/nature11078

Chen Y, Matheson LE, Sakata JT (2016) Mechanisms underlying the social enhancement of vocal learning in songbirds. PNAS USA 113:6641–6646. https://doi.org/10.1073/pnas.1522306113

Chen R, Puzerey PA, Roeser AC, Riccelli TE, Podury A, Maher K, Farhang A, Goldberg J (2019) Songbird ventral basal ganglia sends performance error signals to dopaminergic midbrain. Neuron 103:266–276. https://doi.org/10.1016/j.neuron.2019.04.038

Clayton NS (1989) The effects of cross-fostering on selective song learning in estrildid finches. Behav 109:163–175. https://doi.org/10.1163/156853989X00204

Cornez G, ter Haar SM, Cornil CA, Balthazart J (2015) Anatomically discrete sex differences in neuroplasticity in zebra finches as reflected by perineuronal nets. PLoS One 10:e0123199. https://doi.org/10.1371/journal.pone.0123199

Cornez G, Madison FN, Van A d L, Cornil C, Yoder KM, Ball GF, Balthazart J (2017) Perineuronal nets and vocal plasticity in songbirds: a proposed mechanism to explain the difference between closed-ended and open-ended learning. Dev Neurobiol 77:975–994. https://doi.org/10.1002/dneu.22485

Cornez G, Jonckers E, ter Haar SM, Van der Linden A, Cornil CA, Balthazart J (2018) Timing of perineuronal net development in the zebra finch song control system correlates with developmental song learning. Proc Royal Soc B Biol Sci 285:20180849. https://doi.org/10.1098/rspb.2018.0849

de Villers-Sidani E, Chang EF, Bao S, Merzenich MM (2007) Critical period window for spectral tuning defined in the primary auditory cortex (a1) in the rat. J Neurosci 27:180–189. https://doi.org/10.1523/JNEUROSCI.3227-06.2007

Dhawale AK, Smith MA, Ölveczky BP (2017) The role of variability in motor learning. Ann Rev Neurosci 40:479–498. https://doi.org/10.1146/annurev-neuro-072116-031548

Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. Ann Rev Neurosci 22:567–631. https://doi.org/10.1146/annurev.neuro.22.1.567

Doupe AJ, Solis MM (1997) Song- and order-selective neurons develop in the songbird anterior forebrain during vocal learning. J Neurobiol 33:694–709. https://doi.org/10.1002/(SICI)1097-4695(19971105)33:5<694::AID-NEU13>3.0.CO;2-9

Doupe AJ, Perkel DJ, Reiner A, Stern EA (2005) Birdbrains could teach basal ganglia research a new song. TINS 28:353–363. https://doi.org/10.1016/j.tins.2005.05.005

Doya K, Sejnowski TJ (1995) A novel reinforcement model of birdsong vocalization learning. In: Advances in neural information processing. MIT Press, pp 101–108

Düring DN, Elemans CP (2016) Embodied motor control of avian vocal production. In Vertebrate sound production and acoustic communication. Springer, Cham. pp 119–157

Fagiolini M, Hensch TK (2000) Inhibitory threshold for critical-period activation in primary visual cortex. Nature 404:183–186. https://doi.org/10.1038/35004582

Fee MS, Goldberg JH (2011) A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. Neurosci 198:152–170. https://doi.org/10.1016/j.neuroscience.2011.09.069

Fehér O, Wang H, Saar S, Mitra PP, Tchernichovski O (2009) De novo establishment of wild-type song culture in the zebra finch. Nature 459:564–568. https://doi.org/10.1038/nature07994

Fields HL, Hjelmstad GO, Margolis EB, Nicola SM (2007) Ventral tegmental area neurons in learned appetitive behavior and positive reinforcement. Ann Rev Neurosci 30:289–316. https://doi.org/10.1146/annurev.neuro.30.051606.094341

Funabiki Y, Konishi M (2003) Long memory in song learning by zebra finches. J Neurosci 23:6928–6935. https://doi.org/10.1523/JNEUROSCI.23-17-06928.2003

Gadagkar V, Puzerey PA, Chen R, Baird-Daniel E, Farhang AR, Goldberg JH (2016) Dopamine neurons encode performance error in singing birds. Science 354:1278–1282. https://doi.org/10.1126/science.aah6837

Gahr M, Metzdorf R (1997) Distribution and dynamics in the expression of androgen and estrogen receptors in vocal control systems of songbirds. Brain Res Bull 44:509–517. https://doi.org/10.1016/S0361-9230(97)00233-5

Gale SD, Perkel DJ (2010) A basal ganglia pathway drives selective auditory responses in songbird dopaminergic neurons via disinhibition. J Neurosci 30:1027–1037. https://doi.org/10.1523/JNEUROSCI.3585-09.2010

Gale SD, Person AL, Perkel DJ (2008) A novel basal ganglia pathway forms a loop linking a vocal learning circuit with its dopaminergic input. J Comp Neurol 508:824–839. https://doi.org/10.1002/cne.21700

Gardner TJ, Naef F, Nottebohm F (2005) Freedom and rules: the acquisition and reprogramming of a bird's learned song. Science 308:1046–1049. https://doi.org/10.1126/science.1108214

George JM, Jin H, Woods WS, Clayton DF (1995) Characterization of a novel protein regulated during the critical period for song learning in the zebra finch. Neuron 15:361–372. https://doi.org/10.1016/0896-6273(95)90040-3

Gobes SMH, Bolhuis JJ (2007) Birdsong memory: a neural dissociation between song recognition and production. Curr Biol 17:789–793. https://doi.org/10.1016/j.cub.2007.03.059

Gobes SMH, Jennings RB, Maeda RK (2017) The sensitive period for auditory-vocal learning in the zebra finch: consequences of limited-model availability and multiple-tutor paradigms on song imitation. Behav Process 163:5–12. https://doi.org/10.1016/j.beproc.2017.07.007

Goldstein MH, Schwade JA (2008) Social feedback to infants' babbling facilitates rapid phonological learning. Psychol Sci 19:515–523. https://doi.org/10.1111/j.1467-9280.2008.02117.x

Goldstein MH, King AP, West MJ (2003) Social interaction shapes babbling: testing parallels between birdsong and speech. PNAS USA 100:8030–8035. https://doi.org/10.1073/pnas.1332441100

Hahnloser RH, Kotowicz A (2010) Auditory representations and memory in birdsong learning. Curr Opin Neurobiol 20:332–339. https://doi.org/10.1016/j.conb.2010.02.011

Hamilton KS, King AP, Sengelaub DR, West MJ (1997) A brain of her own: a neural correlate of song assessment in a female songbird. Neurobiol Learn Mem 68:325–332. https://doi.org/10.1006/nlme.1997.3781

Happel MFK, Niekisch H, Castiblanco Rivera LL, Ohl FW, Deliano M, Frischknecht R (2014) Enhanced cognitive flexibility in reversal learning induced by removal of the extracellular matrix in auditory cortex. PNAS USA 111:2800–2805. https://doi.org/10.1073/pnas.1310272111

Hayase S, Wang H, Ohgushi E, Kobayashi M, Mori C, Horita H, Mineta K, Liu W, Wada K (2018) Vocal practice regulates singing activity–dependent genes underlying age-independent vocal learning in songbirds. PLoS Biol 16:e2006537. https://doi.org/10.1371/journal.pbio.2006537

Heinrich JE, Singh TD, Sohrabji F, Nordeen KW, Nordeen EJ (2002) Developmental and hormonal regulation of NR2A mRNA in forebrain regions controlling avian vocal learning. J Neurobiol 51:149–159. https://doi.org/10.1002/neu.10046

Heinrich J, Nordeen K, Nordeen E (2005) Dissociation between extension of the sensitive period for avian vocal learning and dendritic spine loss in the song nucleus LMAN. Neurobiol Learn Mem 83:143–150. https://doi.org/10.1016/j.nlm.2004.11.002

Hensch TK (2004) Critical period regulation. Ann Rev Neurosci 27:549–579. https://doi.org/10.1146/annurev.neuro.27.070203.144327

Hickok G, Houde J, Rong F (2011) Sensorimotor integration in speech processing: computational basis and neural organization. Neuron 69:407–422. https://doi.org/10.1016/j.neuron.2011.01.019

Hikosaka O, Bromberg-Martin E, Hong S, Matsumoto M (2008) New insights on the subcortical representation of reward. Curr Opin Neurobiol 18:203–208. https://doi.org/10.1016/j.conb.2008.07.002

Hisey E, Kearney MG, Mooney R (2018) A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. Nature Neurosci 21:589–597. https://doi.org/10.1038/s41593-018-0092-6

Hoffmann LA, Sober SJ (2014) Vocal generalization depends on gesture identity and sequence. J Neurosci 34:5564–5574. https://doi.org/10.1523/JNEUROSCI.5169-13.2014

Hoffmann LA, Saravanan V, Wood AN, He L, Sober SJ (2016) Dopaminergic contributions to vocal learning. J Neurosci 36:2176–2189. https://doi.org/10.1523/JNEUROSCI.3883-15.2016

Horn G (2004) Pathways of the past: the imprint of memory. Nat Rev Neurosci 5:108–120

Houx BB, Cate CT (1998) Do contingencies with tutor behaviour influence song learning in zebra finches? Behaviour 135:599–614. https://doi.org/10.1163/156853998792897932

Ikeda MZ, Jeon SD, Cowell RA, Remage-Healey L (2015) Norepinephrine modulates coding of complex vocalizations in the songbird auditory cortex independent of local neuroestrogen synthesis. J Neurosci 35:9356–9368. https://doi.org/10.1523/JNEUROSCI.4445-14.2015

Iyengar S, Bottjer SW (2002) The role of auditory experience in the formation of neural circuits underlying vocal learning in zebra finches. J Neurosci 22:946–958. https://doi.org/10.1523/JNEUROSCI.22-03-00946.2002

James LS, Sakata JT (2017) Learning biases underlie “universals” in avian vocal sequencing. Curr Biol 27:3676–3682.e4. https://doi.org/10.1016/j.cub.2017.10.019

Jin H, Clayton DF (1997) Localized changes in immediate-early gene regulation during sensory and motor learning in zebra finches. Neuron 19:1049–1059. https://doi.org/10.1016/S0896-6273(00)80396-7

Kao MH, Doupe AJ, Brainard MS (2005) Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song. Nature 433:638–643

Keller GB, Hahnloser RHR (2009) Neural processing of auditory feedback during vocal practice in a songbird. Nature 457:187–190. https://doi.org/10.1038/nature07467

Knudsen EI (2004) Sensitive periods in the development of the brain and behavior. J Cogn Neurosci 16:1412–1425. https://doi.org/10.1162/0898929042304796

Konishi M (2004) The role of auditory feedback in birdsong. Ann N Y Acad Sci 1016:463–475. https://doi.org/10.1196/annals.1298.010

Kozhevnikov AA, Fee MS (2007) Singing-related activity of identified HVC neurons in the zebra finch. J Neurophysiol 97:4271–4283. https://doi.org/10.1152/jn.00952.2006

Kuhl PK (2007) Is speech learning ‘gated’ by the social brain? Dev Science 10:110–120. https://doi.org/10.1111/j.1467-7687.2007.00572.x

Kuhl PK (2010) Brain mechanisms in early language acquisition. Neuron 67:713–727. https://doi.org/10.1016/j.neuron.2010.08.038

Kuhl PK, Tsao F-M, Liu H-M (2003) Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. PNAS USA 100:9096–9101. https://doi.org/10.1073/pnas.1532872100

Lachlan RF, van Heijningen CAA, ter Haar SM, ten Cate C (2016) Zebra finch song phonology and syntactical structure across populations and continents—a computational comparison. Front Psychol 7:980. https://doi.org/10.3389/fpsyg.2016.00980

Lei H, Mooney R (2010) Manipulation of a central auditory representation shapes learned vocal output. Neuron 65:122–134. https://doi.org/10.1016/j.neuron.2009.12.008

Leonardo A (2004) Experimental test of the birdsong error-correction model. Proc Natl Acad Sci 101:16935–16940. https://doi.org/10.1073/pnas.0407870101

Leonardo A, Konishi M (1999) Decrystallization of adult birdsong by perturbation of auditory feedback. Nature 399:466–470. https://doi.org/10.1038/20933

Lewandowski B, Vyssotski A, Hahnloser RHR, Schmidt M (2013) At the interface of the auditory and vocal motor systems: NIf and its role in vocal processing, production and learning. J Physiol-Paris 107:178–192. https://doi.org/10.1016/j.jphysparis.2013.04.001

Li R, Sakaguchi H (1997) Cholinergic innervation of the song control nuclei by the ventral paleostriatum in the zebra finch: a double-labeling study with retrograde fluorescent tracers and choline acetyltransferase immunohistochemistry. Brain Res 763:239–246. https://doi.org/10.1016/S0006-8993(97)00417-4

Lipkind D, Marcus GF, Bemis DK, Sasahara K, Jacoby N, Takahasi M, Suzuki K, Fehér O, Ravbar P, Okanoya K, Tchernichovski O (2013) Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. Nature 498:104–108. https://doi.org/10.1038/nature12173

Livingston FS, Mooney R (1997) Development of intrinsic and synaptic properties in a forebrain nucleus essential to avian song learning. J Neurosci 17:8997–9009. https://doi.org/10.1523/JNEUROSCI.17-23-08997.1997

London SE, Clayton DF (2008) Functional identification of sensory mechanisms required for developmental song learning. Nature Neurosci 11:579–586. https://doi.org/10.1038/nn.2103

Lorenz KZ (1935) Der kumpan in der umwelt des vogels. der artgenosse als auslösendes moment sozialer verhaltensweisen. J Ornithologie 83: 137–215, 289–413. https://doi.org/10.1007/BF01905355

Lytle SR, Kuhl PK (2017) Social interaction and language acquisition: toward a neurobiological view. In: Fernández EM, Cairns HS (eds) The handbook of psycholinguistics. John Wiley & Sons, Inc., Hoboken, pp 615–634

Mandelblat-Cerf Y, Las L, Denisenko N, Fee MS (2014) A role for descending auditory cortical projections in songbird vocal learning. eLife 3. https://doi.org/10.7554/eLife.02152

Margoliash D, Schmidt MF (2010) Sleep, off-line processing, and vocal learning. Brain Lang 115:45–58. https://doi.org/10.1016/j.bandl.2009.09.005

Marler P (1997) Three models of song learning: evidence from behavior. J Neurobiol 33:501–516

Mets DG, Brainard MS (2018) Genetic variation interacts with experience to determine interindividual differences in learned song. PNAS USA 115:421–426. https://doi.org/10.1073/pnas.1713031115

Mets DG, Brainard MS (2019) Learning is enhanced by tailoring instruction to individual genetic differences. eLife 8. https://doi.org/10.7554/eLife.47216.001

Michel GF, Tyler AN (2005) Critical period: a history of the transition from questions of when, to what, to how. Dev Psychobiol 46:156–162. https://doi.org/10.1002/dev.20058

Mooney R (2009) Neurobiology of song learning. Curr Opin Neurobiol 19:654–660. https://doi.org/10.1016/j.conb.2009.10.004

Moore JM, Woolley SM (2019) Emergent tuning for learned vocalizations in auditory cortex. Nat Neurosci 22:1469–1476. https://doi.org/10.1038/s41593-019-0458-4

Morales M, Margolis EB (2017) Ventral tegmental area: cellular heterogeneity, connectivity and behaviour. Nat Rev Neurosci 18:73–85. https://doi.org/10.1038/nrn.2016.165

Moseley DL, Joshi NR, Prather JF, Podos J, Remage-Healey L (2017) A neuronal signature of accurate imitative learning in wild-caught songbirds (swamp sparrows, *Melospiza georgiana*). Sci Reports 7:17320. https://doi.org/10.1038/s41598-017-17401-2

Murphy K, James LS, Sakata JT, Prather JF (2017) Advantages of comparative studies in songbirds to understand the neural basis of sensorimotor integration. J Neurophysiol 118:800–816. https://doi.org/10.1152/jn.00623.2016

Nelson DA (2000) A preference for own-subspecies' song guides vocal learning in a song bird. PNAS USA 97:13348–13353. https://doi.org/10.1073/pnas.240457797

Nicholson DA, Roberts TF, Sober SJ (2018) Thalamostriatal and cerebellothalamic pathways in a songbird, the Bengalese finch. J Comp Neurol 526:1550–1570. https://doi.org/10.1002/cne.24428

Nick TA (2015) Models of vocal learning in the songbird: historical frameworks and the stabilizing critic. Dev Neurobiol 75:1091–1113. https://doi.org/10.1002/dneu.22189

Nick TA, Konishi M (2005) Neural song preference during vocal learning in the zebra finch depends on age and state. J Neurobiol 62:231–42. https://doi.org/10.1002/neu.20087

Nordeen KW, Nordeen EJ (1992) Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. Behav Neural Biol 57:58–66. https://doi.org/10.1016/0163-1047(92)90757-U

Nordeen EJ, Holtzman DA, Nordeen KW (2009) Increased Fos expression among midbrain dopaminergic cell groups during birdsong tutoring. Eur J Neurosci 30:662–670. https://doi.org/10.1111/j.1460-9568.2009.06849.x

Nowicki S, Searcy WA (2014) The evolution of vocal learning. Curr Opin Neurobiol 28:48–53. https://doi.org/10.1016/j.conb.2014.06.007

O'Connell LA, Hofmann HA (2011) The vertebrate mesolimbic reward system and social behavior network: a comparative synthesis. J Comp Neurol 519:3599–3639. https://doi.org/10.1002/cne.22735

Odom KJ, Hall ML, Riebel K, Omland KE, Langmore NE (2014) Female song is widespread and ancestral in songbirds. Nature Comm 5:3379. https://doi.org/10.1038/ncomms4379

Okubo TS, Mackevicius EL, Payne HL, Lynch GF, Fee MS (2015) Growth and splitting of neural sequences in songbird vocal development. Nature 528:352–357. https://doi.org/10.1038/nature15741

Petkov CI, Jarvis ED (2012) Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. Front Evol Neurosci 4. https://doi.org/10.3389/fnevo.2012.00012

Phan ML, Pytte CL, Vicario DS (2006) Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds. PNAS USA 103:1088–1093. https://doi.org/10.1073/pnas.0510136103

Pidoux L, Le Blanc P, Levenes C, Leblois A (2018) A subcortical circuit linking the cerebellum to the basal ganglia engaged in vocal learning. eLife 7. https://doi.org/10.7554/eLife.32167

Pizzorusso T, Medini P, Landi S, Baldini S, Berardi N, Maffei L (2006) Structural and functional recovery from early monocular deprivation in adult rats. PNAS USA 103:8517–8522. https://doi.org/10.1073/pnas.0602657103

Plamondon SL, Rose GJ, Goller F (2010) Roles of syntax information in directing song development in white-crowned sparrows (*Zonotrichia leucophrys*). J Comp Psychol 124:117–132. https://doi.org/10.1037/a0017229

Prather JF (2013) Auditory signal processing in communication: perception and performance of vocal sounds. Hearing Res 305:144–155. https://doi.org/10.1016/j.heares.2013.06.007

Riede T, Goller F (2014) Morphological basis for the evolution of acoustic diversity in oscine songbirds. Proc R Soc B Biol Sci 281:20132306–20132306. https://doi.org/10.1098/rspb.2013.2306

Roberts TF, Mooney R (2013) Motor circuits help encode auditory memories of vocal models used to guide vocal learning. Hearing Res 303:48–57. https://doi.org/10.1016/j.heares.2013.01.009

Roberts TF, Tschida KA, Klein ME, Mooney R (2010) Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning. Nature 463:948–952. https://doi.org/10.1038/nature08759

Roberts TF, Gobes SMH, Murugan M, Ölveczky BP, Mooney R (2012) Motor circuits are required to encode a sensory model for imitative learning. Nature Neurosci 15:1454–1459. https://doi.org/10.1038/nn.3206

Roberts TF, Hisey E, Tanaka M, Kearney MG, Chattree G, Yang CF, Shah NM, Mooney R (2017) Identification of a motor-to-auditory pathway important for vocal learning. Nature Neurosci 20:978–986. https://doi.org/10.1038/nn.4563

Sakaguchi H (1996) Sex differences in the developmental changes of GABAergic neurons in zebra finch song control nuclei. Exp Brain Res 108:62–68. https://doi.org/10.1007/BF00242904

Sakata JT, Brainard MS (2006) Real-time contributions of auditory feedback to avian vocal motor control. J Neurosci 26:9619–9628. https://doi.org/10.1523/JNEUROSCI.2027-06.2006

Sakata JT, Brainard MS (2008) Online contributions of auditory feedback to neural activity in avian song control circuitry. J Neurosci 28:11378–11390. https://doi.org/10.1523/JNEUROSCI.3254-08.2008

Sara SJ, Bouret S (2012) Orienting and reorienting: the locus coeruleus mediates cognition through arousal. Neuron 76:130–141. https://doi.org/10.1016/j.neuron.2012.09.011

Sasaki A, Sotnikova TD, Gainetdinov RR, Jarvis ED (2006) Social context-dependent singing-regulated dopamine. J Neurosci 26:9010–9014. https://doi.org/10.1523/JNEUROSCI.1335-06.2006

Sato D, Mori C, Sawai A, Wada K (2016) Familial bias and auditory feedback regulation of vocal babbling patterns during early song development. Sci Reports 6. https://doi.org/10.1038/srep30323

Scharff C, Nottebohm F (1991) A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. J Neurosci 11:2896–2913. https://doi.org/10.1523/JNEUROSCI.11-09-02896.1991

Schultz W (2017) Reward prediction error. Curr Biol 27:R369–R371. https://doi.org/10.1016/j.cub.2017.02.064

Shea SD, Margoliash D (2003) Basal forebrain cholinergic modulation of auditory activity in the zebra finch song system. Neuron 40:1213–1226. https://doi.org/10.1016/S0896-6273(03)00723-2

Singh TD, Basham ME, Nordeen EJ, Nordeen KW (2000) Early sensory and hormonal experience modulate age-related changes in NR2B mRNA within a forebrain region controlling avian vocal learning. J Neurobiol 44:82–94. https://doi.org/10.1002/1097-4695(200007)44:1<82::AID-NEU8>3.0.CO;2-W

Snowdon CT, Hausberger M (1997) Social influences on vocal development. Cambridge University Press

Sober SJ, Brainard MS (2009) Adult birdsong is actively maintained by error correction. Nature Neurosci 12:927–931. https://doi.org/10.1038/nn.2336

Solis MM, Doupe AJ (1999) Contributions of tutor and bird's own song experience to neural selectivity in the songbird anterior forebrain. J Neurosci 19:4559–4584. https://doi.org/10.1523/JNEUROSCI.19-11-04559.1999

Sonntag M, Blosa M, Schmidt S, Rübsamen R, Morawski M (2015) Perineuronal nets in the auditory system. Hearing Res 329:21–32. https://doi.org/10.1016/j.heares.2014.12.012

Southwell DG, Froemke RC, Alvarez-Buylla A, Stryker MP, Gandhi SP (2010) Cortical plasticity induced by inhibitory neuron transplantation. Science 327:1145–1148. https://doi.org/10.1126/science.1183962

Stockard C (1921) Developmental rate and structural expression: an experimental study of twins, 'double monsters' and single deformities, and the interaction among embryonic organs during their origin and development. Am J Anat 28:115–277. https://doi.org/10.1002/aja.1000280202

Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. MIT Press

Tachibana RO, Takahasi M, Hessler NA, Okanoya K (2017) Maturation-dependent control of vocal temporal plasticity in a songbird: maturation effect on temporal plasticity. Dev Neurobiol 77:995–1006. https://doi.org/10.1002/dneu.22487

Takesian AE, Hensch TK (2013) Balancing plasticity/stability across brain development. Prog Brain Res 207:3–34. https://doi.org/10.1016/B978-0-444-63327-9.00001-1

Tanaka M, Sun F, Li Y, Mooney R (2018) A mesocortical dopamine circuit enables the cultural transmission of vocal behaviour. Nature 563:117–120. https://doi.org/10.1038/s41586-018-0636-7

Tang Y, Stryker MP, Alvarez-Buylla A, Espinosa JS (2014) Cortical plasticity induced by transplantation of embryonic somatostatin or parvalbumin interneurons. PNAS USA 111:18339–18344. https://doi.org/10.1073/pnas.1421844112

Tchernichovski O, Marcus G (2014) Vocal learning beyond imitation: mechanisms of adaptive vocal development in songbirds and human infants. Curr Opin Neurobiol 28:42–47. https://doi.org/10.1016/j.conb.2014.06.002

Tchernichovski O, Mitra PP, Lints T, Nottebohm F (2001) Dynamics of the vocal imitation process: how a zebra finch learns its song. Science 291:2564–2569. https://doi.org/10.1126/science.1058522

ter Haar SM, Kaemper W, Stam K, Levelt CC, ten Cate C (2014) The interplay of within-species perceptual predispositions and experience during song ontogeny in zebra finches (*Taeniopygia guttata*). Proc R Soc B Biol Sci 281:20141860–20141860. https://doi.org/10.1098/rspb.2014.1860

Thorpe WH (2008) The learning of song patterns by birds, with especial reference to the song of the chaffinch *Fringilla coelebs*. Ibis 100:535–570. https://doi.org/10.1111/j.1474-919X.1958.tb07960.x

Troyer TW, Doupe AJ (2000) An associational model of birdsong sensorimotor learning I. efference copy and the learning of song syllables. J Neurophysiol 84:1204–1223. https://doi.org/10.1152/jn.2000.84.3.1204

Tumer EC, Brainard MS (2007) Performance variability enables adaptive plasticity of ‘crystallized’ adult birdsong. Nature 450:1240–1244. https://doi.org/10.1038/nature06390

Vallentin D, Kosche G, Lipkind D, Long MA (2016) Inhibition protects acquired song segments during vocal learning in zebra finches. Science 351:267–271. https://doi.org/10.1126/science.aad3023

Velho TAF, Lu K, Ribeiro S, Pinaud R, Vicario D, Mello CV (2012) Noradrenergic control of gene expression and long-term neuronal adaptation evoked by learned vocalizations in songbirds. PLoS One 7:e36276. https://doi.org/10.1371/journal.pone.0036276

Vyssotski AL, Stepien AE, Keller GB, Hahnloser RHR (2016) A neural code that is isometric to vocal output and correlates with its sensory consequences. PLoS Biol 14:e2000317. https://doi.org/10.1371/journal.pbio.2000317

Wang Y, Brzozowska-Prechtl A, Karten HJ (2010) Laminar and columnar auditory cortex in avian brain. PNAS USA 107:12676–12681. https://doi.org/10.1073/pnas.1006645107

Wang H, Sawai A, Toji N, Sugioka R, Shibata Y, Suzuki Y, Ji Y, Hayase S, Akama S, Sese J, Wada K (2019) Transcriptional regulatory divergence underpinning species-specific learned vocalization in songbirds. PLoS Biol 17. https://doi.org/10.1371/journal.pbio.3000476

Warren TL, Tumer EC, Charlesworth JD, Brainard MS (2011) Mechanisms and time course of vocal learning and consolidation in the adult songbird. J Neurophysiol 106:1806–1821. https://doi.org/10.1152/jn.00311.2011

Warren TL, Charlesworth JD, Tumer EC, Brainard MS (2012) Variable sequencing is actively maintained in a well learned motor skill. J Neurosci 32:15414–15425. https://doi.org/10.1523/JNEUROSCI.1254-12.2012

Werker JF, Hensch TK (2015) Critical periods in speech perception: new directions. Ann Rev Psychol 66:173–196. https://doi.org/10.1146/annurev-psych-010814-015104

West MJ, King AP (1988) Female visual displays affect the development of male song in the cowbird. Nature 334:244–246. https://doi.org/10.1038/334244a0

West MJ, King AP, Eastzer DH (1981) Validating the female bioassay of cowbird song: relating differences in song potency to mating success. Anim Behav 29:490–501. https://doi.org/10.1016/S0003-3472(81)80110-8

Westermann G, Miranda E (2004) A new model of sensorimotor coupling in the development of speech. Brain Lang 89:393–400. https://doi.org/10.1016/S0093-934X(03)00345-6

Whaling CS, Solis MM, Doupe AJ, Soha JA, Marler P (1997) Acoustic and neural bases for innate recognition of song. PNAS USA 94:12694–12698. https://doi.org/10.1073/pnas.94.23.12694

Wheatcroft D, Qvarnström A (2015) A blueprint for vocal learning: auditory predispositions from brains to genomes. Biol Lett 11:20150155. https://doi.org/10.1098/rsbl.2015.0155

White SA, Livingston FS, Mooney R (1999) Androgens modulate NMDA receptor–mediated EPSCs in the zebra finch song system. J Neurophysiol 82:2221–2234. https://doi.org/10.1152/jn.1999.82.5.2221

Woolley SMN (2017) Early experience and auditory development in songbirds. In: Cramer KS, Coffin AB, Fay RR, Popper AN (eds) Auditory development and plasticity: in honor of Edwin W Rubel. Springer International Publishing, Cham, pp 193–217

Xiao L, Chattree G, Oscos FG, Cao M, Wanat MJ, Roberts TF (2018) A basal ganglia circuit sufficient to guide birdsong learning. Neuron 98:208–221.e5. https://doi.org/10.1016/j.neuron.2018.02.020

Yanagihara S, Hessler NA (2006) Modulation of singing-related activity in the songbird ventral tegmental area by social context. Eur J Neurosci 24:3619–3627. https://doi.org/10.1111/j.1460-9568.2006.05228.x

Yanagihara S, Yazaki-Sugiyama Y (2016) Auditory experience-dependent cortical circuit shaping for memory formation in bird song learning. Nature Comm 7:11946. https://doi.org/10.1038/ncomms11946

Yazaki-Sugiyama Y (2019) Neuronal mechanisms regulating the critical period of sensory experience-dependent song learning. Neurosci Res 140:53–58. https://doi.org/10.1016/j.neures.2018.11.002

Yazaki-Sugiyama Y, Mooney R (2004) Sequential learning from multiple tutors and serial retuning of auditory neurons in a brain area important to birdsong learning. J Neurophysiol 92:2771–2788. https://doi.org/10.1152/jn.00467.2004

Zann R (1993) Structure, sequence and evolution of song elements in wild Australian zebra finches. Auk 110:702–715. https://doi.org/10.2307/4088626

Zevin JD, Seidenberg MS, Bottjer SW (2004) Limits on reacquisition of song in adult zebra finches exposed to white noise. J Neurosci 24:5849–5862. https://doi.org/10.1523/JNEUROSCI.1891-04.2004

Zhao W, Garcia-Oscos F, Dinh D, Roberts TF (2019) Inception of memories that guide vocal learning in the songbird. Science 366:83–89. https://doi.org/10.1126/science.aaw4226

Chapter 3
New Insights into the Avian Song System and Neuronal Control of Learned Vocalizations

Karagh Murphy, Koedi S. Lawley, Perry Smith, and Jonathan F. Prather

Abstract The songbird brain contains a network of structures that are specialized for imitative vocal learning. Over the past few years, many new insights have emerged about the structure and function of that system. Among those insights are understanding how specific pathways contribute to specific aspects of vocal behavior, such as control of the acoustic properties of the vocalizations, control of the sequence in which those sounds are produced, and production of calls as well as songs. New research also has indicated that sites outside of the canonical vocal communication system play key roles in the learning, perception, and performance of the sounds used in vocal communication. This chapter details many of those insights and advocates for an expanded perspective on the vocal communication system as a set of interconnected nuclei that are specialized for the production of not only songs but also multiple types of learned signals used in vocal communication. In addition, several new experimental tools and approaches are highlighted that will allow deeper investigations into fundamental questions about the neural basis of learned vocal communication.

Keywords Auditory · Canary · Cortex · Finch · Learned vocal communication · Songbird · Sparrow · Striatum

---

K. Murphy · K. S. Lawley · P. Smith · J. F. Prather (✉)
Neuroscience Program, Department of Zoology and Physiology, University of Wyoming, Laramie, WY, USA
e-mail: kmurph17@uwyo.edu; klawley1@uwyo.edu; Jonathan.Prather@uwyo.edu

J. T. Sakata et al. (eds.), *The Neuroethology of Birdsong*, Springer Handbook of Auditory Research 71, https://doi.org/10.1007/978-3-030-34683-6_3

3.1 Introduction

Knowledge of the songbird vocal communication system is expanding at an increasing pace. With the advent of new experimental tools and the insights gained from these tools, studies of songbird neurobiology open the door to even broader questions. For example, how does the brain serialize individual actions into behaviorally meaningful sequences? How does the brain compute and employ error signals to refine subsequent behaviors? How does the nervous system store and recall memories to guide behavioral performance? The results and perspectives that are highlighted here make it clear that neurobiological research using songbirds holds the promise of yielding answers to these and other questions regarding the sensorimotor mechanisms that underlie learned behaviors.

3.2 Historical Perspective

3.2.1 The Canonical Song System

The songs performed by thousands of songbird species are learned behaviors that require juvenile auditory experience and ongoing auditory feedback to master and maintain the song. Beginning with the discovery that the songbird brain contains a network of discrete sites and circuits (Fig. 3.1) dedicated to song perception and

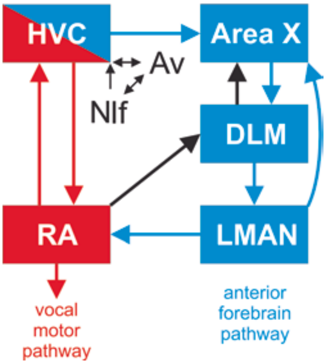


Fig. 3.1 Components of the song system. When the song system was first described, a role for five sites arranged into two interconnected pathways was highlighted (*red*, vocal motor pathway; *blue*, anterior forebrain pathway). In the years since those seminal experiments, additional connections between those sites have been found (additional connections are represented using *black arrows*), and new sites have also been identified as important contributors to song learning and maintenance (additional sites are represented using *black letters*). *Area X*, vocal portion of the avian basal ganglia; *Av*, avalanche; *DLM*, dorsolateral nucleus of the medial thalamus; *HVC*, sensorimotor cortical nucleus; *LMAN*, lateral magnocellular nucleus of the anterior nidopallium; *NIf*, nucleus interfacialis; *RA*, robust nucleus of the arcopallium

Table 3.1 Abbreviations

AFP	anterior forebrain pathway
Av	Avalanche, a ventral region in CM
CM	caudal mesopallium
CMM	caudomedial mesopallium
DLM	medial portion of the dosolateral thalamic nucleus
Field L	primary auditory cortex
HVC	used as proper name for vocal motor nucleus in the nidopallium
HVC_{Av}	HVC cells that project to Av
HVC_{RA}	HVC cells that project to RA
HVC_X	HVC cells that project to Area X
LMAN	lateral magnocellular nucleus of the anterior nidopallium
NCM	caudal medial nidopallium
NIf	nucleus interfacialis of the nidopallium
RA	robust nucleus of the arcopallium
UVA	nucleus uvaeformis
VMP	vocal motor pathway

performance (Nottebohm et al. 1976), researchers have sought to understand how this *song system* enables imitative learning. That seminal experiment, and many studies that followed, revealed five interconnected brain sites that are closely associated with song performance: the sensorimotor cortical nucleus HVC (used as a proper name for the vocal motor nucleus in the nidopallium all abbreviations appear in Table 3.1), the motor cortical nucleus RA (robust nucleus of the arcopallium), and specialized regions of anterior forebrain (lateral magnocellular nucleus of the anterior nidopallium, LMAN), the basal ganglia (Area X) and the thalamus (dorsolateral nucleus of the medial thalamus, DLM) (Reiner et al. 2004; Mooney et al. 2008). Among those sites, HVC plays an especially important role because injury to HVC is associated with deficits in both song performance (Nottebohm et al. 1976) and song perception (Brenowitz 1991; Gentner et al. 2000) and because the activity of HVC neurons plays a central role in song learning (Roberts et al. 2012; Roberts et al. 2017). Consistent with a key role for HVC in learned vocal imitation, HVC activity encodes sensory and motor information about song performance (Katz and Gurney 1981; McCasland and Konishi 1981). Furthermore, HVC is the origin of two pathways (Fig. 3.1): the *vocal motor pathway* (VMP) through the motor cortical nucleus RA, and the *anterior forebrain pathway* (AFP) through the basal ganglia Area X. Both of these pathways play key roles in vocal learning (see Sakata and Yazaki-Sugiyama, Chap. 2; Leblois and Perkel, Chap. 4).

3.2.2 Composition of the Song System and Functions of Specific Components

The VMP includes axonal projections from cell bodies in HVC to the motor nucleus RA (Fig. 3.2). Lesions to either HVC or RA result in vocal deficits such as "permanent and complete elimination of the audible components of song" (Nottebohm et al. 1976). Through these projections from HVC to RA and from RA to targets downstream, the VMP exercises control over brainstem motor neurons that control the muscles of the vocal organ and the associated respiratory structures (Schmidt and Wild 2014; Wild and Botelho 2015). The HVC neurons that project to RA (HVC_{RA} cells) are active during singing in a cascade of activity from one neuron to the next (Hahnloser et al. 2002; Lynch et al. 2016). Focally cooling HVC neurons disrupted the kinetics of HVC activity and resulted in the slowing of song performance (Long and Fee 2008; Andalman et al. 2011); whereas, focally warming HVC had the opposite effect (Long and Fee 2008). This led researchers to suspect that temporal features may be primarily controlled by forebrain circuits including HVC, but some aspects of song timing may be controlled by sites outside of HVC and the VMP (considered further in Sect. 3.3.4).

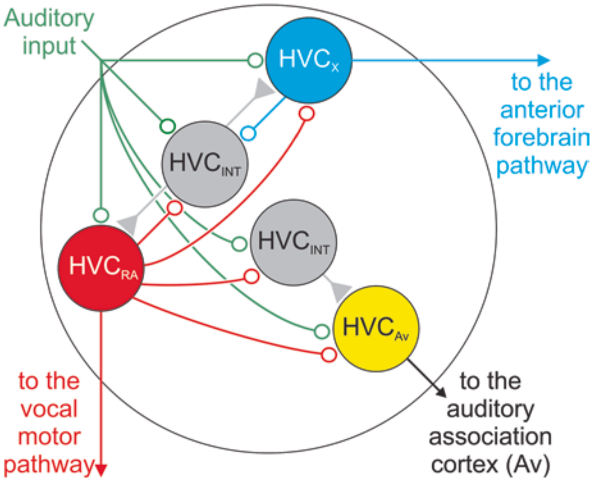


Fig. 3.2 The vocal motor pathway and anterior forebrain pathway emerge from the HVC microcircuit. Three types of neurons that project their axons to sites outside of HVC and at least two classes of locally projecting interneurons (*gray* HVC_{INT}) are arranged in a local HVC microcircuit that gives rise to three pathways: a projection to the vocal motor nucleus RA (*red* HVC_{RA} *cells*), a projection to the avian basal ganglia nucleus Area X (*blue* HVC_X *cells*), and a projection to the auditory association cortical area Avalanche (Av; *yellow* HVC_{Av} *cells*). As noted in Figs. 3.1 and 3.3, this local HVC microcircuit is interconnected with many other sensory and motor sites to enable feedback-dependent imitative vocal learning. See Table 3.1 for all abbreviations. (adapted from Prather 2013)

The motor nucleus RA integrates input from HVC and from another pathway into which HVC neurons project, the AFP (Fig. 3.1). The AFP begins with projection neurons from HVC to the specialized basal ganglia nucleus Area X (HVC_X cells). Area X neurons project to the thalamic nucleus DLM, which projects to the cortical nucleus LMAN, which in turn projects onto the VMP at the level of RA and projects back to Area X (Fig. 3.1). The relative influence of HVC versus LMAN on the activity of RA neurons varies over the course of development: LMAN has the dominant influence during early development, and HVC affects RA far more strongly in the adult state (Aronov et al. 2008; Olveczky et al. 2011). In contrast to the VMP, experimental lesions within the AFP did not prevent birds from singing. Instead, lesioning in AFP impaired a bird's ability to imitate tutor songs and prevented motor exploration for trial-and-error learning. For example, lesions of either LMAN or Area X in zebra finches (*Taeniopygia guttata*) had little or no effect on the production of stereotyped song by adult birds but prevented normal song development in juvenile birds (Sohrabji et al. 1990; Scharff and Nottebohm 1991). Those results suggested that Area X and its projections within the AFP may be important for either learning or improving the bird's imitation of a tutor song through vocal practice. Support for that idea also comes from the observations that exposure to tutor song for as little as one day was sufficient to enhance vocal learning significantly and to stabilize dendritic spines on HVC_X neurons (Roberts et al. 2010; Chen et al. 2016). These data highlight HVC_X neurons as central mechanisms through which experience can influence vocal performance.

Area X neurons project to the thalamic nucleus DLM, and DLM neurons are entrained to their inputs from Area X with very high precision (Goldberg et al. 2012). Furthermore, lesions in DLM largely abolished the normal variation present in the babbling of young birds and caused a dramatic increase in song stereotypy (Goldberg and Fee 2011). While much remains to be explored regarding the function of DLM neurons, those data suggest that DLM also plays a key role in the variation in vocal behavior that occurs during song learning.

DLM neurons project to the cortical nucleus LMAN, and normal activity in LMAN is essential for vocal learning (Bottjer et al. 1984). LMAN neurons that project to RA are thought to introduce variation into the performance of song features that are controlled by activity in HVC and RA, as variance of these features was vastly decreased when the output of the AFP was inactivated (Olveczky et al. 2005). In addition, blocking the output of the AFP during tutoring prevented the gradual improvement in imitation that normally occurs during song learning (Charlesworth et al. 2012; Ali et al. 2013). Thus, both the AFP and the VMP exert their influences on vocal control through their projections onto the motor cortical nucleus RA, and it is through these pathways that the song system controls song learning and performance.

Historically, experience-dependent vocal learning has been most closely related to the five structures within the VMP and AFP (Fig. 3.1). However, those brain areas are interconnected with other sites that are also responsive to auditory stimuli and are active during vocal performance. The following sections highlight advances in knowledge of how sites in the VMP and AFP function and how they may work in concert with additional sites to form the neural basis of learned vocal communication.

3.3 New Insights Into the Function of Sites in the Canonical Song System

3.3.1 Role of HVC in Imitative Learning

Across the population of HVC cells, different projection neurons respond to different portions of the adult song, providing a spatially intermingled representation of the vocal repertoire (Peh et al. 2015). In addition to that sensory representation, some neurons can represent both the adult song and tutor songs that the bird heard only during juvenile development (Prather et al. 2010; Moseley et al. 2017). While those cells respond to both a tutor song and the bird's imitation of that model (Moseley et al. 2017), the co-representation is not simply a result of acoustic similarity between those song types because the strength of auditory response is not predicted by the degree of acoustic similarity between the tutor song and the bird's copy (Prather et al. 2010; Moseley et al. 2017). Intriguingly, the prevalence of these co-responsive cells, termed *bridge cells*, is directly related to the bird's accuracy of imitative learning: birds that possess greater numbers of bridge cells also achieve better learning outcomes (Moseley et al. 2017). These data provide mechanistic insight into how HVC may contribute to imitative vocal learning. Auditory-responsive cells are located throughout HVC, cells that represent different portions of the vocal repertoire are spatially intermingled, and the number of cells that are co-responsive to both the bird's own song and tutor song is correlated with the accuracy of imitative learning (see Sakata and Yazaki-Sugiyama, Chap. 2).

3.3.2 Activity of HVC During Singing

HVC neurons are active when birds produce the sounds and the silences that compose their songs, and projection neurons generate brief bursts of action potentials at very precise points in the song (Hahnloser et al. 2002; Prather et al. 2008). Activity across the population of HVC projection neurons is nearly uniform in its distribution of activity throughout the syllables and gaps that compose song (Lynch et al. 2016). Rather than being tightly organized around specific motor gestures, as had been hypothesized (Amador et al. 2013), the activity of $\mathrm{HVC}_{\mathrm{RA}}$ neurons appears to proceed in a way that is not directly coupled to specific movements. The population of $\mathrm{HVC}_{\mathrm{RA}}$ neurons expresses a continuous cascade of activity throughout the song (Lynch et al. 2016; Picardo et al. 2016): each syllable is driven by a population of $\mathrm{HVC_{RA}}$ neurons that is transiently active at each point in the song, and different populations drive different patterns of activity in the downstream motor pathway.

Several lines of evidence suggest that HVC generates the timing of individual song elements through neurons organized in a synaptic chain. Different mechanisms have been proposed through which that organization may be achieved (Hamaguchi et al. 2016; Galvis et al. 2017), but the common theme in those scenarios is the organization of neurons into chains of sequenced activity that direct song behavior. In these proposed synaptic chains, a population of HVC neurons acts like a clock producing a continuous series of “ticks” in which each tick leads to the activation of an ensemble of neurons in RA to produce the specific features that compose the corresponding portion of the song (Fee et al. 2004; Lynch et al. 2016). Some researchers have proposed that such chains may reside in HVC and control only timing (Long et al. 2010; Lynch et al. 2016). Others have proposed that multiple chains may exist in HVC and control both timing and sequencing (Galvis et al. 2017), and speculations have been mixed regarding the possible contributions of pathways that reside outside of HVC (Hamaguchi et al. 2016; Galvis et al. 2017). Additional research is necessary to more fully understand how these networks may give rise to the different phonology and syntax expressed by different songbird species, but the data suggest that circuits within HVC form part or all of a circuit that underlies the generation of song timing.

3.3.3 Function of the Vocal Motor Pathway in Singing

Within the VMP, manipulation of activity in HVC alters the temporal features of song, whereas manipulation in RA alters acoustic features. For example, altered androgen signaling in RA induced changes in the spectral properties of syllables but not their sequencing; however, similar manipulations in the upstream nucleus HVC induced changes in syllable sequencing but not acoustic features (Alward et al. 2017). In addition, pharmacological suppression of inhibition in RA produced robust and consistent increases in syllable pitch and amplitude; increasing inhibition in RA decreased syllable pitch and amplitude (Miller et al. 2017). Those results indicated that vocal performance is closely related to a precise balance of excitation and inhibition within the local network of RA. Additional experiments revealed that millisecond-scale changes in the activity of neurons in the VMP can result in changes in the output of the muscles used in respiration and control of the vocal organ (Srivastava et al. 2017). Even very subtle changes in the activation of those neurons and the associated muscles can result in song changes ranging from small differences in phonology to impairments of frequency modulation or elimination of certain syllables altogether (Sober et al., 2008; Mencio et al. 2017).

3.3.4 Neural Basis of Vocal Sequencing

Birdsong is like human language in that both skills require that individuals learn and apply syntactical rules. Acquisition of those skills can be quite challenging; for example, grammatical complexity is usually not fully mastered until at least seven years of age (Skeide and Friederici 2016). Similarly, the proper sequencing of individual sounds is essential to communication through song, and birds must learn the syntax of their songs just as they learn the phonology of individual notes (Prather et al. 2017; ten Cate 2018). In that process of learning vocal syntax, birds possess biological predispositions toward certain species-typical patterns, suggesting that neural or motor constraints may contribute to sequence-learning biases (James and Sakata 2017); however, those possible biases are not very restrictive since results from Bengalese finches (*Lonchura striata* var. *domestica*) revealed that birds can still express behavioral flexibility (Warren et al. 2012). This flexibility indicates that sequencing of the sounds that birds use in vocal communication does not reflect hardwired premotor circuitry. Instead, sequencing appears to be at least partially under active control.

Experimental manipulations revealed that activity in HVC plays a very significant role in shaping vocal sequence. Microlesions of the medial portion of HVC resulted in an increase in atypical syllable transitions in zebra finches (Basista et al. 2014). Similarly, cooling HVC in Bengalese finches altered not only the song tempo but also the probability that the bird would perform specific sequences (Zhang et al. 2017). Altered androgen signaling in HVC in canaries (*Serinus canaria*) also is associated with changes in syllable sequencing (Alward et al. 2017). In contrast, neither cooling nor altering the androgen signaling in the downstream motor area RA had any detectable effect on song syntax. Together, these studies make it clear that activity of neurons in HVC plays an important role in arranging individual vocal behaviors into behaviorally meaningful sequences.

HVC exerts its influence on vocal sequencing through activation of RA (VMP) and downstream motor neurons, but whether the AFP may also contribute to the control of vocal sequencing remains unclear. For example, lesions to the input of the AFP (Area X) abolished a zebra finch's ability to learn spectral features but left modification of temporal structure largely unchanged (Ali et al. 2013). In addition, lesions to the output of the AFP (LMAN) affected the spectral properties of Bengalese finch song syllables but not the sequence in which those syllables were produced (Hampton et al. 2009), suggesting that the AFP plays no role in vocal sequencing. In contrast, other authors found that pharmacological manipulations that increased the level of activity in LMAN (Hamaguchi and Mooney 2012) or the expression of specific genes in Area X (Tanaka et al. 2016) could induce sequence variability in adult zebra finches. Thus, findings have been inconsistent regarding the role of the AFP in sequencing, but they suggest the AFP plays little or no role in controlling that aspect of learned vocal behavior.

Studies of the neural basis of vocal sequencing have also suggested that circuits outside of the song system (but that feed into the song system) contribute to vocal sequencing. For example, bilateral lesions of the thalamic nucleus uvaeformis (UVA) caused long-lasting song degradation (Coleman and Vu 2005). Those deficits can improve over time, but subtle changes in syllable sequencing can sometimes persist even after other parts of the song have recovered (Coleman and Vu 2005). The UVA receives synaptic drive from RA and projects to HVC (Coleman et al. 2007), suggesting that UVA could influence activity in the song system and thus play a role in shaping vocal sequencing. However, as in the previous consideration of a possible role for the AFP in vocal sequencing, studies of the effects of UVA have also yielded mixed results. For example, manipulating activity in UVA by focally cooling those neurons induced changes in the timing of individual elements but not in their sequence (Hamaguchi et al. 2016). Therefore, both of the pathways that emerge from HVC and are eventually recurrent onto vocal motor pathways (UVA and the AFP) may contribute to some features of vocal sequencing, but they apparently play a relatively minor role in that process.

In light of lingering uncertainty about the degree to which vocal sequencing is controlled by mechanisms residing in HVC and/or circuitry outside of HVC (that is recurrent onto the song system), it is especially interesting that the integrated control of phonology and vocal sequence may involve contributions from both HVC and the AFP. For birdsong and human speech, the fine phonological details of a specific vocalization can depend on the specific sequence in which that vocalization is embedded (Wohlgemuth et al. 2010; Bolhuis and Everaert 2013). In songbirds, the contributions of HVC and the AFP to this context specificity were realized through experiments in which presentations of song-contingent aversive stimuli were used to evoke changes in the properties of specific targeted syllables in Bengalese finch song (Tian and Brainard 2017). In songs of that species, a specific syllable can be performed as part of many different sequences (e.g., syllable A performed as part of A→B or A→C or A→D). When aversive stimuli were delivered in association with one sequence context (e.g., stimuli delivered when syllable A was performed as part of A→B), learned alterations of the properties of the targeted syllable could generalize across contexts (e.g., changes in A were also evident in A→C transitions). Moreover, this context-specific expression of contingent changes was strongly dependent on activity in the AFP (Tian and Brainard 2017). For example, following context-dependent changes to pitch (e.g., targeted changes to A in A→B that also caused nontargeted changes to A in A→C), inactivation of LMAN caused larger changes to pitch for syllables in the targeted context (A→B) than for the same syllable in the nontarget context (A→C). The implications of those data from Bengalese finches are that the VMP may encode a relatively context-independent representation of a given syllable and that the AFP may provide a sequence-specific biasing signal to the motor pathway that enables the context-specific modulation of syllable features. The biasing signal may gradually modify the motor pathway representation such that the integrated control of spectral and sequential song features reflects a collaboration between mechanisms in the VMP and the AFP (Tian and Brainard 2017). An important goal of future research will be to continue to develop understanding of the neural mechanisms through which individual units of behavior (e.g., syllables) are learned and serialized into behaviorally relevant sequences (see Sect. 3.6).

3.3.5 Expanding the Function of Neurons in the Song System

In addition to songs, male songbirds also produce another form of learned vocalization: *calls*. Calls are typically shorter in duration and less complex in acoustic structure than syllables that compose song, and their relative simplicity initially led researchers to speculate that they are an innate behavior. However, behavioral studies have revealed that at least some aspects of call vocalizations are learned (Simpson and Vicario 1990).

Historically, relatively few studies have investigated the learning and performance of call vocalizations; instead, research has been focused primarily on songs because of their obvious dependence on sensory learning and ongoing auditory feedback (Mooney et al. 2008). However, researchers have become more interested in calls after the discovery that call performance is related to activity in the song system (see Elie and Theunissen, Chap. 7). For example, neural activity in the vocal motor nuclei HVC and RA is correlated with the production of calls by zebra finches (Hahnloser et al. 2002; Ter Maat et al. 2014), and blocking activity in song system structures reduces the precision of call timing and abolishes a bird's ability to avoid being "jammed" by predictable exogenous stimuli (Benichov et al. 2016). Additionally, lesions in HVC induced changes in call production in Bengalese finches (Murphy and Prather 2016; Urbano et al. 2016). Interneurons in HVC and HVC_X neurons that project into the AFP are active both when the bird performs a call and when the bird hears playback of those calls (Murphy and Prather 2016). Together, these findings reveal that the *song system* should be viewed as a set of interconnected nuclei that are not specialized just for song performance but are important for the production of multiple types of learned signals used in vocal communication.

3.4 An Expanded View of the Vocal Communication System

Originally, researchers thought that essential components of song learning (e.g., processing auditory feedback and comparing that information to a memorized song template to improve the quality of imitation) were functions of the canonical song system. These processes are now known to be associated with activity not only in those sites but also in additional areas of the songbird forebrain.

3.4.1 Nucleus Interfacialis: An Interface Between Auditory and Vocal Motor Activity

The nucleus interfacialis of the nidopallium (NIf) is part of a sensorimotor loop that connects auditory inputs to the VMP. The NIf is a major source of auditory input to HVC and is the primary driver of spontaneous activity in HVC neurons (Cardin and

Schmidt 2004; Coleman and Mooney 2004). Moreover, auditory activity in NIf plays a central role in how juvenile birds learn to imitate the sounds that they eventually include in their adult songs (Roberts et al. 2012; Lewandowski et al. 2013). In addition to being active in response to auditory stimuli, NIf neurons also are active during song production, with bursts of activity occurring prior to motor output (McCasland 1987; Lewandowski and Schmidt 2011). Singing-related activity in NIf is thought to play at least some role in how the brain sequences individual syllables to compose songs because bilateral lesions of NIf induced changes in song syntax in Bengalese finches but not in species like zebra finches that produce syntactically simpler songs (Hosino and Okanoya 2000; Cardin et al. 2005). However, experiments in zebra finches documented that transient inactivation (as opposed to permanent lesions) of NIf transformed their usual sequentially stereotyped songs into highly variable and unstructured vocalizations (Otchy et al. 2015; Piristine et al. 2016). These data suggest that the activity of NIf neurons may be an important component of the premotor activity underlying song performance but other components of the song system may eventually be able to compensate for its absence if NIf is permanently removed (see Pratt and Prather 2016).

Manipulations of activity in other regions outside of the song system also influence the activity of NIf and its downstream targets in HVC. For example, microinjection of fadrozole into a secondary auditory area, the caudal medial nidopallium (NCM), to focally inhibit estradiol synthesis induced changes in the selectivity of auditory responses of neurons in NIf and HVC (Pawlisch and Remage-Healey 2015; also see Remage-Healey, Chap. 6). Thus, NIf may also act as a conduit through which other regions outside of the canonical song system can affect the activity of neurons in HVC and its downstream pathways (Coleman and Mooney 2004; Bauer et al. 2008).

3.4.2 The Auditory Lobule Is Essential for Imitative Vocal Learning

Vocal learning and the function of the song system are intimately dependent on auditory processing: auditory input is a key component of imitative learning, and both juvenile learning and adult song maintenance are dependent on auditory feedback (Mooney et al. 2008; Murphy et al. 2017). In songbirds, auditory processing is closely linked not only to HVC and the AFP but also to primary and secondary regions that are analogous to auditory cortical regions in mammals (Jarvis et al. 2005; Wang et al. 2010). These auditory cortical sites in songbirds have been implicated in the formation, storage, and recall of tutor song memory (Prather and Mooney 2004; Prather 2013) and are interconnected with HVC (Fig. 3.3), forming a sensorimotor loop that links the auditory and vocal motor systems (Akutagawa and Konishi 2010; Lewandowski et al. 2013). Because of this functional implication in learning and the anatomical connection with other sensorimotor structures,

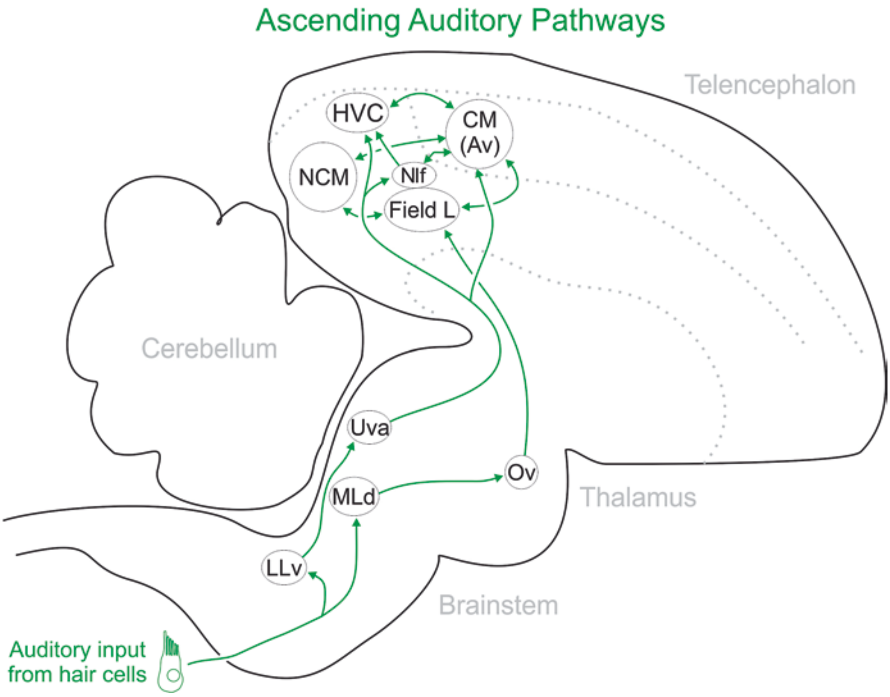


Fig. 3.3 Ascending pathways provide auditory input to the vocal communication system. The HVC and other elements in the canonical song system receive auditory input from hair cells of the inner ear along pathways that ascend in the songbird brain (see Elie and Theunissen, Chap. 7). Those sites play important roles in song learning and exert their effects by influencing activity in HVC and its downstream pathways. See Table 3.1 for all abbreviations. (adapted from Prather 2013)

auditory cortical structures outside of the canonical song system have become the focus of increasing amounts of research (see Sakata and Yazaki-Sugiyama, Chap. 2; Woolley and Woolley, Chap. 5).

The caudal mesopallium (CM) and the NCM are two such regions outside of the canonical song system that play important roles in the memorization and recall of tutor song (Gobes and Bolhuis 2007; London and Clayton 2008). These regions (described here as the auditory lobule) are advantageously positioned to serve such a role because they receive input from the primary auditory cortex (Field L; Fig. 3.3) and send projections to other forebrain sites including HVC (Bauer et al. 2008; Akutagawa and Konishi 2010). Experimental results have made it clear that these regions play a central role in the formation and retention of auditory memories and thus are essential for imitative vocal learning. For example, when cellular cascades are altered in the caudomedial mesopallium (CMM) and NCM during tutor sessions, juveniles produced poor copies of the tutor song (London and Clayton 2008; also see London, Chap. 8). Thus, the auditory cortical areas CMM and NCM are thought to contain the neural substrate for tutor song memory (Gobes and Bolhuis

2007; London and Clayton 2008). Moreover, these areas apparently act in concert with sites within the classically defined song system to shape vocal learning because impairment of song copying also occurred when activity in HVC was disrupted during tutor song exposure (Roberts et al. 2012) and because experience-dependent activity in CM influenced activity in HVC that guided vocal output (Bauer et al. 2008; Roberts et al. 2017).

In support of the idea that activity in the auditory lobule may contribute to vocal learning, neurons in the CMM and NCM appear to be predisposed to respond to songs that birds go on to memorize and learn. For example, CMM and NCM neurons were most likely to respond to songs with species-typical characteristics such as song syntax and rhythmicity (Lampen et al. 2017; Mello et al. 1992). Furthermore, auditory-evoked gene expression in CMM can reflect the behavioral relevance of the sounds because responses to familiar songs were greater than responses to novel stimuli (Gentner et al. 2004; Terpstra et al. 2004). The highly selective auditory responses of auditory lobule neurons are evident even in female songbirds, indicating that selective responsiveness to behaviorally relevant stimuli is an auditory characteristic of the cells in CMM and NCM and does not require the co-expression of learned vocal behavior (Diez et al. 2017). Elsewhere in the songbird brain, data from HVC revealed that the amount of auditory evoked gene expression in HVC was positively correlated with the strength of song learning (Terpstra et al. 2004; Bolhuis et al. 2012). Together, these data suggest that neurons in the auditory lobule, HVC, or a combination of both locations contain the neural representation of a tutor song memory that is compared against ongoing vocalizations to guide the bird's vocal learning.

In addition to evidence from studies of gene expression, electrophysiological recordings also confirm that activity in CMM and NCM is closely related to auditory and social experience. Recordings of neural activity in developing birds revealed that by the time the bird begins to sing, small numbers of neurons are selectively responsive to tutor song, and those cells are distributed across several forebrain regions, including the auditory lobule and HVC (Adret et al. 2012). Among the cells in the auditory lobule, NCM neurons responded differently to songs that were novel versus songs that were at least familiar or perhaps even served as tutor songs (Thompson and Gentner 2010). These data suggest that learning is associated with neuronal responses in which behaviorally relevant stimuli elicit less robust activity in NCM neurons than irrelevant stimuli, providing a possible mechanism through which experience can be tied to learning (but see Adret et al. 2012 and also Yanagihara and Yazaki-Sugiyama 2016). Selective responsiveness was also evident in the activity of CMM neurons, where cells showed increased activity in response to songs that had been learned through operant conditioning than to novel songs (Gentner and Margoliash 2003). Neurons in CMM were more selective and encoded more information about song components than neurons in the lateral portion of CM (Jeanne et al. 2011; Calabrese and Woolley 2015).

Reminiscent of what has been described in HVC, activity in the auditory lobule also extends to calls as well as songs. Specifically, neurons in CMM and NCM were selectively responsive to calls from specific individuals, and some neurons in those regions could use degraded spectral content to discriminate the distance across which that call travelled before reaching the receiver (Menardy et al. 2012; Mouterde et al. 2017). Those neurons may contribute to recognition of not only the location but also the identity of the source of a vocal signal. CMM and NCM are highly interconnected (Vates et al. 1996), and selectivity of auditory responses in each area is greater than that observed in the primary auditory cortex (Field L) (Meliza and Margoliash 2012). Thus, activity coursing through CMM or NCM may constitute different cortical processing streams that encode different yet complimentary features of the learned behaviors that birds use in vocal communication (Meliza et al. 2010).

3.4.3 A Subregion Within the Auditory Lobule Links Sensory and Motor Brain Sites

Within the CM portion of the auditory lobule resides a subpopulation of neurons that have been implicated as playing an especially important role in linking sensory and motor activity in vocal learning. That population of cells resides in the ventromedial portion of CM and is called nucleus Avalanche (Av). Like its surrounding auditory cortical regions, Av is activated by auditory stimuli (Jarvis and Nottebohm 1997), but in contrast to the connectivity of other portions of CM or NC, Av is reciprocally interconnected with HVC (Akutagawa and Konishi 2010; Roberts et al. 2017). The connection from HVC into Av provides a mechanism through which motor-related signals may be integrated with activity in the auditory system. Such a means of conveying information from the vocal production pathway to the auditory system has been theorized to facilitate vocal learning (e.g., Troyer and Doupe 2000), and $\mathrm{HVC}_{\mathrm{Av}}$ neurons are well positioned within the HVC microcircuit to provide that link (Fig. 3.2). In support of a role for $\mathrm{HVC}_{\mathrm{Av}}$ cells in vocal learning, ablation of those neurons in juvenile birds disrupted the bird's ability to imitate features of an adult tutor's song (Roberts et al. 2017). In addition, ablation of $\mathrm{HVC}_{\mathrm{Av}}$ neurons in adult birds interfered with the bird's ability to adaptively modify the duration of vocal elements but did not affect the production of previously learned songs (Roberts et al. 2017). These data reveal that the connection between HVC and Av provides a link between motor and auditory pathways that is essential for vocal plasticity. Interestingly, some neurons in the region of CM where Av resides provide polysynaptic input to midbrain dopaminergic regions that are speculated to generate error signals to guide vocal plasticity in both juvenile and adult birds (Mandelblat-Cerf and Fee 2014; Dunning et al. 2018). One possibility is that information may flow from HVC into Av to be compared against auditory feedback, with the result of that comparison passed to dopaminergic areas and used to reinforce the behavior associated with that feedback (Roberts et al. 2017; Hisey et al. 2018). An important goal of future research will be to test that possibility and to further resolve the causal relation between activity in these pathways and juvenile learning and adult maintenance of the sounds used in vocal communication.

3.5 New Tools to Explore the Function of Neurons in the Vocal Communication System

In the past few years, a number of new and innovative ways have emerged to reveal the function of neural circuits in the songbird brain. For example, there have been significant advancements in methods for electrophysiological recordings from neurons in awake and freely behaving birds, methods of imaging activity across neuronal populations, and methods for stimulating specific pathways using optogenetic techniques. The development and adoption of new experimental tools has opened the door to addressing questions that were previously intractable. These tools have advanced the ways in which songbird neurobiologists can sample or selectively manipulate the activity of individual neurons or populations of cells in the appropriate behavioral context and, thus, continue to reveal the mechanisms and functions of neurons in the expanded song system.

3.5.1 Recording Activity of Identified Neurons in Freely Behaving Birds

The ability to sample neural activity while birds are freely behaving is essential to investigate the neural circuits that underlie behavioral performance. In the early years following the discovery of the canonical song system, researchers were able to record from awake and freely behaving birds, but it was challenging to identify individual neurons among a population of simultaneously recorded cells (McCasland 1987; Yu and Margoliash 1996). The development of a miniature, motorized microdrive and the use of antidromic stimulation techniques facilitated the sampling of extracellular activity from individual, type-identified neurons as birds were engaged in singing and song perception (Fee and Leonardo 2001; Hahnloser et al. 2002). Moreover, the implementation of intracellular recordings in awake, singing birds has enabled the detection not only of action potentials but also subthreshold influences on song control that were previously invisible in extracellular recordings (Long et al. 2010; Hamaguchi et al. 2014). Microdrive recording systems have been adapted to be extremely lightweight (less than 1 g in their motorized configuration) (Jovalekic et al. 2017) and have incorporated additional data collection systems, such as a compass and accelerometer (M. Fee, personal communication), and telemetered recording (e.g., Schregardus et al. 2006; Hasegawa et al. 2015). These technologies continue to advance, and they will provide increasingly high-resolution insights into the neural circuits and patterns of cellular activity that underlie learned vocal communication.

3.5.2 *Optical and Electrophysiological Methods*

Microelectrode arrays and the ability to record from individual neurons over long periods of time have increased our knowledge of how neural activity is related to the acquisition and refinement of song behavior. Using carbon nanofibers arranged in an electrode array, it is possible to record from neurons in HVC for as long as 107 days (Guitchounts et al. 2013). Using electrode arrays, it is possible to simultaneously record from tens or perhaps even hundreds of sites in the brain, providing an electrophysiological approach to recording the activity of many neurons at once in sites that are deeper than imaging techniques may be able to sample. Sampling from neurons over long durations can allow researchers to potentially record from an individual neuron as it matures through plastic song development to stereotyped adult song (Okubo et al. 2015). These types of techniques will be central to our future understanding of how experience shapes motor performance to enable vocal learning and maintenance.

Similar to microelectrode arrays, which can sample many neurons simultaneously, imaging techniques can also discern the activity of many neurons at one time during behavior (Picardo et al. 2016; Katlowitz et al. 2018). Microimaging of cells and calcium imaging of voltage fluctuations have enhanced the scope of what is possible in songbird research. For example, head-mounted microscopes can be used in concert with fluorescent indicators to monitor calcium concentrations in relatively superficial structures such as HVC, and these types of techniques can enable sampling from the same population of neurons in freely behaving birds over the course of several weeks (Liberti et al. 2017). As in the case of electrophysiological recordings, these techniques will continue to provide insight into how the perception and performance of vocal behaviors emerge from the activity of multiple neuronal populations.

3.5.3 *Selectively Activating Neurons to Discern the Function of Specific Pathways*

In addition to recording the activity of neuronal populations, the ability to selectively manipulate the activity in those cells is also a powerful tool to understand how specific circuits contribute to auditory perception and vocal learning. The creation of optogenetic techniques gave the entire field of neuroscience the ability to use light to selectively excite or inhibit neurons with very high spatial and temporal resolution (Deisseroth 2015). This technique has since been adapted for use in songbirds and is beginning to reveal the role of specific cells in juvenile song learning (Roberts et al. 2012; Hisey et al. 2018) and adult song plasticity (Hisey et al. 2018; Xiao et al. 2018). Together with automated methods of monitoring behavior, optogenetics can enable precise stimulation of individual pathways, thus forming a behavior-dependent, closed-loop experimental paradigm to investigate the role of specific pathways in specific aspects of behavior (Tanaka et al. 2018; Xiao et al. 2018).

3.5.4 Manipulating the Genetic Environment

With the continued development of ways to understand and control the genetic environment in songbirds (Agate et al. 2009; Liu et al. 2015), genetic techniques will further enhance our ability to identify the neural structures and functions that underlie vocal control and learning. For example, songbird researchers can induce neurons in specific regions of the songbird brain to express human genes that are thought to play important roles in speech development and pathology (Abe et al. 2015; Tanaka et al. 2016). With those genes expressed in songbirds, researchers can begin to gain insight into the mechanisms through which those gene products may impact cellular function to affect vocal performance (Burkett et al. 2018; Lovell et al. 2018; but see Mueller et al. 2016).

Gene expression varies throughout song learning and development (Olson et al. 2015; Burkett et al. 2018; London, Chap. 8). Manipulating the expression of specific gene products in focal sites (e.g., FOXP2 expression in Area X) during developmental sensorimotor learning can induce incomplete and inaccurate imitations of the tutor song (Haesler et al. 2007) and affect adult song control (Murugan et al. 2013). Because songbirds acquire their songs in a pattern that is strikingly similar to how humans learn the sounds used in speech (Prather et al. 2017), additional studies investigating the role of genes (e.g., FOXP2) in vocal learning deficits can lead to translational applications to human mutations and pathologies. In addition to that clinical relevance, genetic manipulation in songbirds can also be used to discern the mechanisms that regulate the divergence of sexually dimorphic features during development (Zhao et al. 2018). These tools will help researchers continue to reveal not only how insights from songbirds can help to advance human health but also how comparative studies may provide a deeper understanding of how individual gene products and their interactions shape brain development and function (Konopka and Roberts 2016; Mets and Brainard 2018).

3.6 Important Questions and Next Steps

The songbird brain provides an excellent context in which to investigate the neural basis of learned vocal communication. With the advent of many new tools and the strength of the research that has come before, future studies will yield unprecedented insight into how the brain enables an organism to perceive, learn, memorize, recall, and perform a series of complex individual behaviors in a specific sequence.

Those broad themes can be reframed as a series of important questions, each of which are briefly considered in the following sections.

3.6.1 How Does the Brain Serialize Individual Behaviors Into Meaningful Sequences?

The sequencing of individual behavioral units into behaviorally meaningful sequences is integral to many forms of behavior, and that is especially evident in the sounds used in vocal communication. Just as the content of this sentence would be altered if the words or letters were shuffled in their sequence, song syntax would be disrupted if the sequences of notes or phrases were rearranged. That impact is also evident in the activity of individual neurons: neurons in the vocal communication system are more responsive to songs with notes played in their natural sequence (Lewicki 1996; Prather et al. 2008). Similarly, female responses to song are altered when the sequencing of song elements is manipulated: females respond most robustly to the song of their preferred male when his notes are sequenced in his natural order (Dunning and Prather, in preparation).

As noted in Sect. 3.3.4, studies of how the songbird brain controls the phonology and sequencing of individual syllables have yielded varying results. Some findings have suggested that sequencing and phonology are encoded in separate pathways in the brain, while other data suggest that sequence and phonology are encoded through similar circuits. Additional studies incorporating a range of different techniques will be necessary to discern the amount of overlap in the circuits responsible for controlling the phonology and sequencing of individual song syllables. For example, researchers can take advantage of tools to alter singing-related auditory feedback by changing the pitch of that feedback, the timing of that information, or some combination of spectral and temporal changes (Sober and Brainard 2009; Wyatt et al. 2017). By altering different aspects of the feedback that a bird experiences during singing, researchers can induce different forms of adaptive changes in phonology or sequence, thus providing an experimental context that will be helpful in disambiguating the neural control of each of those aspects of learned song performance (Warren et al. 2012; Tian and Brainard 2017).

Insights gained from adult birds will be especially useful in guiding studies of juvenile birds undergoing song development. By comparing the patterns of neural activity detected in juvenile birds undergoing natural song variability (i.e., the *plastic song* phase of late development; Mooney et al. 2008) versus adult birds undergoing feedback-induced song variability, one can gain insight into the degree of similarity in neural substrates underlying these forms of vocal plasticity (Hisey et al. 2018).

3.6.2 How Does the Brain Compute and Use Error Signals to Refine Behavior?

The importance of auditory input is clear in the case of imitative vocal learning in juvenile birds, and auditory feedback continues to play a central role in the refinement of vocal behaviors and their maintenance throughout adulthood (Tschida and Mooney 2012; Murphy et al. 2017). These findings have led researchers to the following central questions: How does the brain compare current performance against a memorized model to guide the development and maintenance of vocal signals, and how does the brain compare motor commands against the associated auditory feedback to enable the refinement and maintenance of vocal behavior? By taking advantage of new tools, such as large-scale array recordings and imaging neural ensembles, songbird researchers will be able to continue resolving the neural circuits involved in the comparison of those signals. With the knowledge of how errors are computed and represented in specific patterns of neural activity, researchers will then be able to investigate the mechanisms through which error signals direct changes in the structures and functions of downstream motor pathways to enable adaptation (Gadagkar et al. 2016; Hisey et al. 2018). Insights into how errors are detected and encoded in the songbird brain will provide a search image for understanding how those processes also occur in the human brain, how sensorimotor interactions enable acquisition of new abilities, and how disruption of activity in those circuits may lead to the emergence of behavioral pathologies.

3.6.3 What Is the Neural Substrate for Learning and Memory?

Before a tutor song model can be recalled and used to guide the development of vocal imitation, a memory of that song must be formed and stored. A long-standing set of questions in the songbird field have been: Where is the memory of a tutor song stored, through what mechanisms and patterns of activity is that memory recalled, and how does that recalled memory act to shape subsequent vocal performances? Songbirds offer many natural behavioral advantages to address these questions. For example, in some species, the period of sensory learning (when the tutor song is committed to memory) is separated by many months from the period of sensorimotor learning (when the bird begins vocalizing and using feedback to refine its imitation of that model) (Mooney et al. 2008; Sakata and Woolley, Chap. 1). That natural pattern of vocal development provides a context in which to disambiguate the formation and storage of a memory versus the recall and use of that memory to guide ongoing behavior. In addition, some species sing a repertoire of song types that are acoustically distinct (Krebs and Kroodsma 1980), providing a context in which to identify how different tutor song models are represented as distinct engrams in the brain (Mooney et al. 2001; Prather et al. 2010). With the advent of new tools to enable researchers to record from individual neurons over extended periods of time, researchers may be able to monitor activity throughout juvenile development and thus investigate the role of specific neurons in memory formation throughout sensory learning and memory recall during sensorimotor learning. Together, the experimental tractability of the songbird vocal communication system and the variety of investigations that are possible in that context highlight the value and broad relevance of songbird neurobiology for understanding the neural basis of learning and memory.

3.6.4 How Can Studying Songbirds Inform Understanding of the Human Brain?

Research has revealed that properties of the avian auditory cortex are more like those of mammalian neocortex than was previously appreciated. Specifically, a study revealed that neurons in a region of the chicken brain that corresponds to the mammalian auditory cortex (a complex including Field L and CM) are arranged in layers in which auditory neurons interact via radially arranged intrinsic connections (Wang et al. 2010). These layers do not correspond precisely to those that have been described in the mammalian neocortex, but they do reveal that defining characteristics of the neocortex are also present in the avian brain (Wild and Krutzfeldt 2010; Woolley and Woolley, Chap. 5).

Additional studies have also shown that, as in the human brain, lateralization of function occurs in the songbird brain (Long et al. 2016; Prather et al. 2017), and these findings may provide insights into the ways that factors such as fluency and early experience influence the neural substrate for language. For example, in birds that are sequentially tutored with two different songs, the pattern of lateralized dominance depends on the proficiency of learning from each tutor in each epoch of development (Olson et al. 2016). The greater the retention of song from their first tutor, the more right dominant the birds are when they are exposed to that song, which is indicated by greater EGR-1 expression in the right hemisphere than the left in response to song played through a speaker. The more the birds learned from their second tutor, the more left dominant they are when they are exposed to that song. Thus, song memories are preserved in a lateralized manner that is dependent on the proficiency of song learning (Olson et al. 2016). These similarities between features of songbird and human brains provide further motivation to use songbirds as an animal model in which to explore how auditory memories are stored, how they are recalled, and how that information is blended with ongoing sensory feedback to influence vocal motor activity.

3.7 Chapter Summary

Following the discovery of a neural network that is specialized for vocal learning in the songbird brain, researchers have sought to understand how that song system enables memorization of a tutor song model followed by comparison of that memory against ongoing behavior to refine and maintain behavior. With the realization that some components of that network are analogous to cortical, striatal, and thalamic areas in the mammalian brain, birdsong research has the potential to help researchers understand not only how the nervous system enables experience-guided learning and memory but also how pathologies of the central nervous system afflict human cognition and behavior.

Acknowledgement This work was supported by NSF IOS CAREER (1453084) to JFP.

Compliance With Ethics Requirements Karagh Murphy declares that she has no conflict of interest.

Koedi S. Lawley declares that she has no conflict of interest.

Perry Smith declares that he has no conflict of interest.

Jonathan F. Prather declares that he has no conflict of interest.

References

Abe K, Matsui S, Watanabe D (2015) Transgenic songbirds with suppressed or enhanced activity of CREB transcription factor. Proc Natl Acad Sci U S A 112(24):7599–7604. https://doi.org/10.1073/pnas.1413484112

Adret P, Meliza CD, Margoliash D (2012) Song tutoring in presinging zebra finch juveniles biases a small population of higher-order song-selective neurons toward the tutor song. J Neurophysiol 108(7):1977–1987. https://doi.org/10.1152/jn.00905.2011

Agate RJ, Scott BB, Haripal B, Lois C, Nottebohm F (2009) Transgenic songbirds offer an opportunity to develop a genetic model for vocal learning. Proc Natl Acad Sci U S A 106(42):17963–17967. https://doi.org/10.1073/pnas.0909139106

Akutagawa E, Konishi M (2010) New brain pathways found in the vocal control system of a songbird. J Comp Neurol 518(15):3086–3100. https://doi.org/10.1002/cne.22383

Ali F, Otchy TM, Pehlevan C, Fantana AL, Burak Y, Olveczky BP (2013) The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. Neuron 80(2)):494–506. https://doi.org/10.1016/j.neuron.2013.07.049

Alward BA, Balthazart J, Ball GF (2017) Dissociable effects on birdsong of androgen signaling in cortex-like brain regions of canaries. J Neurosci 37(36):8612–8624. https://doi.org/10.1523/JNEUROSCI.3371-16.2017

Amador A, Perl YS, Mindlin GB, Margoliash D (2013) Elemental gesture dynamics are encoded by song premotor cortical neurons. Nature 495(7439):59–64. https://doi.org/10.1038/nature11967

Andalman AS, Foerster JN, Fee MS (2011) Control of vocal and respiratory patterns in birdsong: dissection of forebrain and brainstem mechanisms using temperature. PLoS One 6(9):e25461. https://doi.org/10.1371/journal.pone.0025461

Aronov D, Andalman AS, Fee MS (2008) A specialized forebrain circuit for vocal babbling in the juvenile songbird. Science 320(5876):630–634. https://doi.org/10.1126/science.1155140

Basista MJ, Elliott KC, Wu W, Hyson RL, Bertram R, Johnson F (2014) Independent premotor encoding of the sequence and structure of birdsong in avian cortex. J Neurosci 34(50):16821–16834. https://doi.org/10.1523/JNEUROSCI.1940-14.2014

Bauer EE, Coleman MJ, Roberts TF, Roy A, Prather JF, Mooney R (2008) A synaptic basis for auditory-vocal integration in the songbird. J Neurosci 28(6):1509–1522. https://doi.org/10.1523/JNEUROSCI.3838-07.2008

Benichov JI, Benezra SE, Vallentin D, Globerson E, Long MA, Tchernichovski O (2016) The forebrain song system mediates predictive call timing in female and male Zebra finches. Curr biol 26(3):309–318. https://doi.org/10.1016/j.cub.2015.12.037

Bolhuis JJ, Everaert M (eds) (2013) Birdsong, speech and language: exploring the evolution of mind and brain. MIT Press, Cambridge

Bolhuis JJ, Gobes SM, Terpstra NJ, den Boer-Visser AM, Zandbergen MA (2012) Learning-related neuronal activation in the zebra finch song system nucleus HVC in response to the bird's own song. PLoS One 7(7):e41556. https://doi.org/10.1371/journal.pone.0041556

Bottjer SW, Miesner EA, Arnold AP (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. Science 224(4651):901–903. https://doi.org/10.1126/science.6719123

Brenowitz EA (1991) Altered perception of species-specific song by female birds after lesions of a forebrain nucleus. Science 251(4991):303–305. https://doi.org/10.1126/science.1987645

Burkett ZD, Day NF, Kimball TH, Aamodt CM, Heston JB, Hilliard AT, Xiao X, White SA (2018) FoxP2 isoforms delineate spatiotemporal transcriptional networks for vocal learning in the zebra finch. Elife 7. https://doi.org/10.7554/eLife.30649

Calabrese A, Woolley SM (2015) Coding principles of the canonical cortical microcircuit in the avian brain. Proc Natl Acad Sci U S A. 12(11):3517–3522. https://doi.org/10.1073/pnas.1408545112

Cardin JA, Schmidt MF (2004) Auditory responses in multiple sensorimotor song system nuclei are co-modulated by behavioral state. J Neurophys 91(5):2148–2163. https://doi.org/10.1152/jn.00918.2003

Cardin JA, Raksin JN, Schmidt MF (2005) Sensorimotor nucleus NIf is necessary for auditory processing but not vocal motor output in the avian song system. J Neurophysiol 93(4):2157–2166. https://doi.org/10.1152/jn.01001.2004

Charlesworth JD, Warren TL, Brainard MS (2012) Covert skill learning in a cortical-basal ganglia circuit. Nature 486(7402):251–255. https://doi.org/10.1038/nature11078

Chen Y, Matheson LE, Sakata JT (2016) Mechanisms underlying the social enhancement of vocal learning in songbirds. Proc Natl Acad Sci U S A 113(24):6641–6646. https://doi.org/10.1073/pnas.1522306113

Coleman MJ, Mooney R (2004) Synaptic transformations underlying highly selective auditory representations of learned birdsong. J Neurosci 24(33):7251–7265. https://doi.org/10.1523/JNEUROSCI.0947-04.2004

Coleman MJ, Vu ET (2005) Recovery of impaired songs following unilateral but not bilateral lesions of nucleus uvaeformis of adult zebra finches. J Neurobiol 10.1002/neu.2012263(1):70–89. https://doi.org/10.1002/neu.20122

Coleman MJ, Roy A, Wild JM, Mooney R (2007) Thalamic gating of auditory responses in telencephalic song control nuclei. J Neurosci 27(37):10024–10036. https://doi.org/10.1523/JNEUROSCI.2215-07.2007

Deisseroth K (2015) Optogenetics: 10 years of microbial opsins in neuroscience. Nat Neurosci, 18 (9):1213–1225. https://doi.org/10.1038/nn.4091

Diez A, Cui A, MacDougall-Shackleton SA (2017) The neural response of female zebra finches (Taeniopygia guttata) to conspecific, heterospecific, and isolate song depends on early-life song exposure. Behav Processes. https://doi.org/10.1016/j.beproc.2017.12.022

Dunning JL, Maze SE, Atwood EJ, Prather JF (2018) Caudal mesopallial neurons in female songbirds bridge sensory and motor brain regions. J Comp Neurol 526(10):1703–1711. https://doi.org/10.1002/cne.24440

Fee MS, Leonardo A (2001) Miniature motorized microdrive and commutator system for chronic neural recording in small animals. J Neurosci Methods 112(2):83–94. https://doi.org/10.1016/s0165-0270(01)00426-5

Fee MS, Kozhevnikov AA, Hahnloser RH (2004) Neural mechanisms of vocal sequence generation in the songbird. Ann N Y Acad Sci 1016:153–170. https://doi.org/10.1196/annals.1298.022

Gadagkar V, Puzerey PA, Chen R, Baird-Daniel E, Farhang AR, Goldberg JH (2016) Dopamine neurons encode performance error in singing birds. Science 354(6317):1278–1282. https://doi.org/10.1126/science.aah6837

Galvis D, Wu W, Hyson RL, Johnson F, Bertram R (2017) A distributed neural network model for the distinct roles of medial and lateral HVC in zebra finch song production. J Neurophysiol 118(2):677–692. https://doi.org/10.1152/jn.00917.2016

Gentner TQ, Margoliash D (2003) Neuronal populations and single cells representing learned auditory objects. Nature 424(6949):669–674. https://doi.org/10.1038/nature01731

Gentner TQ, Hulse SH, Bentley GE, Ball GF (2000) Individual vocal recognition and the effect of partial lesions to HVC on discrimination, learning, and categorization of conspecific song in adult songbirds. J Neurobiol 42(1):117–133. https://doi.org/10.1002/(SICI)1097-4695(200001)42:1<117::AID-NEU11>3.0.CO;2-M

Gentner TQ, Hulse SH, Ball GF (2004) Functional differences in forebrain auditory regions during learned vocal recognition in songbirds. J comp Physiol A Neuroethol Sens Neural Behav Physiol 190(12):1001–1010. https://doi.org/10.1007/s00359-004-0556-x

Gobes SM, Bolhuis JJ (2007) Birdsong memory: a neural dissociation between song recognition and production. Curr Biol 17(9):789–793. https://doi.org/10.1016/j.cub.2007.03.059

Goldberg JH, Fee MS (2011) Vocal babbling in songbirds requires the basal ganglia-recipient motor thalamus but not the basal ganglia. J Neurophysiol 105(6):2729–2739. https://doi.org/10.1152/jn.00823.2010

Goldberg JH, Farries MA, Fee MS (2012) Integration of cortical and pallidal inputs in the basal ganglia-recipient thalamus of singing birds. J Neurophysiol. https://doi.org/10.1152/jn.00056.2012

Guitchounts G, Markowitz JE, Liberti WA, Gardner TJ (2013) A carbon-fiber electrode array for long-term neural recording. J Neural Eng 10(4):046016. https://doi.org/10.1088/1741-2560/10/4/046016

Haesler S, Rochefort C, Georgi B, Licznerski P, Osten P, Scharff C (2007) Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus area X. PLoS Biol 5(12):e321. https://doi.org/10.1371/journal.pbio.0050321

Hahnloser RH, Kozhevnikov AA, Fee MS (2002) An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature 419(6902):65–70. https://doi.org/10.1038/nature00974

Hamaguchi K, Mooney R (2012) Recurrent interactions between the input and output of a songbird cortico-basal ganglia pathway are implicated in vocal sequence variability. J Neurosci 32(34):11671–11687. https://doi.org/10.1523/JNEUROSCI.1666-12.2012

Hamaguchi K, Tschida KA, Yoon I, Donald BR, Mooney R (2014) Auditory synapses to song premotor neurons are gated off during vocalization in zebra finches. Elife 3:e01833. https://doi.org/10.7554/eLife.01833

Hamaguchi K, Tanaka M, Mooney R (2016) A distributed recurrent network contributes to temporally precise vocalizations. Neuron. Neuron. 91(3):680–693. https://doi.org/10.1016/j.neuron.2016.06.019

Hampton CM, Sakata JT, Brainard MS (2009) An avian basal ganglia-forebrain circuit contributes differentially to syllable versus sequence variability of adult Bengalese finch song. J Neurophysiol 101(6):3235–3245. https://doi.org/10.1152/jn.91089.2008

Hasegawa T, Fujimoto H, Tashiro K, Nonomura M, Tsuchiya A, Watanabe D (2015) A wireless neural recording system with a precision motorized microdrive for freely behaving animals. Sci Rep 5(7853). https://doi.org/10.1038/srep07853

Hisey E, Kearney MG, Mooney R (2018) A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. Nat Neurosci 21(4):589–597. https://doi.org/10.1038/s41593-018-0092-6

Hosino T, Okanoya K (2000) Lesion of a higher-order song nucleus disrupts phrase level complexity in Bengalese finches. Neuroreport 11(10):2091–2095. https://doi.org/10.1097/00001756-200007140-00007

James LS, Sakata JT (2017) Learning biases underlie "universals" in avian vocal sequencing. Curr biol 27(23):3676–3682. e3674. https://doi.org/10.1016/j.cub.2017.10.019

Jarvis E, Gunturkun O, Bruce L, Csillag A, Karten H, Kuenzel W, Medina L, Paxinos G, Perkel DJ, Shimizu T, Striedter G, Wild JM, Ball GF, Dugas-Ford J, Durand SE, Hough GE, Husband S, Kubikova L, Lee DW, Mello CV, Powers A, Siang C, Smulders TV, Wada K, White SA, Yamamoto K, Yu J, Reiner A, Butler AB, Consorti ABN (2005) Avian brains and a new understanding of vertebrate brain evolution. Nat Rev Neurosci 6(2):151–159. https://doi.org/10.1038/nrn1606

Jarvis ED, Nottebohm F (1997) Motor-driven gene expression. Proc Natl Acad Sci U S A 94(8):4097–4102. https://doi.org/10.1073/pnas.94.8.4097

Jeanne JM, Thompson JV, Sharpee TO, Gentner TQ (2011) Emergence of learned categorical representations within an auditory forebrain circuit. J Neurosci 31(7):2595–2606. https://doi.org/10.1523/JNEUROSCI.3930-10.2011

Jovalekic A, Cave-Lopez S, Canopoli A, Ondracek JM, Nager A, Vyssotski AL, Hahnloser RH (2017) A lightweight feedback-controlled microdrive for chronic neural recordings. J neural Eng 14(2):026006. https://doi.org/10.1088/1741-2552/aa5848

Katlowitz KA, Picardo MA, Long MA (2018) Stable sequential activity underlying the maintenance of a precisely executed skilled behavior. Neuron 98(6):1133–1140. e1133. https://doi.org/10.1016/j.neuron.2018.05.017

Katz LC, Gurney ME (1981) Auditory responses in the zebra finch's motor system for song. Brain Res 221(1):192–197. https://doi.org/10.1016/0006-8993(81)91073-8

Konopka G, Roberts TF (2016) Insights into the neural and genetic basis of vocal communication. Cell 164(6):1269–1276. https://doi.org/10.1016/j.cell.2016.02.039

Krebs JR, Kroodsma DE (1980) Repertoires and geographical variation in bird song. Adv Study Behav 11:143–177. https://doi.org/10.1016/S0065-3454(08)60117-5

Lampen J, McAuley JD, Chang SE, Wade J (2017) Neural activity associated with rhythmicity of song in juvenile male and female zebra finches. Behav Processes. https://doi.org/10.1016/j.beproc.2017.12.003

Lewandowski BC, Schmidt M (2011) Short bouts of vocalization induce long-lasting fast gamma oscillations in a sensorimotor nucleus. J Neurosci 31(39):13936–13948. https://doi.org/10.1523/JNEUROSCI.6809-10.2011

Lewandowski B, Vyssotski A, Hahnloser RH, Schmidt M (2013) At the interface of the auditory and vocal motor systems: NIf and its role in vocal processing, production and learning. J Physiol Paris 107(3):178–192. https://doi.org/10.1016/j.jphysparis.2013.04.001

Lewicki MS (1996) Intracellular characterization of song-specific neurons in the zebra finch auditory forebrain. J Neurosci 16(18):5855–5863. https://doi.org/10.1523/JNEUROSCI.16-18-05854.1996

Liberti WA, Perkins LN, Leman DP, Gardner TJ (2017) An open source, wireless capable miniature microscope system. J Neural Eng 14(4). https://doi.org/10.1088/1741-2552/aa6806

Liu WC, Kohn J, Szwed SK, Pariser E, Sepe S, Haripal B, Oshimori N, Marsala M, Miyanohara A, Lee R (2015) Human mutant huntingtin disrupts vocal learning in transgenic songbirds. Nat Neurosci 18(11):1617–1622. https://doi.org/10.1038/nn.4133

London SE, Clayton DF (2008) Functional identification of sensory mechanisms required for developmental song learning. Nat Neurosci 11(5):579–586. https://doi.org/10.1038/nn.2103

Long MA, Fee MS (2008) Using temperature to analyse temporal dynamics in the songbird motor pathway. Nature 456(7219):189–194. https://doi.org/10.1038/nature07448

Long MA, Jin DZ, Fee MS (2010) Support for a synaptic chain model of neuronal sequence generation. Nature 468(7322):394–399. https://doi.org/10.1038/nature09514

Long MA, Katlowitz KA, Svirsky MA, Clary RC, Byun TM, Majaj N, Oya H, Howard MA 3rd, Greenlee JDW (2016) Functional segregation of cortical regions underlying speech timing and articulation. Neuron 89(6):1187–1193. https://doi.org/10.1016/j.neuron.2016.01.032

Lovell PV, Huizinga NA, Friedrich SR, Wirthlin M, Mello CV (2018) The constitutive differential transcriptome of a brain circuit for vocal learning. BMC Genomics 19(1):231. https://doi.org/10.1186/s12864-018-4578-0

Lynch GF, Okubo TS, Hanuschkin A, Hahnloser RH, Fee MS (2016) Rhythmic continuous-time coding in the songbird analog of vocal motor cortex. Neuron 90(4):877–892. https://doi.org/10.1016/j.neuron.2016.04.021

Mandelblat-Cerf Y, Fee MS (2014) An automated procedure for evaluating song imitation. Plos One 9(5):e96484. https://doi.org/10.1371/journal.pone.0096484

McCasland JS (1987) Neuronal control of bird song production. J Neurosci 7(1):23–39. https://doi.org/10.1523/JNEUROSCI.07-01-00023

McCasland JS, Konishi M (1981) Interaction between auditory and motor activities in an avian song control nucleus. Proc Natl Acad Sci U S A 78(12):7815–7819. https://doi.org/10.1073/pnas.78.12.7815

Meliza CD, Margoliash D (2012) Emergence of selectivity and tolerance in the avian auditory cortex. J Neurosci 32(43):15158–15168. https://doi.org/10.1523/JNEUROSCI.0845-12.2012

Meliza CD, Chi Z, Margoliash D (2010) Representations of conspecific song by starling secondary forebrain auditory neurons: toward a hierarchical framework. J Neurophysiol 103(3):1195–1208. https://doi.org/10.1152/jn.00464.2009

Mello CV, Vicario DS, Clayton DF (1992) Song presentation induces gene-expression in the songbird forebrain. PNAS 89(15):6818–6822. https://doi.org/10.1073/pnas.89.15.6818

Menardy F, Touiki K, Dutrieux G, Bozon B, Vignal C, Mathevon N, Del Negro C (2012) Social experience affects neuronal responses to male calls in adult female zebra finches. Eur J Neurosci 35(8):1322–1336. https://doi.org/10.1111/j.1460-9568.2012.08047.x

Mencio C, Kuberan B, Goller F (2017) Contributions of rapid neuromuscular transmission to the fine control of acoustic parameters of birdsong. J Neurophysiol 117(2):637–645. https://doi.org/10.1152/jn.00843.2015

Mets DG, Brainard MS (2018) Genetic variation interacts with experience to determine interindividual differences in learned song. Proc Natl Acad Sci U S A 115(2):421–426. https://doi.org/10.1073/pnas.1713031115

Miller MN, Cheung CYJ, Brainard MS (2017) Vocal learning promotes patterned inhibitory connectivity. Nat Commun 8(1):2105. https://doi.org/10.1038/s41467-017-01914-5

Mooney R, Hoese W, Nowicki S (2001) Auditory representation of the vocal repertoire in a songbird with multiple song types. Proc Natl Acad Sci U S A 98(22):12778–12783. https://doi.org/10.1073/pnas.221453298

Mooney R, Prather JF, Roberts T (2008) Neurophysiology of birdsong learning. In: Eichenbaum H (ed) Learning and memory: a comprehensive reference, Memory systems, vol 3. Elsevier, Oxford, pp 441–474

Moseley DL, Joshi NR, Prather JF, Podos J, Remage-Healey L (2017) A neuronal signature of accurate imitative learning in wild-caught songbirds (swamp sparrows, Melospiza georgiana). Sci Rep 7(1):17320. https://doi.org/10.1038/s41598-017-17401-2

Mouterde SC, Elie JE, Mathevon N, Theunissen FE (2017) Single neurons in the avian auditory cortex encode individual identity and propagation distance in naturally degraded communication calls. J Neurosci 37(13):3491–3510. https://doi.org/10.1523/JNEUROSCI.2220-16.2017

Mueller KL, Murray JC, Michaelson JJ, Christiansen MH, Reilly S, Tomblin JB (2016) Common genetic variants in FOXP2 are not associated with individual differences in language development. Plos One 11(4):e0152576. https://doi.org/10.1371/journal.pone.0152576

Murphy K, Prather JF (2016) Representation of calls in the activity of neurons in the songbird premotor nucleus HVC. In: *2016* Society for Neuroscience annual meeting, San Diego

Murphy K, James LS, Sakata JT, Prather JF (2017) Advantages of comparative studies in songbirds to understand the neural basis of sensorimotor integration. J Neurophysiol 118(2):800–816. https://doi.org/10.1152/jn.00623.2016

Murugan M, Harward S, Scharff C, Mooney R (2013) Diminished FoxP2 levels affect dopaminergic modulation of corticostriatal signaling important to song variability. Neuron 80(6):1464–1476. https://doi.org/10.1016/j.neuron.2013.09.021

Nottebohm F, Stokes TM, Leonard CM (1976) Central control of song in the canary, Serinus canarius. J Comp Neurol 165(4):457–486. https://doi.org/10.1002/cne.901650405

Okubo TS, Mackevicius EL, Payne HL, Lynch GF, Fee MS (2015) Growth and splitting of neural sequences in songbird vocal development. Nature 528(7582):352–357. https://doi.org/10.1038/nature15741

Olson CR, Hodges LK, Mello CV (2015) Dynamic gene expression in the song system of zebra finches during the song learning period. Dev Neurobiol 75(12):1315–1338. https://doi.org/10.1002/dneu.22286

Olson EM, Maeda RK, Gobes SM (2016) Mirrored patterns of lateralized neuronal activation reflect old and new memories in the avian auditory cortex. Neurosci 330:395–402. https://doi.org/10.1016/j.neuroscience.2016.06.009

Olveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS biol 3 (5):e153. https://doi.org/10.1371/journal.pbio.0030153

Olveczky BP, Otchy TM, Goldberg JH, Aronov D, Fee MS (2011) Changes in the neural control of a complex motor sequence during learning. J Neurophysiol 106(1):386–397. https://doi.org/10.1152/jn.00018.2011

Otchy TM, Wolff SB, Rhee JY, Pehlevan C, Kawai R, Kempf A, Gobes SM, Ölveczky BP (2015) Acute off-target effects of neural circuit manipulations. Nature 528(7582):358–363. https://doi.org/10.1038/nature16442

Pawlisch BA, Remage-Healey L (2015) Neuroestrogen signaling in the songbird auditory cortex propagates into a sensorimotor network via an 'interface' nucleus. Neurosci 284:522–535. https://doi.org/10.1016/j.neuroscience.2014.10.023

Peh WY, Roberts TF, Mooney R (2015) Imaging auditory representations of song and syllables in populations of sensorimotor neurons essential to vocal communication. J Neurosci 35(14):5589–5605. https://doi.org/10.1523/JNEUROSCI.2308-14.2015

Picardo MA, Merel J, Katlowitz KA, Vallentin D, Okobi DE, Benezra SE, Clary RC, Pnevmatikakis EA, Paninski L, Long MA (2016) Population-level representation of a temporal sequence underlying song production in the Zebra finch. Neuron 90(4):866–876. https://doi.org/10.1016/j.neuron.2016.02.016

Piristine HC, Choetso T, Gobes SM (2016) A sensorimotor area in the songbird brain is required for production of vocalizations in the song learning period of development. Dev Neurobiol 76(11):1213–1225. https://doi.org/10.1002/dneu.22384

Prather JF (2013) Auditory signal processing in communication: perception and performance of vocal sounds. Hear Res 305:144–155. https://doi.org/10.1016/j.heares.2013.06.007

Prather JF, Mooney R (2004) Neural correlates of learned song in the avian forebrain: simultaneous representation of self and others. Curr Opin Neurobiol 14(4):496–502. https://doi.org/10.1016/j.conb.2004.06.004

Prather JF, Peters S, Nowicki S, Mooney R (2008) Precise auditory-vocal mirroring in neurons for learned vocal communication. Nature 451(7176):305–310. https://doi.org/10.1038/nature06492

Prather JF, Peters S, Nowicki S, Mooney R (2010) Persistent representation of juvenile experience in the adult songbird brain. J Neurosci 30(31):10586–10598. https://doi.org/10.1523/JNEUROSCI.6042-09.2010

Prather JF, Okanoya K, Bolhuis JJ (2017) Brains for birds and babies: neural parallels between birdsong and speech acquisition. Neurosci Biobehav rev 81(Pt B):225–237. https://doi.org/10.1016/j.neubiorev.2016.12.035

Pratt KG, Prather JF (2016) Systems neuroscience: how the cortex contributes to skilled movements. Curr Biol 26(2):64–66. https://doi.org/10.1016/j.cub.2015.11.044

Reiner A, Perkel DJ, Bruce LL, Butler AB, Csillag A, Kuenzel W, Medina L, Paxinos G, Shimizu T, Striedter G, Wild M, Ball GF, Durand S, Gunturkun O, Lee DW, Mello CV, Powers A, White SA, Hough G, Kubikova L, Smulders TV, Wada K, Dugas-Ford J, Husband S, Yamamoto K, Yu J, Siang C, Jarvis ED (2004) Revised nomenclature for avian telencephalon and some related brainstem nuclei. J Comp Neurol 473(3):377–414. https://doi.org/10.1002/cne.20118

Roberts TF, Tschida KA, Klein ME, Mooney R (2010) Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning. Nature 463(7283):948–952. https://doi.org/10.1038/nature08759

Roberts TF, Gobes SM, Murugan M, Olveczky BP, Mooney R (2012) Motor circuits are required to encode a sensory model for imitative learning. Nat Neurosci 15(10):1454–1459. https://doi.org/10.1038/nn.3206

Roberts TF, Hisey E, Tanaka M, Kearney MG, Chattree G, Yang CF, Shah NM, Mooney R (2017) Identification of a motor-to-auditory pathway important for vocal learning. Nat Neurosci 20(7):978–986. https://doi.org/10.1038/nn.4563

Scharff C, Nottebohm F (1991) A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. J Neurosci 11(9):2896–2913. https://doi.org/10.1523/JNEUROSCI.11-09-02896.1991

Schmidt MF, Wild JM (2014) The respiratory-vocal system of songbirds: anatomy, physiology, and neural control. In: Breathing, emotion and evolution, vol 212, pp 297–335. https://doi.org/10.1016/B978-0-444-63488-7.00015-X

Schregardus DS, Pieneman AW, Ter Maat A, Jansen RF, Brouwer TJ, Gahr ML (2006) A lightweight telemetry system for recording neuronal activity in freely behaving small animals. J Neurosci Methods 155(1):62–71. https://doi.org/10.1016/j.jneumeth.2005.12.028

Simpson HB, Vicario DS (1990) Brain pathways for learned and unlearned vocalizations differ in zebra finches. J Neurosci 10(5):1541–1556. https://doi.org/10.1523/JNEUROSCI.10-05-01541.1990

Skeide MA, Friederici AD (2016) The ontogeny of the cortical language network. Nat Rev Neurosci 17(5):323–332. https://doi.org/10.1038/nrn.2016.23

Sober SJ, Brainard MS (2009) Adult birdsong is actively maintained by error correction. Nature Neuroscience 12(7):927–931. https://doi.org/10.1038/nn.2336

Sober SJ, Wohlgemuth MJ, Brainard MS (2008) Central contributions to acoustic variation in birdsong. J. Neuroscience, 28(41):10370–10379. https://doi.org/10.1523/JNEUROSCI.2448-08.2008

Sohrabji F, Nordeen EJ, Nordeen KW (1990) Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. Behav Neural Biol 53(1):51–63. https://doi.org/10.1016/0163-1047(90)90797-A

Srivastava KH, Holmes CM, Vellema M, Pack AR, Elemans CP, Nemenman I, Sober SJ (2017) Motor control by precisely timed spike patterns. Proc Natl Acad Sci U S A 114(5):1171–1176. https://doi.org/10.1073/pnas.1611734114

Tanaka M, Singh Alvarado J, Murugan M, Mooney R (2016) Focal expression of mutant huntingtin in the songbird basal ganglia disrupts cortico-basal ganglia networks and vocal sequences. Proc Natl Acad Sci U S A 113 (12):E1720–1727. Doi: https://doi.org/10.1073/pnas.1523754113

Tanaka M, Sun F, Li Y, Mooney R (2018) A mesocortical dopamine circuit enables the cultural transmission of vocal behavior. Nature 563:117–120. https://doi.org/10.1038/s41586-018-0636-7

ten Cate C (2018) The comparative study of grammar learning mechanisms: birds as models. Curr Opin. Behav Sci 21:13–18. https://doi.org/10.1016/j.cobeha.2017.11.008

Ter Maat A, Trost L, Sagunsky H, Seltmann S, Gahr M (2014) Zebra finch mates use their forebrain song system in unlearned call communication. Plos One 9(10):e109334. https://doi.org/10.1371/journal.pone.0109334

Terpstra NJ, Bolhuis JJ, den Boer-Visser AM (2004) An analysis of the neural representation of birdsong memory. J Neurosci 24(21):4971–4977. https://doi.org/10.1523/JNEUROSCI.0570-04.2004

Thompson JV, Gentner TQ (2010) Song recognition learning and stimulus-specific weakening of neural responses in the avian auditory forebrain. J Neurophysiol 103(4):1785–1797. https://doi.org/10.1152/jn.00885.2009

Tian LY, Brainard MS (2017) Discrete circuits support generalized versus context-specific vocal learning in the songbird. Neuron 96(5):1168–1177. e1165. https://doi.org/10.1016/j.neuron.2017.10.019

Troyer TW, Doupe AJ (2000) An associational model of birdsong sensorimotor learning I. Efference copy and the learning of song syllables. J Neurophysiol 84(3):1204–1223. https://doi.org/10.1152/jn.2000.84.3.1204

Tschida K, Mooney R (2012) The role of auditory feedback in vocal learning and maintenance. Curr Opin Neurobiol 22(2):320–327. https://doi.org/10.1016/j.conb.2011.11.006

Urbano CM, Aston AE, Cooper BG (2016) HVC contributes toward conspecific contact call responding in male Bengalese finches. Neuroreport 27(7):481–486. https://doi.org/10.1097/WNR.0000000000000567

Vates GE, Broome BM, Mello CV, Nottebohm F (1996) Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches. J Comp Neurol 366(4):613–642. https://doi.org/10.1002/(SICI)1096-9861(19960318)366:4<613::AID-CNE5>3.0.CO;2-7

Wang Y, Brzozowska-Prechtl A, Karten HJ (2010) Laminar and columnar auditory cortex in avian brain. Proc Natl Acad Sci U S A 107(28):12676–12681. https://doi.org/10.1073/pnas.1006645107

Warren TL, Charlesworth JD, Tumer EC, Brainard MS (2012) Variable sequencing is actively maintained in a well learned motor skill. J Neurosci 32(44):15414–15425. https://doi.org/10.1523/JNEUROSCI.1254-12.2012

Wild JM, Botelho JF (2015) Involvement of the avian song system in reproductive behaviour. Biol Lett 11(12):20150773. https://doi.org/10.1098/rsbl.2015.0773

Wild JM, Krutzfeldt NO (2010) Neocortical-like organization of avian auditory 'cortex'. Commentary on Wang Y, Brzozowska-Prechtl A, Karten HJ (2010): Laminar and columnar auditory cortex in avian brain. Proc Natl Acad Sci USA 107:12676–12681. Brain Behav Evol 76(2):89–92. https://doi.org/10.1159/000320215

Wohlgemuth MJ, Sober SJ, Brainard MS (2010) Linked control of syllable sequence and phonology in birdsong. J Neurosci 30(39):12936–12949. https://doi.org/10.1523/JNEUROSCI.2690-10.2010

Wyatt M, Berthiaume EA, Kelly CW, Sober SJ (2017) The effects of pitch shifts on delay-induced changes in vocal sequencing in a songbird. eNeuro 4(1). https://doi.org/10.1523/ENEURO.0254-16.2017

Xiao L, Chattree G, Oscos FG, Cao M, Wanat MJ, Roberts TF (2018) A Basal Ganglia Circuit Sufficient to Guide Birdsong Learning. Neuron. https://doi.org/10.1016/j.neuron.2018.02.020

Yanagihara S, Yazaki-Sugiyama Y (2016) Auditory experience dependent cortical circuit shaping for memory formation in bird song learning. Nat Commun. https://doi.org/10.1038/ncomms11946

Yu AC, Margoliash D (1996) Temporal hierarchical control of singing in birds. Science 273(5283):1871–1875. https://doi.org/10.1126/science.273.5283.1871

Zhang YS, Wittenbach JD, Jin DZ, Kozhevnikov AA (2017) Temperature manipulation in songbird brain implicates the premotor nucleus HVC in birdsong syntax. J Neurosci 37(10):2600–2611. https://doi.org/10.1523/JNEUROSCI.1827-16.2017

Zhao Y, Zhang X, Wang R, Bing J, Wu F, Zhang Y, Xu J, Han Z, Zeng S (2018) Erbin and ErbB2 play roles in the sexual differentiation of the song system nucleus HVC in bengalese finches (Lonchura Striata var. domestica). Dev Neurobiol 78(1):15–38. https://doi.org/10.1002/dneu.22551

Chapter 4
The Song Circuit as a Model of Basal Ganglia Function

Arthur Leblois and David J. Perkel

Abstract Songbirds possess a discrete basal ganglia (BG)-thalamocortical circuit dedicated to song learning and plasticity, which makes them a particularly valuable model for studying the function of basal ganglia in sensorimotor learning. As the basal ganglia are highly conserved across vertebrates, understanding gained in songbirds will generalize to other vertebrate taxa, including humans. Current knowledge about the similarities and differences in the BG circuit in birds and mammals is reviewed at the biochemical, anatomical, and physiological levels to highlight the possible parallels that may be drawn between species and also to reveal the limitations of these parallels. Building on these comparisons, the current hypotheses concerning BG function in mammals and birds are examined in light of current evidence collected in songbirds. Finally, suggestions are made for future experimental and theoretical investigations of BG function that could be conducted in songbirds.

Keywords Area X · Dopamine · Reinforcement learning · Songbird · Striatum · Trial-and-error · VTA · Ventral tegmental area · Vocal learning

---

A. Leblois
Institut des Maladies Neurodégénératives, CNRS, UMR 5293, Université de Bordeaux,
Bordeaux, France
e-mail: arthur.leblois@u-bordeaux.fr

D. J. Perkel (✉)
Departments of Biology and Otolaryngology, University of Washington,
Seattle, Washington, USA
e-mail: Perkel@uw.edu

J. T. Sakata et al. (eds.), *The Neuroethology of Birdsong*, Springer Handbook
of Auditory Research 71, https://doi.org/10.1007/978-3-030-34683-6_4

4.1 Introduction

The basal ganglia (BG; see Table 4.1 for all abbreviations in this chapter) are highly conserved across vertebrates, from lampreys to primates, with similar basic circuits and functional connections (Doupe et al. 2005; Stephenson-Jones et al. 2011). Studying BG function in a variety of species can provide a more general understanding of the conserved nature of this fundamental component of vertebrate brains. Further, species differences in BG function may reflect specializations for particular functions or could reflect divergent evolution (Doupe et al. 2005; Yartsev 2017).

Songbirds possess a set of interconnected brain nuclei, known as the *song system*, that are dedicated to song learning and production. The song system consists of two circuits: the *motor pathway*, which is necessary and sufficient for song production, and the *anterior forebrain pathway* (AFP), a cortico-BG-thalamocortical circuit that is necessary for song learning and plasticity but not required for basic song production. The AFP contains a single BG nucleus called Area X, the medial portion of the dorsolateral nucleus of the thalamus (DLM), and the lateral magnocellular nucleus of the anterior nidopallium (LMAN). The LMAN is located in the avian *pallium*, which is the functional analogue of the mammalian cortex (Reiner et al. 2004). The cortico-BG-thalamocortical circuit is a particularly valuable model for studying BG function because it is involved in the learning of a single sensorimotor skill (song).

Table 4.1 Abbreviations

AFP	Anterior forebrain pathway
AIv	Ventral portion of the intermediate arcopallium
BG	Basal ganglia
BOS	bird's own song
DLM	Medial portion of the dorsolateral thalamic nucleus
GABA	Gamma-amino butyric acid
Glu	Glutamate
GPe	External segment of the globus pallidus
GPi	Internal segment of the globus pallidus
HVC	Used as proper name for vocal motor nucleus in the nidopallium
$\mathrm{HVC_X}$	HVC neurons that project to area X
RA	Robust nucleus of the arcopallium
RL	Reinforcement learning
SN	Spiny neuron
SNc	Substantia nigra pars compacta
SP	Substance P
STN	Subthalamic nucleus
VP	Ventral pallidum
VTA	Ventral tegmental area

Revealing the parallels between birds and mammals will likely allow bidirectional benefits for understanding BG function in both taxa. Although the evolutionary distance between birds and mammals is great (~300 million years ago they shared a common ancestor), their neural circuits may display significant differences in organization and function. This review highlights the strengths of the song system for the study of BG function and discusses recent progress in revealing the role of the BG in song learning and production and the homologies between the BG-cortical loops in songbirds and mammals. Furthermore, this review illuminates aspects of songbird and mammalian BG that differ to varying degrees. Because theoretical frameworks are increasingly used to guide experimental work, this review incorporates the theoretical advances with descriptions of the empirical evidence for or against specific hypotheses.

4.2 Similarities in the Basal Ganglia Between Birds and Mammals

Much of what is known about the song-related BG circuit is guided by a comparative approach. Briefly, the evidence underlying the homology between BG circuits in birds and mammals comes from multiple sources. First, numerous neurochemical markers distinguish the BG from the overlying pallium or cortex. For example, there is rich dopaminergic innervation from the midbrain (Bottjer 1993), strong and extensive expression of dopamine receptors (Casto and Ball 1994; Kubikova et al. 2014), and dense acetylcholinesterase activity in the BG relative to the surrounding tissue (Karten and Dubbeldam 1973). Second, anatomical connectivity studies in birds have shown that the pallium projects to striatal regions, which project to pallidal regions, which project to thalamic or midbrain targets just as in mammals (Reiner et al. 2004; Jarvis et al. 2005). The cellular composition of the avian BG, at morphological, electrophysiological, and neurochemical/molecular levels also parallels that seen in mammals, with mostly minor variations in some cases (see Sect. 4.2.2) (Kawaguchi et al. 1995; Farries and Perkel 2000). Functional connectivity among the structures in the avian circuit also matches that in mammals, with glutamatergic and excitatory cortical projections to the striatum and GABAergic and inhibitory striatal and pallidal outputs. Finally, to the extent that the functional roles of these structures are well characterized, they appear to have similar roles in generating and modulating behavior (Brainard and Doupe 2013; Ding and Perkel 2014).

Comparisons between the AFP and the mammalian cortico-BG-thalamocortical loop have provided valuable insights into the function of the AFP. The striking similarity of the circuitry in Area X in birds and the motor BG in mammals supports a unified view of the role of BG circuitry. This view has guided work in this field for two decades and will continue to do so, particularly in the realm of reinforcement learning for vocal plasticity (see Sect. 4.3). However, there are also important differences between the AFP and mammalian BG circuitry. This section highlights four aspects for which it is not yet clear to what degree the AFP differs from mammalian BG: organization into direct and indirect pathways, anatomical organization of the loop, functional mode of BG output, and the role of dopamine.

4.2.1 Direct and Indirect Pathways

A guiding principle for how the mammalian BG regulate the selection, initiation, and maintenance of movement is the notion of *direct* and *indirect pathways* (Fig. 4.1) (Albin et al. 1989; DeLong 1990). Corticostriatal projections contact two classes of striatal spiny neurons (in similar proportions), which are distinct in their expression of dopamine receptors and neuropeptides and in their intrinsic electrophysiological properties. Spiny neurons that express predominantly D1 dopamine receptors and substance P (SP) form part of the direct pathway and project to the output structures of the BG, the internal segment of the globus pallidus (GPi), and the substantia nigra pars reticulata. In contrast, spiny neurons that express predominantly D2 dopamine receptors and enkephalin form part of the indirect pathway and project to the external segment of the globus pallidus (GPe). The GPe, in turn, projects to the subthalamic nucleus (STN), which projects to BG output nuclei where projections from the indirect and direct pathways converge (Fig. 4.1).

The direct pathway is thought to have a net facilitatory effect on movement, and the indirect pathway is thought to oppose this influence. The balance of activity between the direct and indirect pathways is regulated by midbrain dopamine inputs to the striatum, and they critically modulate movement onset and speed. In addition to the direct and indirect pathways, there is a connection known as the *hyperdirect pathway*, which involves a direct projection from the cortex to the STN, relays excitation to the GPi, and impedes movement. While this paradigm is a gross oversimplification (as other anatomical pathways have been discovered), it still holds heuristic value for understanding the functional roles of each major pathway and

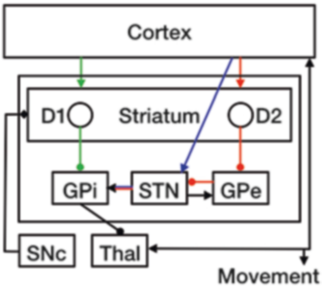


Fig. 4.1 Simplified schematic diagram of the main connections of mammalian basal ganglia (BG) circuitry: direct pathway, *green*; indirect pathway, *red*; hyperdirect pathway, *blue*. *Triangles* indicate glutamatergic excitatory connections; *filled circles* indicate GABAergic inhibitory connections; *diamonds* represent dopaminergic connections. (see Table 4.1 for all abbreviations)

helps guide clinical thinking with respect to therapies for Parkinson's disease and other movement disorders (Kravitz et al. 2010; Dudman and Krakauer 2016).

The avian Area X has the cellular composition of the striatum and globus pallidus combined and shares some aspects of the mammalian circuitry (Farries and Perkel 2002; Carrillo and Doupe 2004). Many important elements of the mammalian BG circuitry, however, have not been investigated in Area X (Fig. 4.2). For example, whether Area X has parallel direct and indirect pathway circuits is not fully known. There is evidence that about half of Area X neurons express both D1

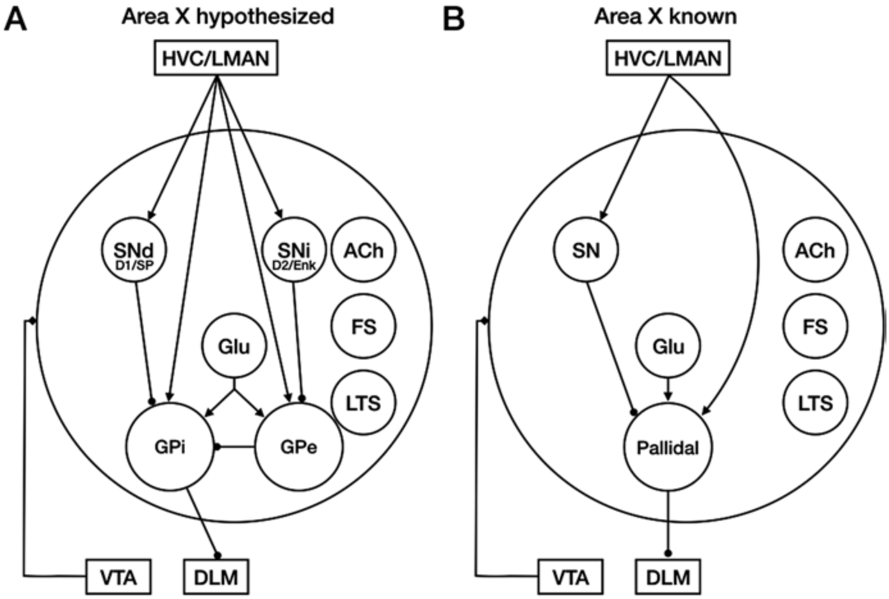


Fig. 4.2 Diagram of hypothesized and known cellular and synaptic aspects of Area X microcircuitry. *Triangles* indicate glutamatergic excitatory connections; *filled circles* indicate GABAergic inhibitory connections; *diamonds* represent dopaminergic connections. (**A**) *Hypothesized circuitry* within Area X based on mammalian striatal circuitry. Two types of spiny neurons (SN), corresponding to those in the direct and indirect pathway in mammals (*SNd* and *SNi*, respectively), are hypothesized to receive inputs from *HVC* and *LMAN*. SNd neurons are hypothesized to express *D1* dopamine receptors and substance P (*SP*), and SNi neurons are hypothesized to express *D2* dopamine receptors and enkephalin (*ENK*). Each of these SN types projects to a specific type of pallidal neuron, corresponding to the internal and external segments of the globus pallidum (*GPi* and *GPe* neurons, respectively). The GPi neuron type projects to *DLM*, while the GPe neuron makes local connections to the GPi neuron and possibly to the glutamatergic neuron (*Glu*), which could be an analog of subthalamic nucleus neurons. (**B**) *Known cell types and connections* within Area X. Spiny neurons (*SN*) receive glutamatergic excitatory input from *HVC* and *LMAN*. They contact and inhibit pallidal neurons, which project to *DLM*. Both SNs and pallidal neurons are GABAergic. Excitatory inputs also directly contact pallidal neurons. The three classes of striatal interneurons known in mammals are also present in Area X, including the cholinergic neuron (*ACh*), fast-spiking (*FS*) neuron, and low-threshold spike (*LTS*) neuron. The connections received or made by these interneurons are unknown. Evidence for a glutamatergic neuron type (*Glu*) in Area X that excites pallidal neurons has recently emerged. (see Table 4.1 for additional abbreviations)

and D2 receptor mRNA, about 30% express only D1 receptors, and about 15% express only D2 receptors (Kubikova and Kostál 2010). Assuming that the vast majority of these neurons are spiny neurons, this suggests that about half of Area X spiny neurons fit the mammalian model of segregated direct and indirect pathway spiny neurons, but the other half does not. Moreover, it is not known whether there is the same pattern of co-expression of neuropeptide markers, such as substance P and enkephalin, as in mammals.

In addition, different electrophysiological classes of Area X spiny neurons may exist. While there have been reports of heterogeneity in Area X spiny neurons (Farries and Perkel 2000, 2002), testing specifically for differences in intrinsic properties, such as those described in rodents, will require a molecular marker for D1-receptor or D2-receptor expressing Area X spiny neurons (Kreitzer and Malenka 2008). These are critical aspects of the accepted mammalian configuration of direct and indirect pathways, and they must be addressed directly in songbirds to confirm or refute potential anatomical and functional parallels with mammals.

Another tenet of the direct/indirect pathway configuration is that distinct and separate populations of pallidal neurons interact to affect the balance of activity of the output neurons. In mammals, the key site of convergence is the GPi, where GPe and spiny neurons project to and inhibit GPi output neurons that project to thalamus. Because pallidal and striatal neurons are mixed within Area X, identifying cell types corresponding to GPi and GPe in birds is complicated. Initial reports, based on in vitro recordings, described a single population of pallidal output neurons (Farries and Perkel 2002); however, some Area X output neurons received contacts from axon collaterals of other pallidal neurons. Moreover, comparison of the estimated number of Area X neurons projecting to DLM with the estimated number of cells with pallidal properties suggested that not all Area X pallidal neurons project to DLM (Farries et al. 2005). Together, these data led to the hypothesis that Area X output neurons correspond to GPi neurons and that Area X pallidal neurons that do not project out but project to the output neurons correspond to GPe neurons (Farries et al. 2005). This hypothesis gained additional support from in vivo recordings during singing. In particular, Goldberg et al. (2010) recorded two classes of pallidal neurons in Area X that were distinguished by their firing rates when birds were silent or singing. In a subset of neurons these properties correlated very well with whether the Area X neuron could be antidromically activated from DLM (indicating likely projection to DLM). Comparison of similar data from GPi and GPe neurons in primates supported this parallel in cellular composition similarity. Analyzing the expression of specific markers of each pallidal cell type, such as Nkx2.1 (Carrillo and Doupe 2004), in different populations of Area X projection neurons will allow further testing of parallels between songbirds and mammals.

4.2.2 Anatomical Organization of Basal Ganglia-Cortical Loops

A number of anatomical features that characterize the mammalian BG remain to be explored in Area X, including parallel loops, a diversity of interneuron types, patch-matrix components, and specific microcircuitry. BG-thalamocortical circuits are organized in parallel loops involving different areas of the pallium in birds or of the cortex in mammals (Alexander et al. 1986). These multiple loops likely evolved through a process of exaptation whereby the ancestral core unit, already present before lampreys, diverged from the main vertebrate lineage, and has been co-opted for multiple functions (Grillner et al. 2005; Dudman and Krakauer 2016). In mammals, the BG-thalamocortical loops have been divided into three parallel loops along the dorsoventral axis in the striatum: a *motor loop* (connected to the dorsolateral striatum), a *cognitive/associative loop* (connected to the dorsomedial striatum), and a *limbic loop* (connected to the ventral striatum/nucleus accumbens). One important outstanding question is where Area X fits among these parallel loops. Resolving this question will require further careful analysis of the molecular, anatomical, and functional properties of Area X. In particular, examining the relationship between Area X neuronal activity and reward or reinforcement could help clarify this issue.

Another principle of mammalian BG organization is the division of striatal tissue into patch and matrix components with different expression levels of particular proteins, such as acetylcholinesterase and calbindin, and different connectivity patterns (Graybiel et al. 1987). Hints of such organization are present in the developing Area X (Kosubek-Langer et al. 2017), but to what degree such organization is present in the adult Area X remains unclear. Further studies with additional markers will be critical to address the degree to which Area X demonstrates patch and matrix components.

One key difference between Area X and mammalian BG is the fact that there is not a connection between Area X and the avian STN. While the STN in zebra finches does have connections with more caudal and lateral parts of the BG, STN lacks direct connections with Area X (Person et al. 2008). An indirect pathway following the mammalian model would involve the STN, which provides a critical set of glutamatergic neurons in the BG circuitry. Interestingly, a population of glutamatergic neurons has been discovered in Area X (Budzillo et al. 2017). They provide strong excitatory drive onto Area X pallidal neurons and could represent a functional analog of STN neurons. Unlike cholinergic interneurons in the mammalian striatum, which can also release glutamate (Tecuapetla et al. 2010; Higley et al. 2011), Area X glutamatergic neurons are distinct from Area X cholinergic neurons (Budzillo et al. 2017). Learning more about the connectivity and intrinsic and functional properties of these glutamatergic neurons will allow a further assessment of the hypothesis that they have taken on a role similar to that of mammalian STN neurons.

Whatever the precise anatomical circuitry is within Area X in songbirds and whatever differences there are with BG circuitry in mammals, functional connectivity between the input structures (HVC [used as a proper name] and LMAN in songbirds) and the output neurons (DLM-projecting pallidal cells) appears to be surprisingly similar across songbirds and mammals. Indeed, excitatory inputs from HVC to Area X evoke a fast excitatory response in pallidal neurons that can precede the inhibitory response mediated by local GABAergic neurons (Leblois et al. 2009). This effect is similar to the response mediated by the hyperdirect pathway through the STN in primates (Nambu et al. 2000), which may be mediated by different specific circuitry in songbirds. The precise pathway mediating this excitation is unknown but most likely represents monosynaptic input from HVC. While LMAN also makes glutamatergic projections into Area X, which Area X neurons receive this input monosynaptically is not known, and the overall functional effect of this excitatory projection remains to be determined (Woolley et al. 2014).

4.2.3 Pallido-Thalamic Transmission

Striatal neurons are believed to inhibit the tonic firing of GABAergic pallidal neurons in mammals, thereby disinhibiting thalamic target neurons and allowing them to respond to excitatory inputs from the cortex, reinforce cortical firing, and, perhaps, help to select and initiate particular actions over other options (Deniau and Chevalier 1985; Hikosaka et al. 2000). In songbirds, the output from Area X to DLM appears to use similar principles but with important specializations. In brain slices, DLM neurons show strong rebound firing after single input pulses or trains of action potentials in Area X axons (Person and Perkel 2005). This is due to several specializations of this connection, including the extreme potency of the unitary claw-like or calyx-like GABAergic ending on each DLM neuron (Luo and Perkel 1999; Person and Perkel 2005), the extremely negative reversal potential of GABA_A receptor-mediated synaptic response (Person and Perkel 2005), and the strong post-inhibitory rebound properties of DLM neurons. Thus, in the absence of glutamatergic input, patterned activity of the Area X input to DLM can drive patterned DLM firing with very high temporal precision. In vivo, the large presynaptic terminal from Area X causes a powerful enough electrical signal in DLM that a single extracellular electrode can detect both pre-synaptic and post-synaptic action potentials (Person and Perkel 2007). This fact has allowed direct observation of the input-output relationship at this connection in anesthetized (Person and Perkel 2007; Kojima and Doupe 2009) and behaving animals (Goldberg and Fee 2012), which revealed that DLM spikes occurred during and immediately following brief pauses in presynaptic firing.

A simple interpretation of these results might be that post-inhibitory rebound (following brief pauses in Area X pallidal cell firing) is the main driver of DLM neuron firing. However, Goldberg and Fee (2012) demonstrated a projection from the robust nucleus of the arcopallium (RA) to DLM that provides excitatory inputs critical for DLM firing. Moreover, even after lesions in Area X, the RA input to DLM could drive patterned firing of DLM neurons. How these convergent inputs interact remains unclear. One possibility is that the excitatory input from RA drives patterned firing in DLM. Another possibility is that RA provides tonic excitation to DLM that, when combined with pauses in inhibitory input from Area X, could accelerate DLM firing. Further experiments will be needed to sort out the details of this interesting mechanism.

Despite superficial differences, recent findings have shown greater similarity between this Area X-DLM connection and mammalian circuitry than previously thought. For example, inputs to the thalamus from the pretectum (Bokor et al. 2005; Wanaverbecq et al. 2008) and substantia nigra (Bodor et al. 2008) have multiple presynaptic active zones, are physiologically potent, and are capable of driving postsynaptic thalamic firing. Thus, extrathalamic GABAergic inputs can give rise to temporally precise firing patterns in thalamic target neurons (Halassa and Acsády 2016). In a mouse brain slice preparation, optogenetic stimulation of nigro-thalamic inputs at low frequencies did not routinely evoke rebound bursts in motor thalamic neurons, but higher-frequency stimulation could drive firing (Edgerton and Jaeger 2014). In vivo, thalamic neurons fired at the end of optogenetically created pallidal bursts (Kim et al. 2017). This suggests that post-inhibitory rebound also shapes thalamic firing in mammals.

4.2.4 Role of Dopaminergic Inputs to Area X

Both the mammalian BG and songbird Area X receive projections from midbrain dopaminergic areas. In mammals, the substantia nigra pars compacta (SNc) projects to the dorsal striatum, and the ventral tegmental area (VTA) projects to the ventral striatum (also known as the nucleus accumbens). In finches, the SNc and VTA are continuous. The midbrain dopaminergic neurons that project to Area X are located between the SNc and the VTA (Gale and Perkel 2010a; Gadagkar et al. 2016). While they are commonly called VTA neurons, a rigorous assessment of possible differences between SNc and VTA has not been undertaken.

Given these neuroanatomical connections, a role for dopamine in song learning has long been hypothesized (Bottjer 1993; Doya and Sejnowski 1998). Suggestive evidence has come from several studies. Midbrain neurons in behaving adult finches show singing-related activity modulated by social context (Yanagihara and Hessler 2012). Dopaminergic neurons have selective responses to the bird's own song (BOS) that depend on excitatory transmission in Area X, consistent with a pathway from Area X to the ventral pallidum to midbrain dopaminergic neurons (Gale and Perkel 2010a).

Studies manipulating and recording activity of dopaminergic neurons provide more direct evidence that dopamine terminals in Area X are essential for adult song plasticity. For example, depletion of dopaminergic inputs into Area X altered song plasticity in a pitch-contingent auditory-feedback perturbation task (Hoffmann et al. 2016). In addition, the firing of dopaminergic neurons projecting to Area X is in line with an internal error prediction signal that provides an evaluation of song quality (Gadagkar et al. 2016). Finally, song changes can be driven by activating or inhibiting dopaminergic input to Area X, as predicted by a reinforcement learning model (see Sect. 4.3.3) of song learning (Xiao et al. 2018). Thus, the available evidence concerning the role of dopamine in song learning is entirely consistent with the current view of dopamine function in motor learning in mammals (Schultz et al. 1997; Doya 2000).

In the mammalian BG, it is common to distinguish tonic dopamine levels, which appear to be essential to support movement initiation and speed, from phasic dopamine, which appears to signal salience or reward prediction error. Dopamine levels measured using microdialysis probes in Area X rose when the bird was in the presence of a female and sang directed song (Sasaki et al. 2006). Phasic dopamine release is suggested indirectly from recordings of midbrain dopaminergic neurons, which can show behaviorally related burst activity (Yanagihara and Hessler 2006; Gadagkar et al. 2016). Measurements of dopamine levels with high temporal resolution (e.g., with fast scan cyclic voltammetry) will allow testing of specific hypotheses about the role of tonic or phasic dopamine in Area X (Woolley 2019).

In conclusion, many of the similarities between the AFP and mammalian BG-cortical circuitry clearly derive from common evolutionary origins (Grillner et al. 2005). But it is not clear whether the differences, such as the intermingled location of striatal and pallidal cell types or the absence of a connection between Area X and the STN, reflect fundamental differences in the functional capabilities of the circuit, or whether alternative solutions using similar principles have arisen in the different taxa through convergent evolution. Detailed analyses of the development of BG and the adult configuration of different portions of the BG across multiple taxa will be necessary to address this issue.

Interestingly, some differences between the avian and mammalian BG may relate to changes driven by evolution toward a specialization of the circuit for vocal learning. Vocalizations are rapid and require tight coordination of vocal musculature during singing (Elemans 2014). The surprisingly large axons (Leblois et al. 2009), fast signaling, and timing precision (Person and Perkel 2005; Goldberg and Fee 2012) that allow much faster transmission along the song-related BG, as compared to the homologous BG motor loop in mammals, may represent a specialization for movement speed. Careful comparative anatomy and physiology will be required to confirm that evolutionary pressure for fast muscle control has affected the song system.

4.3 Functional Hypotheses about the Role of the Basal Ganglia

The first evidence of the involvement of BG circuits in motor control came from studies of patients with Parkinson's or Huntington's disease. Based on the observation of clinical manifestations of such BG-related movement disorders and postmortem histological damage, BG circuits were proposed to be hyperkinetic or hypokinetic, depending on the pathological condition (Albin et al. 1989; DeLong 1990). As anatomical and physiological data have accumulated in various animal models (mostly rodents and nonhuman primates), the BG have been assigned a number of functions in motor control and in sensorimotor learning. Moreover, experimental evidence in songbirds has shed new light on the classical hypotheses of BG function and also led to novel hypotheses of BG function. This section will detail six of the most prominent models for BG function in light of data collected on songbirds. In particular, the section discusses classic models postulating a role for the BG in (1) action selection, (2) automation of learned motor plans, and (3) reinforcement learning. The section highlights how experimental evidence in songbirds has shed new light on these hypotheses and led to novel hypotheses regarding a role for the BG in evaluation of current performance and the generation of variability.

4.3.1 Action Selection

The tonic GABAergic output of the BG largely inhibits thalamocortical circuits involved in movement execution. A release of tonic inhibition, called *the disinhibitory process* (Deniau and Chevalier 1985), could underlie action initiation. Relatedly, the pattern of focused disinhibition and broad inhibition during movement execution could reflect the simultaneous release of a given motor program during the inhibition of unwanted programs (Mink and Thach 1993).

Early experimental evidence collected in songbirds did not support the hypothesis that the song-related BG-cortical loop (AFP) is involved in any action initiation or action selection process during singing. Indeed, lesions of Area X do not prevent normal song production in adult zebra finches (Scharff and Nottebohm 1991). However, more recent data suggest the contribution of BG circuitry in songbirds to action selection, in particular, to the on-line selection of syllables during song production (i.e., *syllable sequencing*). For example, partial lesions of Area X induce changes in the song sequence, in particular, an increase in the number of syllable repetitions in both zebra finches (Kubikova et al. 2014) and Bengalese finches (Kobayashi et al. 2001). In addition, acute perturbations of LMAN activity can modify syllable sequencing in zebra finches (Hamaguchi and Mooney 2012), and LMAN lesions modify the variability of syllable sequences in canaries (Alliende et al. 2017), which suggests a contribution of the BG-thalamocortical loop to syllable sequencing during singing. Until now, this question had been addressed primarily in songbird species with a limited song repertoire. The BG-cortical loop could also play some role in song selection in species that produce multiple song types and that select different song types depending on their environment (Brenowitz and Beecher 2005). Further investigation of the role of the BG in songbird species with large song repertoires is required to test a possible role in selecting specific song types over others.

4.3.2 Motor Routine and Transfer from Basal Ganglia to Cortex

Work in songbirds challenges a relatively recent idea that habits are encoded in the BG (Graybiel 2005). Adult finches can produce over one thousand stereotyped renditions of their song in a single day, and song production can easily be considered as a motor routine or habit (Fee and Goldberg 2011). However, adult song production does not require the integrity of the BG-cortical loop (Bottjer et al. 1984; Scharff and Nottebohm 1991). These early findings already informed the idea that the song motor program is not stored in BG-cortical loops. Rather, the BG-cortical loop is involved in the acquisition (Bottjer et al. 1984; Scharff and Nottebohm 1991) and plasticity (Brainard and Doupe 2000) of the song. Later findings have further demonstrated that small changes in the song program that allow the bird to correct motor errors are implemented initially by the cortico-BG loop, but then are transferred rapidly (< 24 hrs) to the premotor networks in the pallium, where the motor program is stored (Andalman and Fee 2009; Warren et al. 2011).

A similar idea is now emerging in the mammalian literature (Piron et al. 2016) concerning the storage of habit motor programs (Graybiel 2008; Desmurget and Turner 2008). Indeed, striatal activity initially implements adaptive changes in behavior to maximize reward or correct motor errors in mammals and ultimately drives downstream consolidation of those changes in cortical circuits that encode habitual behaviors in a BG-independent manner (Atallah et al. 2004; Pasupathy and Miller 2005). The progressive transfer of the motor control of habitual behaviors from the BG to cortex during learning (i.e., the consolidation of learning in the cortical network) is consistent with the ability of the brain to chunk complex motor sequences into elementary movements (Graybiel 1998; Barnes et al. 2005). A repertoire of elementary movements could be learned early in development (e.g., grasping and reaching) and be stored in the cortex after consolidation. Consistently, the BG encode movement information during the initial stages of motor learning, but they signal only the start and end of a movement sequence after learning (Barnes et al. 2011). Therefore, it now seems more likely that the BG are involved in starting and ending habitual actions, but that the program for the habitual movement itself is encoded in cortex. This fits better with the situation in songbirds, though not perfectly.

4.3.3 Reinforcement Learning

As with other complex sensorimotor skills, birdsong is learned through a trial-and-error process (Marler and Waser 1977; Doya and Sejnowski 1995). Juvenile birds first produce an imprecise and variable version of their future song, called *plastic song*, and gradually progress toward *crystallized song*. The reinforcement learning (RL) framework, developed initially for robotics and machine-learning purposes

(Sutton and Barto 1981), is an appealing theoretical model to understand how neural activity in the BG-cortical loop drives learning both in mammals (Doya 2000; Hikosaka 2007) and in songbirds (Fee and Goldberg 2011). As the latter review article has already provided an extensive overview of the experimental evidence supporting the implementation of RL in the song-related BG-cortical circuits (Fig. 4.3), this review focuses on more recent evidence in support of the RL hypothesis and the questions remaining with respect to its implementation for this circuit.

Standard RL algorithms are well suited to low-dimensional tasks such as choosing between discrete options. It remains unclear whether these algorithms can be used to learn high-dimensional and continuous tasks, such as performing a given gesture or producing a given sound (Parr and Russell 1998; Dhawale et al. 2017). Moreover, it is likely to be difficult for RL algorithms to learn to drive motor output

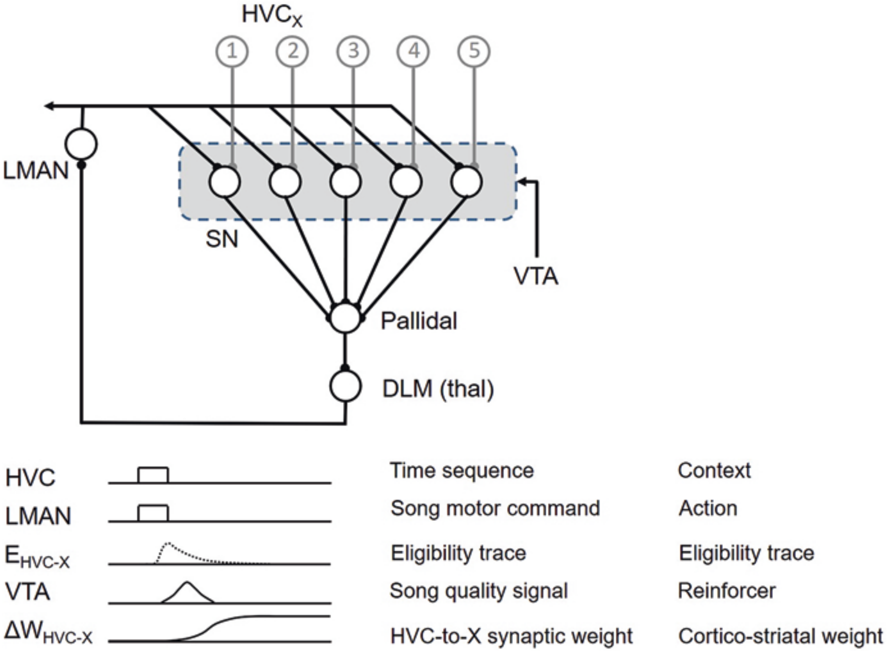


Fig. 4.3 Schematic diagram of reinforcement learning (RL) circuitry and processes in songbirds, referring to the classical concepts from the RL framework. HVC inputs to Area X spiny neurons (SN) provide temporal information (context in the RL theory), and LMAN inputs provide an efference copy of the motor command selected action. The activity from *HVC* and *LMAN* leads to an *eligibility trace*, or capacity for synaptic plasticity, whose duration outlasts the electrical synaptic activity. Based on the degree of song match to the tutor song template, the *VTA* is hypothesized to provide an evaluation signal or reinforcer. When these signals converge, for example after a high-quality song is produced, changes in synaptic strength or weight at the connections from HVC to Area X (corresponding to corticostriatal synapses in mammals) are hypothesized to occur. (see Table 4.1 for abbreviations; redrawn from Fee and Goldberg 2011; used with permission)

through a highly redundant motor system with many degrees of freedom (Lashley 1933; Bernstein 1967). Previous attempts to model the song acquisition process with RL algorithms have relied on a simplified vocal output and a minimal tutor song (Fiete et al. 2007). Building a RL model able to learn even the relatively simple zebra finch vocalizations with a realistic biomechanical representation of the birds' vocal apparatus remains an important challenge (Düring et al. 2013).

The following sections discuss which circuit may implement the song evaluation (also called the *critic* in RL *actor-critic theory*, Sect. 4.3.3.1) and describe the circuits generating and modulating the behavioral variability needed for exploration during such trial-and-error learning (Sect. 4.3.3.2).

4.3.3.1 Song Evaluation

Reinforcement learning relies on the ability of the agent to reinforce successful actions and thereby requires an evaluation of current performance. The mammalian BG literature mostly concentrates on instrumental conditioning, whereby actions are evaluated externally and outcome performance is measured as the amount of reward received. On the contrary, most natural sensorimotor learning paradigms, including birdsong, rely on an internal evaluation of the subject's performance.

An early hypothesis about song learning was that the AFP generates an evaluation of the bird's own song (*template comparison*) during the sensorimotor phase of song learning (Konishi 2004). The initial suggestion for this role came from the findings, in anesthetized birds, that neurons in LMAN and Area X respond more to playback of the BOS than to other songs (Doupe 1997; Solis and Doupe 1997).

Because the BOS and the tutor's song are highly similar under normal circumstances, it is difficult to determine whether AFP neurons encode the BOS or the tutor's song. To address this, Solis and Doupe (2000) manipulated the tracheosyringeal nerve, thereby altering BOS and enabling it to be differentiated from the tutor song. Following the manipulation, most neurons in LMAN and Area X responded preferentially to playback of the altered BOS rather than to tutor song. On the surface, this finding argued against the hypothesis that AFP neurons contribute to evaluating the degree of match between BOS and tutor song. However, a small subset of neurons clearly showed stronger responses to the tutor song than to the BOS, suggesting that the BG have access to a memory of the tutor song and could help signal the similarity between the BOS and the tutor song.

These tutor song-selective responses and BOS-selective responses are also conveyed to midbrain dopaminergic neurons, which are hypothesized to transmit reinforcement signals to other parts of the brain (Schultz et al. 1997). Indeed, some Area X neurons that project to thalamic nucleus DLM also send axon collaterals to a region of ventral pallidum (VP) that project to the substantia nigra and VTA (Gale and Perkel 2010b). Playback of the BOS, but not other songs, reduced the activity of VP neurons and excited midbrain dopaminergic neurons. These responses in the VP and midbrain depend on glutamatergic synaptic transmission in Area X. Together, these data suggest that the AFP could provide a comparison signal to dopaminergic neurons to help birds detect when they produced better copies of the tutor song, which is essential for reinforcement learning (see Sakata and Yazaki-Sugiyama, Chap. 2).

It is too simplistic, however, to think that AFP neurons serve mainly to evaluate the BOS against the tutor song memory. For example, neurons in HVC, Area X, and LMAN responded to BOS playback in anesthetized and sleeping zebra finches but responded much less, if at all, in awake animals (Cardin and Schmidt 2003; Hamaguchi et al. 2014), and these neurons very clearly showed singing-related firing, even in deafened animals (Hessler and Doupe 1999). Moreover, distortions of auditory feedback during singing did not affect the motor-related firing, suggesting that these neurons do not provide an online evaluation of song (Kozhevnikov and Fee 2007; Hamaguchi et al. 2014). What function does this highly selective auditory representation subserve? A clue may be found in considering an inverse model (see Sect. 4.3.4).

Dopaminergic neurons in the midbrain project back to Area X, and recent experimental data provide strong support for the role of dopamine as a reinforcer during song learning, signaling internally evaluated song quality. First, dopamine depletion confined to Area X impeded the ability of adult birds to shift the fundamental frequency of a given syllable in response to distorted auditory feedback (Hoffmann et al. 2016), and dopaminergic neurons projecting to Area X were necessary for proper copying of tutor song in juveniles (Hisey et al. 2018). These results suggested that adjusting vocal output based on auditory feedback requires proper dopamine delivery in Area X. Secondly, dopaminergic neurons that project to Area X signal the internally evaluated song quality in a manner consistent with encoding *reward prediction error* (Gadagkar et al. 2016). These studies used an artificial adult song plasticity paradigm; whether natural variation in plastic song causes changes in VTA firing that is consistent with this model during normal song acquisition remains unknown. Regardless, as predicted by a reinforcement learning model, activating dopaminergic fibers in Area X contingent on a high fundamental frequency rendition of a syllable resulted in an increase in the syllable's fundamental frequency; inhibition of the terminals contingent on high frequency resulted in a decrease in the frequency of that syllable (Hisey et al. 2018; Xiao et al. 2018).

In addition to BG neurons, song-evaluation signals from other pallial areas could be relayed to dopaminergic neurons in the VTA. For example, neurons in the ventral portion of the intermediate arcopallium (AIv) not only receive auditory information and project to the VTA, they also respond to disruptive auditory feedback presented during singing and they are necessary for juvenile song acquisition (Gale and Perkel 2010a; Mandelblat-Cerf et al. 2014). Neurons in other brain regions display changes in singing-related firing when the auditory feedback of the bird is distorted (e.g. noise-masked syllables); for example, the caudolateral mesopallium includes neurons that respond only to distorted auditory feedback during singing (Keller and Hahnloser 2009). Whether these neurons transmit song evaluation signals to DA neurons through the AIv remains to be determined.

As for a biologically realistic RL implementation, the timing of song-evaluation signals poses a serious and well-known challenge called the *temporal credit* *assignment problem*. Because reinforcing signals arrive after production of a specific movement, they could reinforce the motor signals corresponding to the next movement or interfere with the motor signals (Florian 2007). This is particularly critical in the context of birdsong production, as vocal gestures are fast and brief (on the order of a few tens of milliseconds). A key observation in this respect is the timing of dopaminergic neuron responses to disturbed auditory feedback: these responses are fast enough (~50 ms latency) to be assigned to a given vocal gesture, or note, in the song (Gadagkar et al. 2016). However, a trace of the corresponding motor signals (the *eligibility trace*) would need to remain present in the synapses involved so that *note-specific plasticity* can be modulated or induced by the dopamine signal. Slow synapse-specific calcium signaling could be the substrate for such an eligibility trace (Fiete et al. 2007; Wanjerkhede and Bapi 2011). Alternatively, persistence of motor signals that generate behavioral variability could allow reinforcement by a delayed dopamine signal (Darshan et al. 2017).

A related issue is the spatial specificity of the dopamine-based reinforcement signal. Given the topographic organization of the BG-cortical loop (Iyengar et al. 1999; Luo et al. 2001) and the downstream motor pathway (from RA to muscles) (Vicario and Nottebohm 1988; Vicario 1991), Area X likely displays a functional organization reflecting the vocal output of the various syrinx muscles. In an ideal world, each muscle would impact a different acoustic parameter and be modulated in Area X by a specific dopamine signal. However, the contribution of muscles to acoustic parameters is complex, nonlinear, and gesture dependent (Fee et al. 1998; Srivastava et al. 2015), and the VTA-Area X pathway does not display any clear topographic organization (Luo et al. 2001; Gale and Perkel 2010b).

4.3.3.2 Generation and Modulation of Variability

Variability in movement kinematics during performance is necessary for many forms of motor learning and is referred to as *motor exploration* in the RL framework (Dhawale et al. 2017). During song acquisition, young birds transition from initially producing erratic vocalizations (*subsong*, similar to infant babbling) to producing variable but temporally structured songs (*plastic song*), until the end of the song-learning period when they are able to produce stereotyped, crystallized song (Marler and Pickert 1984; Tchernichovski et al. 2000). Song variability continues to decrease with age in adult songbirds, though at a much slower rate than during juvenile learning (Sakata and Vehrencamp 2012; James and Sakata 2014). In addition, song variability can change acutely in adults depending on the social context (Woolley and Kao 2015). In particular, male finches produce a highly stereotyped version of their song when singing toward a female (*directed song*) but a more variable song when they sing by themselves (*undirected song*) (Kao et al. 2005; Sakata et al. 2008).

High levels of song variability could be advantageous during training to allow motor exploration (during undirected song and, to a greater extent, during the early phase of juvenile song acquisition) but not during performance (courtship songs and possibly during territorial defense). Indeed, female finches prefer the sound of the more stereotyped directed song over the more variable undirected song (Woolley and Doupe 2008; Chen et al. 2017). Importantly, multiple aspects of song are variable: acoustic structure of syllables (e.g., fluctuation in the syllable's fundamental frequency), syllable order (sequence variability), and syllable and gap duration (timing variability). It is necessary to evaluate them separately and in the appropriate species (e.g., adult zebra finches do not produce songs with much sequence variability).

The ability of songbirds to modulate song variability with context suggests that the brain actively injects variability for motor exploration. The current understanding of neural signals driving song production favors a division of labor between the motor pathway (guiding song sequence and timing through HVC inputs to RA) and the BG-cortical loop (contributing solely to acoustic modulation of syllables) (see Sect. 4.3.1) (Mooney 2009). Consistent with this idea, the neural mechanisms involved in the generation and/or modulation of different types of variability appear to be segregated: timing variability relies mostly on the fluctuations in HVC inputs to RA (Ali et al. 2013) and acoustic variability relies on the integrity of the AFP (Kao et al. 2005; Olveczky et al. 2005).

The circuits generating and modulating sequence variability, however, remain controversial. No changes in syllable sequencing were observed in adult zebra finches (Bottjer et al. 1984; Kao and Brainard 2006) or in adult Bengalese finches (Hampton et al. 2009) following LMAN lesions. On the contrary, other studies have reported a reduction in sequence variability following LMAN lesion or inactivation in juvenile zebra finches (Scharff and Nottebohm 1991; Olveczky et al. 2005) and in canaries outside the breeding season (Alliende et al. 2017). In addition, perturbations in activity in LMAN or Area X increase sequence variability in zebra finches (Hamaguchi and Mooney 2012; Tanaka et al. 2016).

Multiple lines of experimental evidence support a role of the song-related BG-cortical loop in driving and modulating acoustic variability in song, and recent data even suggest a mechanism for the decrease in acoustic variability during development. Lesions of the cortical output of the loop (LMAN) in juveniles reduced the acoustic variability of plastic song and abolished context-dependent modulation of acoustic variability in adult finches (Scharff and Nottebohm 1991; Kao and Brainard 2006). Moreover, singing-related activity of neurons in LMAN was much more variable in undirected song than in directed song (Kao et al. 2008), suggesting that trial-to-trial variability in LMAN firing drives song variability. The same circuit also drives babbling in juveniles (Aronov et al. 2008) and song fluctuations (around a stereotyped song pattern) during plastic song in older juveniles (Olveczky et al. 2005; Warren et al. 2011). During song acquisition, the singing related activity of RA neurons gradually changes from highly variable firing patterns to precise and sparse bursts of spikes locked to song motifs (Olveczky et al. 2011). LMAN input is necessary for the expression of RA firing variability, and the change from variable to stereotyped firing patterns in RA throughout development could be explained by the strengthening and pruning of HVC inputs to RA (that drive stereotyped patterns) while LMAN inputs remain unchanged (Garst-Orozco et al. 2014). Indeed, as HVC input to RA becomes stronger, it drives stronger bursting in RA, interleaved with periods of inhibition-driven silence (Olveczky et al. 2011). LMAN inputs to RA are mediated mostly through voltage-dependent NMDA receptors (Mooney and Konishi 1991; Stark and Perkel 1999), and the influence of LMAN on RA firing is weak when HVC inputs are strong. The influence of LMAN on RA is thus diminished with age, resulting in a progressively more stereotyped song. In summary, strong influence of LMAN on RA early in development initially drives subsong and then drives the variability in plastic song; thereafter, HVC plays an increasingly strong role in driving the stereotyped firing of RA as the bird approaches crystallization.

Area X also plays a central role in modulating song variability, likely through its influence on LMAN. Context-related changes in song variability are associated with changes in the firing variability of Area X pallidal neurons (Hessler and Doupe 1999). Area X spiny neurons, which provide convergent inhibition on pallidal cells, display increased firing rates during undirected singing (as compared to directed singing), possibly explaining changes in pallidal firing variability (Woolley et al. 2014). Interestingly, other inhibitory interneurons in Area X also change their pattern of activity with social context (Woolley 2016). In contrast, HVC neurons projecting to Area X show no change in firing with social context (Woolley et al. 2014), suggesting that changes in song-related patterns of activity related to social context emerge in the BG-cortical loop.

The dopaminergic input to Area X displays context-dependence (Sasaki et al. 2006; Yanagihara and Hessler 2006) and could trigger changes in song variability. Indeed, the activation of D1 receptors in Area X decreases song-evoked activity in the BG-cortical loop (Leblois et al. 2010), and blocking D1 transmission or depleting dopamine in Area X blocks the modulation of song acoustic variability with social context (Leblois et al. 2010; Leblois and Perkel 2012). Surprisingly, partially lesioning dopaminergic terminals in Area X, using the 6-OHDA neurotoxin, decreased song variability while leaving social context modulation of this variability unaffected (Miller et al. 2015). One challenge with partial neurotoxic lesions is that the long-term depletion of dopamine in Area X may reflect a pathological condition rather than the acute effects of dopaminergic signaling in Area X.

Given that LMAN lesions abolish song acoustic variability, the brain region involved in generating this variability must be in, or upstream from, LMAN. Two different hypotheses have been proposed. The first hypothesis proposes that acoustic variability is generated in the thalamocortical targets of the song-related BG (DLM and LMAN), consistent with the evidence that Area X lesions did not prevent variability in the song of juvenile birds (Fee and Goldberg 2011; Chen et al. 2014), and the lesions induce only transient changes in song variability in adults (Ali et al. 2013). Moreover, a recent theoretical study proposed that the topographic organization of the downstream motor pathway from LMAN to RA could be sufficient to drive large and spatially coherent fluctuations of activity, provided that LMAN and RA have strongly connected recurrent networks with balanced inhibition and excitation (Darshan et al. 2017). An alternative hypothesis is that song variability is generated within Area X (Woolley et al. 2014; Budzillo et al. 2017). Identifying the specific cellular and circuit mechanisms underlying the generation and regulation of song variability remains a crucial goal.

Whatever the precise site for the generation of song variability, the BG-cortical loop actively injects variability for motor exploration. In a RL paradigm, following evaluation of the current song (see Sect. 4.3.3.1), successful vocal performance would be reinforced to optimize the vocal output. If the BG-cortical loop is indeed the site where such reinforcement learning is implemented, the BG network must receive a copy of the current motor command or at least be instructed about the exploration signals that are causally related to the reinforcement (Fee 2014). In the classical *actor-critic model* (Houk and Wise 1995; Suri and Schultz 1999) the BG-cortical loop is the actor that drives actions and therefore has access to the motor command. In songbirds, the BG-cortical loop is not necessary for song production and therefore is not driving the motor program itself.

LMAN is thought to drive acoustic variability and could send an efference copy of this acoustic variability to Area X for its evaluation with respect to dopaminergic reinforcement signals (Fee and Goldberg 2011; Fee 2014). Several recent findings may challenge this concept, however. First, blocking LMAN input into RA (and thereby silencing the influence of LMAN on song) does not impede the ability of the song-related BG-cortical loop to adaptively modify its output to optimize song (Charlesworth et al. 2012). In this experimental context, the BG-cortical loop must rely on an efference copy of motor fluctuation that arises from somewhere other than LMAN. Secondly, LMAN neurons display very little noise correlation in their activity during singing (Darshan et al. 2017); only in RA are fluctuations from LMAN input amplified and spatially correlated, giving rise to behaviorally relevant variability in the motor command. Accordingly, the correlation between LMAN single-neuron activity and acoustic variability during singing is likely very low compared to what is observed in RA (Sober et al. 2008). Further exploration of the link between LMAN neural signals and the vocal output are needed to test this assertion.

If LMAN does not transmit signals related to pitch variation on a trial-by-trial basis, where could the BG-cortical loop obtain this information? One possibility is that RA transmits an efference copy of the song to DLM through a recently discovered pathway (Goldberg and Fee 2012). Alternatively, HVC neurons projecting to Area X could, in principle, transmit such information, although no correlation between HVC firing and the fluctuations in vocal output has been reported (Kozhevnikov and Fee 2007). In the latter case, any variability emerging downstream from HVC would not be transmitted to the BG-cortical loop, impeding the ability of the BG to correlate song fluctuations with the associated reinforcement signals.

There is some evidence that the BG-cortical loop modulates and/or generates motor variability for the learning and maintenance of motor skills in mammals. Indeed, striatal activity displayed large fluctuations during the early stages of a sensorimotor learning task, and the fluctuations decreased as performance improved (Barnes et al. 2005). Moreover, dopamine signaling appears to be involved in the regulation of behavioral variability in mammals, including humans. Dopamine-regulating genes are responsible for interindividual differences in exploration and exploitation behaviors (Frank et al. 2009), and animals treated with drugs that enhance the dopamine signal (such as cocaine or amphetamines) exhibited behavioral stereotypy (Canales and Graybiel 2000; Aliane et al. 2009). Differences in movement latency associated with different reward values, possibly reflecting a transition from exploration (when the action outcome is uncertain) to exploitation (when the action outcome is desirable), rely on D1 receptor transmission in the BG in primates (Nakamura and Hikosaka 2006). Finally, dopamine depletion is associated with increased firing variability in BG output neurons (Boraud et al. 2002), and large fluctuations of dopamine levels caused by dopamine replacement therapy can cause dyskinesia, which is characterized by unwanted erratic movements (Voon 2017). While further experimentation in mammals is important, understanding the generation and control of variability in movements in songbirds may help guide research on the role of the BG in that process in mammals and the role of variability in learning in general.

4.3.4 Inverse Learning

Vocal learning requires the integration of auditory and motor signals in the brain to support the imitation of the tutor's vocalizations. Different models of learning, however, rely on different interactions between auditory and motor signals. For example, in the RL framework discussed previously, auditory signals are funneled into a single *song quality* reinforcement signal that guides motor learning. However, an alternative, though not mutually exclusive, model proposes that learning does not rely directly on sensory feedback to guide trial-and-error learning, but rather that learning relies on the transformation of a sensory representation of the song into a motor representation (Hahnloser and Ganguli 2013). Such an *inverse model*, or auditory motor map, may be built early in life and used later to reproduce any type of sound. A relatively simple mechanism by which an inverse model may be learned is through associative learning during motor production: coactivation of the neuronal assembly driving the motor command and the neuronal assembly encoding the sensory representation of the produced gesture lead to tightening of their connections (Heyes 2001). Through this process, the neurons involved in the sensory representation of a gesture build stronger and stronger connections with the motor neurons that are driving its production. After sufficient training, the sensory representation of a sound could drive the motor pattern necessary for the production of the sound. Previous modeling work proposed a similar mechanism for arm movements (Chaminade et al. 2008; Rolf et al. 2010) or speech learning (Westermann and Miranda 2004; Oudeyer 2005). One critical element of an inverse model in the brain is the presence of neurons that have both sensory and motor representations of the same motor gestures (Oztop et al. 2006; Arbib 2008). In primates, these *mirror neurons* have been recorded in the premotor cortex when animals performed a motor action or saw the same action performed by another individual (Rizzolatti et al. 1996; Iacoboni 1999).

Neurons with similar features have been reported in songbirds. Indeed, neurons in HVC show both sensory-related and motor-related activity (McCasland 1987; Margoliash 1997). More recently, HVC neurons projecting to Area X reportedly displayed similar responses when the bird heard a syllable (auditory response) or when it produced the same syllable (motor drive) (Prather et al. 2008; Hamaguchi et al. 2014). While such a mechanism may be successfully implemented in a simplistic representation of the song system (Hahnloser and Ganguli 2013; Hanuschkin et al. 2013), whether it could drive song learning under realistic conditions is unclear, and further exploration of this idea is needed.

Beyond the ability to display sensory and motor responses to the same motor gesture, a critical feature of mirror neurons is the precise temporal alignment of firing relative to the acoustic signal while hearing it or singing it. A recent theoretical investigation of inverse model implementation suggests two different scenarios (Hanuschkin et al. 2013). In one scenario, the inverse model is learned through random exploration of the motor space. In this case, the sensory activity evoked during the production of a sound lags the motor command by a constant delay referred to as the *sensorimotor loop delay*. This delay includes the transmission delay from motor neurons to muscles, the time involved in vocal production, and the auditory neural delays between the auditory cue and the evoked activity in auditory neurons (Giret et al. 2014). This constant lag between the causal motor activity and the evoked auditory activity patterns will lead to an association between the motor pattern driving a syllable and the auditory pattern evoked by the preceding syllable (with a time interval corresponding to the sensorimotor loop delay) and, therefore, a lag between auditory and motor patterns of the mirror neurons. Interestingly, responses in the cortical target of the BG loop, LMAN, show correspondence between motor activity and auditory-evoked responses that display similarity with a time-lag corresponding to the sensorimotor delay (Giret et al. 2014). This finding argues for the existence of a causal inverse model upstream from LMAN.

In the second scenario, the inverse model is learned while producing many renditions of a stereotyped motor output. During the production of a stereotyped sequence, the motor command of a given motor gesture always coincides with the auditory activity evoked by the preceding gesture in the sequence. Associative learning leads to a *predictive inverse model* that maps sensory activity to future motor action and is associated with mirror representation with no lag. Interestingly, auditory and motor responses of neurons in HVC that provide input to Area X display mirror responses to single syllables (either sung or played back) with no apparent lag (Prather et al. 2008; Hamaguchi et al. 2014), which supports a predictive inverse model involving the HVC-to-Area X connection.

In general, the existence of an inverse model upstream from the motor areas of the song system could naturally explain why selective auditory responses can be driven in the whole motor network (e.g., in sleeping or anaesthetized birds). Evidence is lacking with regard to where in the brain the inverse model is implemented and what types of circuitry and plasticity mechanisms would support the existence and development of an inverse model. Also, the existence of an inverse model in the avian brain would predict the ability to perform one-trial learning, which is not observed in songbirds, leaving open the question of its existence. Finally, RL and inverse models are not necessarily mutually exclusive, and exploration of how they may be integrated may lead to novel insights into the mechanisms of vocal learning and plasticity.

4.4 Future Directions

While the past two decades have seen substantial growth in the understanding of how the BG contribute to song learning and plasticity, many questions remain unresolved. In addition to the application of techniques and approaches applied broadly across neuroscience, such as gene editing and big data, our understanding of BG function and dysfunction will particularly benefit from focused research in specific areas.

4.4.1 Cellular Components and Microcircuitry in the Songbird Basal Ganglia

It will be critical to continue the detailed determination of the cellular components and micro-circuitry of Area X and the AFP. Some of this work is necessarily descriptive but bears directly on how the circuit contributes to song learning. Such work also addresses questions related to how similar the AFP is to mammalian BG circuits and will allow the generation of more biologically plausible computational models. Moreover, this careful descriptive analysis is also likely to help elucidate which portion of the mammalian BG is most comparable to which portion of Area X (e.g., dorsal versus ventral striatum).

4.4.2 Cellular and Synaptic Plasticity Mechanisms

Several forms of synaptic plasticity have been described in the BG circuit. Coincident activity in afferents from HVC and/or LMAN along with postsynaptic depolarization of Area X neurons leads to long-term synaptic potentiation (LTP) of the HVC and LMAN excitatory inputs to Area X (Ding and Perkel 2004). This form of synaptic plasticity requires D1 receptors and was observed in brain slices from both juvenile and adult zebra finches. However, the degree to which this form of synaptic plasticity plays a role in song learning is unclear. Knowing the precise temporal contingencies (e.g., the temporal window when plasticity is possible after dopamine release) will shed light on the duration of the *eligibility trace* that this mechanism could provide. In LMAN, LTP of recurrent synapses made by LMAN neurons on their neighbors and long-term depression of afferents from DLM have been described (Boettiger and Doupe 2001). These forms of plasticity were restricted to the ages when song learning occurs.

Several forms of thalamocortical synaptic plasticity have been described in mammalian sensory systems (e.g., Crair and Malenka 1995), and it seems fruitful to explore parallels between these systems. While Fee (2012) hypothesized a form of synaptic plasticity at the connections from LMAN to Area X, it has not been addressed experimentally. Further work testing for other forms of synaptic plasticity and in vivo studies testing their role in vocal learning will be essential.

4.4.3 Functional Roles of Different Circuits

As researchers in the field make progress toward understanding the functional contributions made by each portion of the song system circuit, some notes of caution are needed. First, much of the work that purports to describe the avian BG has involved manipulations of LMAN, an area that is an output structure of a BG loop but is not actually part of the BG. It is critical to consider the entire loop in studies of the song system. One strength of the study of avian vocal learning is that, unlike cortical-BG circuits for other behaviors, there is precise knowledge of which portions of the pallium provide input to the BG for this specific behavior, which portions of the BG are involved, and which thalamic target receives the BG output.

Second, it is widely appreciated that interpreting the results of lesion experiments is fraught with potential confounds. For example, normal performance of a behavior after lesioning a particular structure suggests that the structure is not essential for production of the behavior. However, such data do not necessarily imply that the structure does not take part normally. Conversely, altered behavior after a lesion does not necessarily mean that the damaged structure is critical for the behavior. These ideas are highlighted in a study on the short-term and long-term effects of lesions of the nucleus interfacialis (NIf) (Otchy et al. 2015). Lesions result in immediate, severe disruption of song, which suggests NIf is essential for song production. However, because song eventually recovers following NIf lesions, the data are more consistent with a permissive role rather than an instructive role of NIf inputs to HVC in generating the neural activity pattern for song.

Another potential concern while interpreting lesion studies is how any retrograde effect may affect circuitry in a nucleus projecting to a lesioned structure. For example, if damage to Area X leads to sprouting of local axon collaterals in HVC that project to Area X (HVC_X neurons), such plasticity could alter intra-HVC circuitry and affect signals sent to RA. Moreover, the presence of loops in the song system further complicates interpretation of the results of lesion experiments. For example, the projection from LMAN to Area X, via collaterals of axons projecting to RA, has an unknown function. Lesions of LMAN in adult zebra finches did not eliminate differences in firing rate or variability among pallidal neurons between directed and undirected social contexts (Woolley et al. 2014). More data are needed to understand the function of those collateral inputs.

4.4.4 New Experimental Technologies

Myriad new technologies show great potential for studying song learning. Manipulation of gene expression through viral vectors is increasingly common (Haesler et al. 2004), and transgenic finches have been created, although the technology remains far from routine (Agate et al. 2009; Liu et al. 2017). Use of optogenetic methods (Roberts et al. 2012; Xiao et al. 2018) and chemogenetic manipulations (Yazaki-Sugiyama et al. 2015; Heston et al. 2018) of song system neurons has begun (Hisey et al. 2018; Tanaka et al. 2018). Refinements in viral gene delivery and cell-type specific targeting will enhance the usefulness of these approaches. Another promising approach is optical recordings of calcium transients using lightweight head-mounted cameras (Liberti et al. 2016; Picardo et al. 2016). As the sensitivity and temporal resolution of genetically encoded calcium and, eventually, voltage sensors improve, these approaches could provide long-term measurements of neuronal activity from large numbers of neurons, especially in relatively superficial structures.

With respect to studying the BG in songbirds, Area X poses both advantages and disadvantages. Because of the clear homology between the avian Area X and the mammalian BG, experiments involving manipulation of gene expression offer the potential for stronger generalizations of interpretations than are possible for the pallial regions. For studying microcircuitry, the fact that striatal and pallidal neurons are intermingled in Area X may facilitate study in single tissue slices or sections. On the other hand, this intermingling of cells makes it essential to have cell-type-specific markers and promoters to be able to fully tease apart the roles of different cell types within the circuit.

4.5 Conclusions

One of the principal strengths of the song system for allowing mechanistic understanding of complex learned social behavior is the ease of crossing levels of organization. Because of the well-described birdsong circuitry with discrete nuclei and connections, it has been possible to relate phenomena at molecular, synaptic, cellular, and circuit levels to specific aspects of learning, producing, and processing vocal signals. There is every reason to believe that those advantages will continue to favor this system. Further integration of concepts from other areas of neuroscience and from other parts of the brain into the current working models of the song system will be fruitful. For example, work on temporal sequences of firing in the hippocampus may inform the understanding of similar issues in the song system. Similarly, understanding the role of the cerebellum in learning precisely timed movements and in detecting motor errors could guide progress into understanding the remarkable ability of songbirds to copy songs from other individuals. Although potential homologies between the pallial structures in songbirds and mammals remain unclear, those for the BG are much more certain. Thus, progress in understanding the role of the BG in song learning and plasticity is likely to generalize to other forms of sensorimotor learning and plasticity across the vertebrates.

Compliance with Ethics Requirements Arthur Leblois declares that he has no conflict of interest.

David Perkel declares that he has no conflict of interest.

References

Agate RJ, Scott BB, Haripal B et al (2009) Transgenic songbirds offer an opportunity to develop a genetic model for vocal learning. Proc Natl Acad Sci 106:17963–17967. https://doi.org/10.1073/pnas.0909139106

Albin RL, Young AB, Penney JB (1989) The functional anatomy of basal ganglia disorders. Trends Neurosci 12:366–375. https://doi.org/10.1016/0166-2236(89)90074-X

Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci 9:357–381. https://doi.org/10.1146/annurev.ne.09.030186.002041

Ali F, Otchy TM, Pehlevan C et al (2013) The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. Neuron 80:1–13. https://doi.org/10.1016/j.neuron.2013.07.049

Aliane V, Pérez S, Nieoullon A et al (2009) Cocaine-induced stereotypy is linked to an imbalance between the medial prefrontal and sensorimotor circuits of the basal ganglia. Eur J Neurosci 30:1269–1279. https://doi.org/10.1111/j.1460-9568.2009.06907.x

Alliende J, Giret N, Pidoux L et al (2017) Seasonal plasticity of song behavior relies on motor and syntactic variability induced by a basal ganglia–forebrain circuit. Neuroscience 359:49–68. https://doi.org/10.1016/j.neuroscience.2017.07.007

Andalman AS, Fee MS (2009) A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. Proc Natl Acad Sci U S A 106:12518–12523. https://doi.org/10.1073/pnas.0903214106

Arbib MA (2008) From grasp to language: embodied concepts and the challenge of abstraction. J Physiol Paris 102:4–20. https://doi.org/10.1016/j.jphysparis.2008.03.001

Aronov D, Andalman AS, Fee MS (2008) A specialized forebrain circuit for vocal babbling in the juvenile songbird. Science 320:630–634. https://doi.org/10.1126/science.1155140

Atallah HE, Frank MJ, O'Reilly RC (2004) Hippocampus, cortex, and basal ganglia: insights from computational models of complementary learning systems. Neurobiol Learn Mem 82:253–267. https://doi.org/10.1016/j.nlm.2004.06.004

Barnes TD, Kubota Y, Hu D et al (2005) Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature 437:1158–1161. https://doi.org/10.1038/nature04053

Barnes TD, Mao J-B, Hu D et al (2011) Advance cueing produces enhanced action-boundary patterns of spike activity in the sensorimotor striatum. J Neurophysiol 105:1861–1878. https://doi.org/10.1152/jn.00871.2010

Bernstein NA (1967) The co-ordination and regulation of movements: conclusions towards the study of motor co-ordination. Biodyn Locomot:104–113

Bodor AL, Giber K, Rovo Z et al (2008) Structural correlates of efficient GABAergic transmission in the basal ganglia-thalamus pathway. J Neurosci 28:3090–3102. https://doi.org/10.1523/JNEUROSCI.5266-07.2008

Boettiger CA, Doupe AJ (2001) Developmentally restricted synaptic plasticity in a songbird nucleus required for song learning. Neuron 31:809–818. https://doi.org/10.1016/S0896-6273(01)00403-2

Bokor H, Frère SGA, Eyre MD et al (2005) Selective GABAergic control of higher-order thalamic relays. Neuron 45:929–940. https://doi.org/10.1016/j.neuron.2005.01.048

Boraud T, Bezard E, Bioulac B, Gross CE (2002) From single extracellular unit recording in experimental and human parkinsonism to the development of a functional concept of the role played by the basal ganglia in motor control. Prog Neurobiol 66:265–283. https://doi.org/10.1016/S0301-0082(01)00033-8

Bottjer SW (1993) The distribution of tyrosine hydroxylase immunoreactivity in the brains of male and female zebra finches. J Neurobiol 24:51–69. https://doi.org/10.1002/neu.480240105

Bottjer SW, Miesner EA, Arnold AP (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. Science 224:901–903. https://doi.org/10.1126/science.6719123

Brainard MS, Doupe AJ (2000) Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. Nature 404:762–766. https://doi.org/10.1038/35008083

Brainard MS, Doupe AJ (2013) Translating birdsong: songbirds as a model for basic and applied medical research. Annu Rev Neurosci 36:489–517. https://doi.org/10.1146/annurev-neuro-060909-152826

Brenowitz EA, Beecher MD (2005) Song learning in birds: diversity and plasticity, opportunities and challenges. Trends Neurosci 28:127–132. https://doi.org/10.1016/j.tins.2005.01.004

Budzillo A, Duffy A, Miller KE et al (2017) Dopaminergic modulation of basal ganglia output through coupled excitation–inhibition. Proc Natl Acad Sci 114:5713–5718. https://doi.org/10.1073/pnas.1611146114

Canales JJ, Graybiel AM (2000) A measure of striatal function predicts motor stereotypy. Nat Neurosci 3:377–383. https://doi.org/10.1038/73949

Cardin JA, Schmidt MF (2003) Song system auditory responses are stable and highly tuned during sedation, rapidly modulated and unselective during wakefulness, and suppressed by arousal. J Neurophysiol 90:2884–2899. https://doi.org/10.1152/jn.00391.2003

Carrillo GD, Doupe AJ (2004) Is the songbird area X striatal, pallidal, or both? An anatomical study. J Comp Neurol 473:415–437. https://doi.org/10.1002/cne.20099

Casto JM, Ball GF (1994) Characterization and localization of D1 dopamine receptors in the sexually dimorphic vocal control nucleus, area X, and the basal ganglia of european starlings. J Neurobiol 25:767–780. https://doi.org/10.1002/neu.480250703

Chaminade T, Oztop E, Cheng G, Kawato M (2008) From self-observation to imitation: visuomotor association on a robotic hand. Brain Res Bull 75:775–784. https://doi.org/10.1016/j.brainresbull.2008.01.016

Charlesworth JD, Warren TL, Brainard MS (2012) Covert skill learning in a cortical-basal ganglia circuit. Nature 486:251–255. https://doi.org/10.1038/nature11078

Chen JR, Stepanek L, Doupe AJ (2014) Differential contributions of basal ganglia and thalamus to song initiation, tempo, and structure. J Neurophysiol 111:248–257. https://doi.org/10.1152/jn.00584.2012

Chen Y, Clark O, Woolley SC (2017) Courtship song preferences in female zebra finches are shaped by developmental auditory experience. Proc R Soc B Biol Sci 284:20170054. https://doi.org/10.1098/rspb.2017.0054

Crair MC, Malenka RC (1995) A critical period for long-term potentiation at thalamocortical synapses. Nature 375:325–328. https://doi.org/10.1038/375325a0

Darshan R, Wood WE, Peters S et al (2017) A canonical neural mechanism for behavioral variability. Nat Commun 8. https://doi.org/10.1038/ncomms15415

DeLong MR (1990) Primate models of movement disorders of basal ganglia origin. Trends Neurosci 13:281–285. https://doi.org/10.1016/0166-2236(90)90110-V

Deniau JM, Chevalier G (1985) Disinhibition as a basic process in the expression of striatal functions. II. The striato-nigral influence on thalamocortical cells of the ventromedial thalamic nucleus. Brain Res 334:227–233. https://doi.org/10.1016/0006-8993(85)90214-8

Desmurget M, Turner RS (2008) Testing basal ganglia motor functions through reversible inactivations in the posterior internal globus pallidus. J Neurophysiol 99:1057–1076. https://doi.org/10.1152/jn.01010.2007

Dhawale AK, Smith MA, Ölveczky BP (2017) The role of variability in motor learning. Annu Rev Neurosci 40:479–498. https://doi.org/10.1146/annurev-neuro-072116-031548

Ding L, Perkel DJ (2004) Long-term potentiation in an avian basal ganglia nucleus essential for vocal learning. J Neurosci 24:488–494. https://doi.org/10.1523/JNEUROSCI.4358-03.2004

Ding L, Perkel DJ (2014) Two tales of how expectation of reward modulates behavior. Curr Opin Neurobiol 29:142–147. https://doi.org/10.1016/j.conb.2014.07.011

Doupe AJ (1997) Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. J Neurosci 17:1147–1167. https://doi.org/10.1523/JNEUROSCI.17-03-01147.1997

Doupe AJ, Perkel DJ, Reiner A, Stern EA (2005) Birdbrains could teach basal ganglia research a new song. Trends Neurosci 28:353–363. https://doi.org/10.1016/j.tins.2005.05.005

Doya K (2000) Complementary roles of basal ganglia and cerebellum in learning and motor control. Curr Opin Neurobiol 10:732–739. https://doi.org/10.1016/S0959-4388(00)00153-7

Doya K, Sejnowski T (1995) A novel reinforcement model of birdsong vocalization learning. Adv Neural Inf Process Syst 130:101–108. https://doi.org/10.1093/brain/awm214

Doya K, Sejnowski TJ (1998) A computational model of birdsong learning by auditory experience and auditory feedback. Cent Audit Process Neural Model:77–88. https://doi.org/10.1007/978-1-4615-5351-9_8

Dudman JT, Krakauer JW (2016) The basal ganglia: from motor commands to the control of vigor. Curr Opin Neurobiol 37:158–166. https://doi.org/10.1016/j.conb.2016.02.005

Düring DN, Ziegler A, Thompson CK et al (2013) The songbird syrinx morphome: a three-dimensional, high-resolution, interactive morphological map of the zebra finch vocal organ. BMC Biol 11:1. https://doi.org/10.1186/1741-7007-11-1

Edgerton JR, Jaeger D (2014) Optogenetic activation of nigral inhibitory inputs to motor thalamus in the mouse reveals classic inhibition with little potential for rebound activation. Front Cell Neurosci 8:1–11. https://doi.org/10.3389/fncel.2014.00036

Elemans CPH (2014) The singer and the song: the neuromechanics of avian sound production. Curr Opin Neurobiol 28:172–178. https://doi.org/10.1016/j.conb.2014.07.022

Farries MA, Perkel DJ (2000) Electrophysiological properties of avian basal ganglia neurons recorded in vitro. J Neurophysiol 84:2502–2513. https://doi.org/10.1152/jn.2000.84.5.2502

Farries MA, Perkel DJ (2002) A telencephalic nucleus essential for song learning contains neurons with physiological characteristics of both striatum and globus pallidus. J Neurosci 22:3776–3787. https://doi.org/10.1523/JNEUROSCI.22-09-03776.2002

Farries MA, Ding L, Perkel DJ (2005) Evidence for "direct" and "indirect" pathways through the song system basal ganglia. J Comp Neurol 484:93–104. https://doi.org/10.1002/cne.20464

Fee MS (2012) Oculomotor learning revisited: a model of reinforcement learning in the basal ganglia incorporating an efference copy of motor actions. Front Neural Circuits 6:1–18. https://doi.org/10.3389/fncir.2012.00038

Fee MS (2014) The role of efference copy in striatal learning. Curr Opin Neurobiol 25:194–200. https://doi.org/10.1016/j.conb.2014.01.012

Fee MS, Goldberg JH (2011) A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. Neuroscience 198:1–19. https://doi.org/10.1016/j.neuroscience.2011.09.069

Fee MS, Shraiman B, Pesaran B, Mitra PP (1998) The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird. Nature 395:67–71. https://doi.org/10.1038/25725

Fiete IR, Fee MS, Seung HS (2007) Model of birdsong learning based on gradient estimation by dynamic perturbation of neural conductances. J Neurophysiol 98:2038–2057. https://doi.org/10.1152/jn.01311.2006

Florian RV (2007) Reinforcement learning through modulation of spike-timing-dependent synaptic plasticity. Neural Comput 19:1468–1502. https://doi.org/10.1162/neco.2007.19.6.1468

Frank MJ, Doll BB, Oas-Terpstra J, Moreno F (2009) Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. Nat Neurosci 12:1062–1068. https://doi.org/10.1038/nn.2342

Gadagkar V, Puzerey PA, Chen R et al (2016) Dopamine neurons encode performance error in singing birds. Science (80-) 354:1278–1282. https://doi.org/10.1126/science.aah6837

Gale SD, Perkel DJ (2010a) A basal ganglia pathway drives selective auditory responses in songbird dopaminergic neurons via disinhibition. J Neurosci 30:1027–1037. https://doi.org/10.1523/JNEUROSCI.3585-09.2010

Gale SD, Perkel DJ (2010b) Anatomy of a songbird basal ganglia circuit essential for vocal learning and plasticity. J Chem Neuroanat 39:124–131. https://doi.org/10.1016/j.jchemneu.2009.07.003

Garst-Orozco J, Babadi B, Ölveczky BP (2014) A neural circuit mechanism for regulating vocal variability during song learning in zebra finches. elife 3:e03697. https://doi.org/10.7554/eLife.03697

Giret N, Kornfeld J, Ganguli S, Hahnloser RHR (2014) Evidence for a causal inverse model in an avian cortico-basal ganglia circuit. Proc Natl Acad Sci 111:6063–6068. https://doi.org/10.1073/pnas.1317087111

Goldberg JH, Fee MS (2012) A cortical motor nucleus drives the basal ganglia-recipient thalamus in singing birds. Nat Neurosci 15:620–627. https://doi.org/10.1038/nn.3047

Goldberg JH, Adler A, Bergman H, Fee MS (2010) Singing-related neural activity distinguishes two putative pallidal cell types in the songbird basal ganglia: comparison to the primate internal and external pallidal segments. J Neurosci 30:7088–7098. https://doi.org/10.1523/JNEUROSCI.0168-10.2010

Graybiel AM (1998) The basal ganglia and chunking of action repertoires. Neurobiol Learn Mem 70:119–136

Graybiel AM (2005) The basal ganglia: learning new tricks and loving it. Curr Opin Neurobiol 15:638–644. https://doi.org/10.1016/j.conb.2005.10.006

Graybiel AM (2008) Habits, rituals, and the evaluative brain. Annu Rev Neurosci 31:359–387. https://doi.org/10.1146/annurev.neuro.29.051605.112851

Graybiel AM, Hirscht EC, Agidt YA (1987) Differences in tyrosine hydroxylase-like immunoreactivity characterize the mesostriatal innervation of striosomes and extrastriosomal matrix at maturity. Neurobiology 84:303–307. https://doi.org/10.1073/pnas.84.1.303

Grillner S, Hellgren J, Ménard A et al (2005) Mechanisms for selection of basic motor programs--roles for the striatum and pallidum. Trends Neurosci 28:364–370. https://doi.org/10.1016/j.tins.2005.05.004

Haesler S, Wada K, Nshdejan A et al (2004) FoxP2 expression in avian vocal learners and non-learners. J Neurosci 24:3164–3175. https://doi.org/10.1523/JNEUROSCI.4369-03.2004

Hahnloser RHR, Ganguli S (2013) Vocal learning with inverse models. Princ Neural Coding:547–564. https://doi.org/10.1201/b14756-32

Halassa MM, Acsády L (2016) Thalamic inhibition: diverse sources, diverse scales. Trends Neurosci 39:680–693. https://doi.org/10.1016/j.tins.2016.08.001

Hamaguchi K, Mooney R (2012) Recurrent interactions between the input and output of a songbird Cortico-basal ganglia pathway are implicated in vocal sequence variability. J Neurosci 32:11671–11687. https://doi.org/10.1523/JNEUROSCI.1666-12.2012

Hamaguchi K, Tschida KA, Yoon I et al (2014) Auditory synapses to song premotor neurons are gated off during vocalization in zebra finches. elife 3. https://doi.org/10.7554/eLife.01833

Hampton CM, Sakata JT, Brainard MS (2009) An avian basal ganglia-forebrain circuit contributes differentially to syllable versus sequence variability of adult Bengalese finch song. J Neurophysiol 101:3235–3245. https://doi.org/10.1152/jn.91089.2008

Hanuschkin A, Ganguli S, Hahnloser RHR (2013) A Hebbian learning rule gives rise to mirror neurons and links them to control theoretic inverse models. Front Neural Circuits 7:1–15. https://doi.org/10.3389/fncir.2013.00106

Hessler NA, Doupe AJ (1999) Singing-related neural activity in a dorsal forebrain-basal ganglia circuit of adult zebra finches. J Neurosci 19:10461–10481. https://doi.org/10.1523/JNEUROSCI.19-23-10461.1999

Heston JB, Simon J, Day NF et al (2018) Bidirectional scaling of vocal variability by an avian cortico-basal ganglia circuit. Physiol Rep 6:e13638. https://doi.org/10.14814/phy2.13638

Heyes C (2001) Causes and consequences of imitation. Trends Cogn Sci 5:253–261. https://doi.org/10.1016/S1364-6613(00)01661-2

Higley MJ, Gittis AH, Oldenburg IA et al (2011) Cholinergic interneurons mediate fast VGluT3-dependent glutamatergic transmission in the striatum. PLoS One 6:e19155. https://doi.org/10.1371/journal.pone.0019155

Hikosaka O (2007) Basal ganglia mechanisms of reward-oriented eye movement. Ann N Y Acad Sci 1104:229–249. https://doi.org/10.1196/annals.1390.012

Hikosaka O, Takikawa Y, Kawagoe R (2000) Role of the basal ganglia in the control of purposive saccadic eye movements. Physiol Rev 80:953–978. https://doi.org/10.1152/physrev.2000.80.3.953

Hisey E, Kearney MG, Mooney R (2018) A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. Nat Neurosci 21:589–597. https://doi.org/10.1038/s41593-018-0092-6

Hoffmann LA, Saravanan V, Wood AN et al (2016) Dopaminergic contributions to vocal learning. J Neurosci 36:2176–2189. https://doi.org/10.1523/JNEUROSCI.3883-15.2016

Houk JC, Wise SP (1995) Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex their role in planning and controlling action. Cereb Cortex 5:95–110. https://doi.org/10.1093/cercor/5.2.95

Iacoboni M (1999) Cortical mechanisms of human imitation. Science 286:2526–2528. https://doi.org/10.1126/science.286.5449.2526

Iyengar S, Viswanathan SS, Bottjer SW (1999) Development of topography within song control circuitry of zebra finches during the sensitive period for song learning. J Neurosci 19:6037–6057. https://doi.org/10.1523/JNEUROSCI.19-14-06037.1999

James LS, Sakata JT (2014) Vocal motor changes beyond the sensitive period for song plasticity. J Neurophysiol 112:2040–2052. https://doi.org/10.1152/jn.00217.2014

Jarvis E, Güntürkün O, Bruce L et al (2005) Avian brains and a new understanding of vertebrate brain evolution. Nat Rev Neurosci 6:151–159. https://doi.org/10.1038/nrn1606

Kao MH, Brainard MS (2006) Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. J Neurophysiol 96:1441–1455. https://doi.org/10.1152/jn.01138.2005

Kao MH, Doupe AJ, Brainard MS (2005) Contribution of an avian basal ganglia-forebrain circuit to real-time modulation of song. Nature 433:638–643. https://doi.org/10.1038/nature03127

Kao MH, Wright BD, Doupe AJ (2008) Neurons in a forebrain nucleus required for vocal plasticity rapidly switch between precise firing and variable bursting depending on social context. J Neurosci 28:13232–13247. https://doi.org/10.1523/JNEUROSCI.2250-08.2008

Karten HJ, Dubbeldam JL (1973) The organization and projections of the paleostriatal complex in the pigeon (columba livia). J Comp Neurol 148:61–89. https://doi.org/10.1002/cne.901480105

Kawaguchi Y, Wilson CJ, Augood SJ, Emson PC (1995) Striatal interneurones: chemical, physiological and morphological characterization. Trends Neurosci 18:527–535. https://doi.org/10.1016/0166-2236(95)98374-8

Keller GB, Hahnloser RHR (2009) Neural processing of auditory feedback during vocal practice in a songbird. Nature 457:187–190. https://doi.org/10.1038/nature07467

Kim J, Kim Y, Nakajima R et al (2017) Inhibitory basal ganglia inputs induce excitatory motor signals in the thalamus. Neuron 95:1181–1196. https://doi.org/10.1016/j.neuron.2017.08.028

Kobayashi K, Uno H, Okanoya K (2001) Partial lesions in the anterior forebrain pathway affect song production in adult Bengalese finches. Neuroreport 12:353–358. https://doi.org/10.1097/00001756-200102120-00034

Kojima S, Doupe AJ (2009) Activity propagation in an avian basal ganglia-thalamocortical circuit essential for vocal learning. J Neurosci 29:4782–4793. https://doi.org/10.1523/JNEUROSCI.4903-08.2009

Konishi M (2004) The role of auditory feedback in birdsong. Ann N Y Acad Sci 1016:463–475. https://doi.org/10.1196/annals.1298.010

Kosubek-Langer J, Schulze L, Scharff C (2017) Maturation, behavioral activation, and connectivity of adult-born medium spiny neurons in a striatal song nucleus. Front Neurosci 11:1–12. https://doi.org/10.3389/fnins.2017.00323

Kozhevnikov AA, Fee MS (2007) Singing-related activity of identified HVC neurons in the zebra finch. J Neurophysiol 97:4271–4283. https://doi.org/10.1152/jn.00952.2006

Kravitz AV, Freeze BS, Parker PRL et al (2010) Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. Nature 466:622–626. https://doi.org/10.1038/nature09159

Kreitzer AC, Malenka RC (2008) Striatal plasticity and basal ganglia circuit function. Neuron 60:543–554. https://doi.org/10.1016/j.neuron.2008.11.005

Kubikova L, Kostál L (2010) Dopaminergic system in birdsong learning and maintenance. J Chem Neuroanat 39:112–123. https://doi.org/10.1016/j.jchemneu.2009.10.004

Kubikova L, Bosikova E, Cvikova M et al (2014) Basal ganglia function, stuttering, sequencing, and repair in adult songbirds. Sci Rep 4. https://doi.org/10.1038/srep06590

Lashley KS (1933) Integrative functions of the cerebral cortex. Physiol Rev 13:1–43. https://doi.org/10.1152/physrev.1933.13.1.1

Leblois A, Perkel DJ (2012) Striatal dopamine modulates song spectral but not temporal features through D1 receptors. Eur J Neurosci 35:1–11. https://doi.org/10.1111/j.1460-9568.2012.08095.x

Leblois A, Bodor AL, Person AL, Perkel DJ (2009) Millisecond timescale disinhibition mediates fast information transmission through an avian basal ganglia loop. J Neurosci 29:15420–15433. https://doi.org/10.1523/JNEUROSCI.3060-09.2009

Leblois A, Wendel BJ, Perkel DJ (2010) Striatal dopamine modulates basal ganglia output and regulates social context-dependent behavioral variability through D1 receptors. J Neurosci 30:5730–5743. https://doi.org/10.1523/JNEUROSCI.5974-09.2010

Liberti WA, Markowitz JE, Perkins LN et al (2016) Unstable neurons underlie a stable learned behavior. Nat Neurosci 19:1665–1671. https://doi.org/10.1038/nn.4405

Liu W-C, Hruska-Plochan M, Miyanohara A (2017) Lentiviral-mediated Transgenesis in songbirds. Methods Mol Biol 1650:149–165. https://doi.org/10.1007/978-1-4939-7216-6_9

Luo M, Perkel DJ (1999) A GABAergic, strongly inhibitory projection to a thalamic nucleus in the zebra finch song system. J Neurosci 19:6700–6711. https://doi.org/10.1523/JNEUROSCI.19-15-06700.1999

Luo M, Ding L, Perkel DJ (2001) An avian basal ganglia pathway essential for vocal learning forms a closed topographic loop. J Neurosci 21:6836–6845. https://doi.org/10.1523/JNEUROSCI.21-17-06836.2001

Mandelblat-Cerf Y, Las L, Denisenko N, Fee MS (2014) A role for descending auditory cortical projections in songbird vocal learning. elife 3:1–23. https://doi.org/10.7554/eLife.02152

Margoliash D (1997) Functional organization of forebrain pathways for song production and perception. J Neurobiol 33:671–693. https://doi.org/10.1002/(SICI)1097-4695(19971105)33:5<671::AID-NEU12>3.0.CO;2-C

Marler P, Pickert R (1984) Species-universal microstructure in the learned song of the swamp sparrow (Melospiza georgiana). Anim Behav 32:673–689. https://doi.org/10.1016/S0003-3472(84)80143-8

Marler P, Waser MS (1977) Role of auditory feedback in canary song development. J Comp Physiol Psychol 91:8–16. https://doi.org/10.1037/h0077303

McCasland JS (1987) Neuronal control of bird song production. J Neurosci 7:23–39. https://doi.org/10.1523/JNEUROSCI.07-01-00023.1987

Miller JE, Hafzalla GW, Burkett ZD et al (2015) Reduced vocal variability in a zebra finch model of dopamine depletion: implications for Parkinson disease. Physiol Rep 3:e12599. https://doi.org/10.14814/phy2.12599

Mink JW, Thach WT (1993) Basal ganglia intrinsic circuits and their role in behavior. Curr Opin Neurobiol 3:950–957. https://doi.org/10.1016/0959-4388(93)90167-W

Mooney R (2009) Neural mechanisms for learned birdsong. Learn Mem 16:655–669. https://doi.org/10.1101/lm.1065209

Mooney R, Konishi M (1991) Two distinct inputs to an avian song nucleus activate different glutamate receptor subtypes on individual neurons. Proc Natl Acad Sci U S A 88:4075–4079. https://doi.org/10.1073/pnas.88.10.4075

Nakamura K, Hikosaka O (2006) Facilitation of saccadic eye movements by Postsaccadic electrical stimulation in the primate caudate. J Neurosci 26:12885–12895. https://doi.org/10.1523/JNEUROSCI.3688-06.2006

Nambu A, Tokuno H, Hamada I et al (2000) Excitatory cortical inputs to pallidal neurons via the subthalamic nucleus in the monkey. J Neurophysiol 84:289–300. https://doi.org/10.1152/jn.2000.84.1.289

Olveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS Biol 3:e153. https://doi.org/10.1371/journal.pbio.0030153

Olveczky BP, Otchy TM, Goldberg JH et al (2011) Changes in the neural control of a complex motor sequence during learning. J Neurophysiol 106:386–397. https://doi.org/10.1152/jn.00018.2011

Otchy TM, Wolff SBE, Rhee JY et al (2015) Acute off-target effects of neural circuit manipulations. Nature 528:358–363. https://doi.org/10.1038/nature16442

Oudeyer PY (2005) The self-organization of speech sounds. J Theor Biol 233:435–449. https://doi.org/10.1016/j.jtbi.2004.10.025

Oztop E, Kawato M, Arbib M (2006) Mirror neurons and imitation: a computationally guided review. Neural Netw 19:254–271. https://doi.org/10.1016/j.neunet.2006.02.002

Parr R, Russell S (1998) Reinforcement learning with hierarchies of machines. Neural Inf Process Syst 104:1043–1049

Pasupathy A, Miller EK (2005) Different time courses of learning-related activity in the prefrontal cortex and striatum. Nature 433:873–876. https://doi.org/10.1038/nature03287

Person AL, Perkel DJ (2005) Unitary IPSPs drive precise thalamic spiking in a circuit required for learning. Neuron 46:129–140. https://doi.org/10.1016/j.neuron.2004.12.057

Person AL, Perkel DJ (2007) Pallidal neuron activity increases during sensory relay through thalamus in a songbird circuit essential for learning. J Neurosci 27:8687–8698. https://doi.org/10.1523/JNEUROSCI.2045-07.2007

Person AL, Gale SD, Farries MA, Perkel DJ (2008) Organization of the songbird basal ganglia, including area X. J Comp Neurol 508:840–866. https://doi.org/10.1002/cne.21699

Picardo MA, Merel J, Katlowitz KA et al (2016) Population-level representation of a temporal sequence underlying song production in the Zebra finch. Neuron 90:866–876. https://doi.org/10.1016/j.neuron.2016.02.016

Piron C, Kase D, Topalidou M et al (2016) The globus pallidus pars interna in goal-oriented and routine behaviors: resolving a long-standing paradox. Mov Disord 31:1146–1154. https://doi.org/10.1002/mds.26542

Prather JF, Peters S, Nowicki S, Mooney R (2008) Precise auditory-vocal mirroring in neurons for learned vocal communication. Nature 451:305–310. https://doi.org/10.1038/nature06492

Reiner A, Perkel DJ, Bruce LL et al (2004) Revised nomenclature for avian telencephalon and some related brainstem nuclei. J Comp Neurol 473:377–414. https://doi.org/10.1002/cne.20118

Rizzolatti G, Fadiga L, Gallese V, Fogassi L (1996) Premotor cortex and the recognition of motor actions. Brain Res Cogn Brain Res 3:131–141

Roberts TF, Gobes SMH, Murugan M et al (2012) Motor circuits are required to encode a sensory model for imitative learning. Nat Neurosci 15:1454–1459. https://doi.org/10.1038/nn.3206

Rolf M, Steil JJ, Gienger M (2010) Goal babbling permits direct learning of inverse kinematics. IEEE Trans Auton Ment Dev 2:216–229. https://doi.org/10.1109/TAMD.2010.2062511

Sakata JT, Vehrencamp SL (2012) Integrating perspectives on vocal performance and consistency. J Exp Biol 215:201–209. https://doi.org/10.1242/jeb.056911

Sakata JT, Hampton CM, Brainard MS (2008) Social modulation of sequence and syllable variability in adult birdsong. J Neurophysiol 99:1700–1711. https://doi.org/10.1152/jn.01296.2007

Sasaki A, Sotnikova TD, Gainetdinov RR, Jarvis ED (2006) Social context-dependent singing-regulated dopamine. J Neurosci 26:9010–9014. https://doi.org/10.1523/JNEUROSCI.1335-06.2006

Scharff C, Nottebohm F (1991) A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. J Neurosci 11:2896–2913. https://doi.org/10.1523/JNEUROSCI.11-09-02896.1991

Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275:1593–1599. https://doi.org/10.1126/science.275.5306.1593

Sober SJ, Wohlgemuth MJ, Brainard MS (2008) Central contributions to acoustic variation in birdsong. J Neurosci 28:10370–10379. https://doi.org/10.1523/JNEUROSCI.2448-08.2008

Solis MM, Doupe AJ (1997) Anterior forebrain neurons develop selectivity by an intermediate stage of birdsong learning. J Neurosci 17:6447–6462. https://doi.org/10.1523/JNEUROSCI.17-16-06447.1997

Srivastava KH, Elemans CPH, Sober SJ (2015) Multifunctional and context-dependent control of vocal acoustics by individual muscles. J Neurosci 35:14183–14194. https://doi.org/10.1523/JNEUROSCI.3610-14.2015

Stark LL, Perkel DJ (1999) Two-stage, input-specific synaptic maturation in a nucleus essential for vocal production in the zebra finch. J Neurosci 19:9107–9116. https://doi.org/10.1523/JNEUROSCI.19-20-09107.1999

Stephenson-Jones M, Samuelsson E, Ericsson J et al (2011) Evolutionary conservation of the basal ganglia as a common vertebrate mechanism for action selection. Curr Biol 21:1081–1091. https://doi.org/10.1016/j.cub.2011.05.001

Suri RE, Schultz W (1999) A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. Neuroscience 91:871–890. https://doi.org/10.1016/S0306-4522(98)00697-6

Sutton RS, Barto AG (1981) Toward a modern theory of adaptive networks: expectation and prediction. Psychol Rev 88:135–170. https://doi.org/10.1037/0033-295X.88.2.135

Tanaka M, Alvarado JS, Murugan M, Mooney R (2016) Focal expression of mutant huntingtin in the songbird basal ganglia disrupts cortico-basal ganglia networks and vocal sequences. Proc Natl Acad Sci 113:E1720–727. https://doi.org/10.1073/pnas.1523754113

Tanaka M, Sun F, Li Y, Mooney R (2018) A mesocortical dopamine circuit enables the cultural transmission of vocal behaviour. Nature 563:117–120. https://doi.org/10.1038/s41586-018-0636-7

Tchernichovski O, Nottebohm F, Ho C et al (2000) A procedure for an automated measurement of song similarity. Anim Behav 59:1167–1176. https://doi.org/10.1006/anbe.1999.1416

Tecuapetla F, Patel JC, Xenias H et al (2010) Glutamatergic signaling by mesolimbic dopamine neurons in the nucleus accumbens. J Neurosci 30:7105–7110. https://doi.org/10.1523/JNEUROSCI.0265-10.2010

Vicario DS (1991) Organization of the zebra finch song control system: functional organization of outputs from nucleus robustus archistriatalis. J Comp Neurol 309:486–494. https://doi.org/10.1002/cne.903090405

Vicario DS, Nottebohm F (1988) Organization of the zebra finch song control system: I. representation of syringeal muscles in the hypoglossal nucleus. J Comp Neurol 271:346–354. https://doi.org/10.1002/cne.902710305

Voon V (2017) Chapter 24 – Decision-making and impulse control disorders in Parkinson's disease. In: Dreher JC, Tremblay L (eds) Decision Neuroscience. Academic, pp 305–314. https://doi.org/10.1016/B978-0-12-805308-9.00024-5

Wanaverbecq N, Bodor AL, Bokor H et al (2008) Contrasting the functional properties of GABAergic axon terminals with single and multiple synapses in the thalamus. J Neurosci 28:11848–11861. https://doi.org/10.1523/JNEUROSCI.3183-08.2008

Wanjerkhede SM, Bapi RS (2011) Role of CAMKII in reinforcement learning: a computational model of glutamate and dopamine signaling pathways. Biol Cybern 104:397–424. https://doi.org/10.1007/s00422-011-0439-5

Warren TL, Tumer EC, Charlesworth JD, Brainard MS (2011) Mechanisms and time Courseof vocal learning and consolidation in the adult songbird. J Neurophysiol 106:1806–1821. https://doi.org/10.1152/jn.00311.2011

Westermann G, Miranda ER (2004) A new model of sensorimotor coupling in the development of speech. Brain Lang 89:393–400. https://doi.org/10.1016/S0093-934X(03)00345-6

Woolley SC (2016) Social context differentially modulates activity of two interneuron populations in an avian basal ganglia nucleus. J Neurophysiol 116:2831–2840. https://doi.org/10.1152/jn.00622.2016

Woolley SC (2019) Dopaminergic regulation of vocal-motor plasticity and performance. Curr Opin Neurobiol 54:127–133. https://doi.org/10.1016/j.conb.2018.10.008

Woolley SC, Doupe AJ (2008) Social context-induced song variation affects female behavior and gene expression. PLoS Biol 6:525–537. https://doi.org/10.1371/journal.pbio.0060062

Woolley SC, Kao MH (2015) Variability in action: CONTRIBUTIONS of a songbird cortical-basal ganglia circuit to vocal motor learning and control. Neuroscience 296:39–47. https://doi.org/10.1016/j.neuroscience.2014.10.010

Woolley SC, Rajan R, Joshua M, Doupe AJ (2014) Emergence of context-dependent variability across a basal ganglia network. Neuron 82:208–223. https://doi.org/10.1016/j.neuron.2014.01.039

Xiao L, Chattree G, Oscos FG et al (2018) A basal ganglia circuit sufficient to guide birdsong learning. Neuron 98:208–221. https://doi.org/10.1016/j.neuron.2018.02.020

Yanagihara S, Hessler NA (2006) Modulation of singing-related activity in the songbird ventral tegmental area by social context. Eur J Neurosci 24:3619–3627. https://doi.org/10.1111/j.1460-9568.2006.05228.x

Yanagihara S, Hessler NA (2012) Phasic basal ganglia activity associated with high-gamma oscillation during sleep in a songbird. J Neurophysiol 107:424–432. https://doi.org/10.1152/jn.00790.2011

Yartsev MM (2017) The emperor's new wardrobe: rebalancing diversity of animal models in neuroscience research. Science 358:466–469. https://doi.org/10.1126/science.aan8865

Yazaki-Sugiyama Y, Yanagihara S, Fuller PM, Lazarus M (2015) Acute inhibition of a cortical motor area impairs vocal control in singing zebra finches. Eur J Neurosci 41:97–108. https://doi.org/10.1111/ejn.12757

Chapter 5
Integrating Form and Function in the Songbird Auditory Forebrain

Sarah C. Woolley and Sarah M. N. Woolley

Abstract Vocal communication is critical for reproduction and survival across a wide range of species. For vocal communication systems to function, receivers must perform a range of auditory tasks to decode and process acoustic signals. In songbirds, learned vocal signals (songs) can be used by receivers to gain information about the species, sex, identity, and even motivation of the singer. Moreover, young songbirds must hear and memorize songs during development to use them as templates for song learning. This chapter reviews research on the structure and function of the songbird auditory system. In particular, the relationships between the organization, connections, and information-coding properties of the auditory pallium are described and how the functions of those circuits allow birds to perform a range of auditory tasks is considered, including individual recognition, tutor song learning, auditory memory, and mate choice processes.

Keywords Auditory cortex · Birdsong perception · CMM · Field L · Learning · Memory · NCM · Song preference · Territoriality

5.1 Introduction

Birdsong is an acoustic communication signal used in a wide range of contexts that include courtship interactions and territory advertisement. Song behavior varies substantially among the over 5000 songbird species, with species-specific variation in vocal learning, sex-specific patterns in song use, the number of songs that one

---

S. C. Woolley (✉)
Department of Biology and Centre for Research in Brain, Language, and Music,
McGill University, Montreal, QC, Canada
e-mail: sarah.woolley@mcgill.ca

S. M. N. Woolley
Zuckerman Institute and Department of Psychology, Columbia University,
Jerome L. Greene Science Center, New York, NY, USA
e-mail: sw2277@columbia.edu

J. T. Sakata et al. (eds.), *The Neuroethology of Birdsong*, Springer Handbook
of Auditory Research 71, https://doi.org/10.1007/978-3-030-34683-6_5

individual sings, and the acoustic features of song (see Sakata and Woolley, Chap. 1). Moreover, songbirds are one of only a few taxa that learn their songs. Consequently, many of the acoustic features of an individual's song are unique to that individual, making song a signal that conveys individual identity in addition to species, sex, location, and breeding condition. Songbirds of both sexes use the unique songs of familiar individuals to maintain social relationships with mates and territory neighbors (Catchpole and Slater 2008). The immense diversity of song behavior across species and individuals provides the opportunity to identify functional relationships between the neural circuits for auditory processing and vocal communication behavior. For example, the unique songs of individuals can be used as probes to investigate the neural mechanisms of vocal perception, including those that underlie learning, memory, sensorimotor integration, vocal production, and mate choice.

Despite the impressive diversity of birdsong across species and individuals, there exist common principles of auditory processing underlying song learning, perception, and production among species. Across many songbird species, males learn to sing as juveniles and use their adult songs to court females and to engage in aggressive exchanges with other males (see Sakata and Yazaki-Sugiyama, Chap. 2). The acoustic features of male song convey honest information about reproductive fitness to listeners (Beecher and Brenowitz 2005; Richner 2016). Additionally, the acoustic properties of male song drive female attraction to males; females use song as a mate choice cue (Riebel 2009). Consequently, females evaluate songs to choose males that will contribute to the next generation, which places a premium on auditory processing by females and on song learning and performance by males (see Podos and Seung, Chap. 9).

While the importance of hearing for song perception is obvious, determining the importance of hearing during development and the degree to which it shapes song production and perception has required experimental studies that manipulate auditory experience and analyze the effects of those manipulations on song and preference behavior. Those studies have shown that auditory exposure to adult song is required for song to develop normally and that auditory feedback is required for both song development and maintenance (Brainard and Doupe 2000; Murphy, Lawley, Smith, and Prather, Chap. 3). Additionally, song exposure is necessary for the display of some species-typical song preferences in adulthood (Lauay et al. 2004; Chen et al. 2017). The multiple ways in which song behavior depends on hearing illustrate that auditory coding is a fundamental form of neural processing in song communication.

This chapter describes the structure and function of song with particular focus on the song-related auditory tasks that birds perform to perceive and process communication signals. In light of those behavioral functions, the chapter then describes the organization, connections, and information-coding properties of the auditory pallium with particular emphasis on its roles in species and individual recognition, tutor song learning, and mate choice processes. Throughout, the chapter highlights the homologies between avian and mammalian auditory systems and the unique advantages that songbirds afford to the study of auditory processing.

5.2 Structure of Song

The complex acoustic structure of birdsong provides the dimensionality for the diversity that has been documented across species and individuals. A bird's *song* is a sequence of complex sound units, hierarchically organized into notes, syllables, motifs, and bouts (Fig. 5.1). *Notes* are the smallest acoustic units in song and may be produced alone or grouped in time to form *syllables*. Syllables are therefore composed of one or multiple notes. *Motifs* (also called phrases or strophes) are stereotyped sequences of syllables. Birds of some species produce multiple different motifs; others repeat the same motif multiple times in singing *bouts*. Figure 5.1A shows a spectrogram of a zebra finch song bout in which the motif is repeated multiple times with the notes, syllables, and motifs labeled.

As with the mating vocalizations of many animals, song structure is species specific. The spectrograms in Fig. 5.1 show the acoustic features that distinguish the songs of two closely related species: the zebra finch (*Taeniopygia guttata*) and the long-tailed finch (*Poephila acuticauda*). Whereas zebra finch song is characterized by harmonic and noisy syllables (Fig. 5.1A, B), long-tailed finch song is dominated by syllables with nearly tonal frequency-modulated sweeps (Fig. 5.1C).

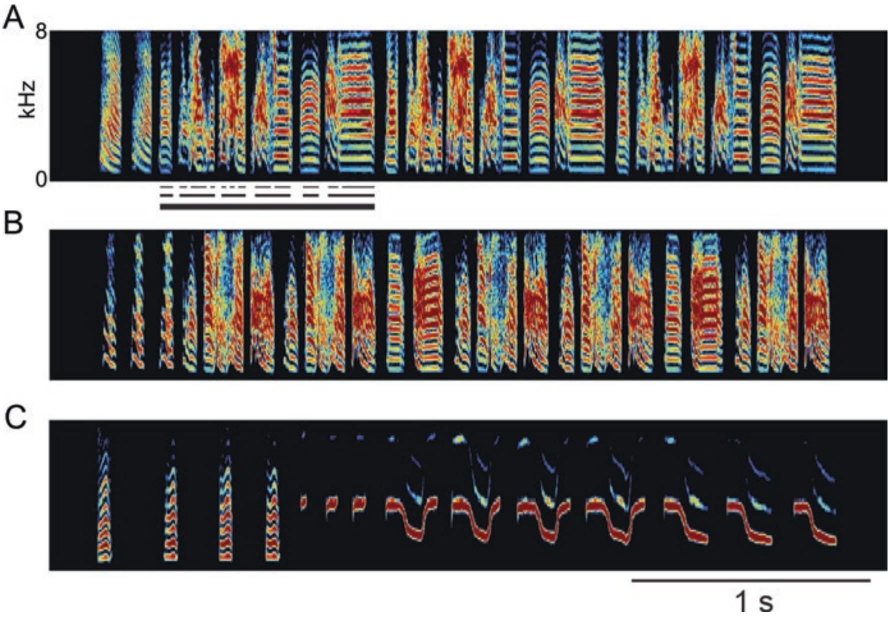


Fig. 5.1 Spectrograms of songs highlight differences in song structure between individuals and species. Song spectrograms from two different zebra finches (A, B; *Taeniopygia guttata*) and a closely related species, the long-tailed finch (C; *Poephila acuticauda*). *Color* indicates intensity: *blue* is low and *red* is high. *Lines below the top spectrogram* label the different components of song including notes (*top*), syllables (*middle*), and a motif (*bottom*)

Unlike the mating vocalizations of most animals, the structure of birdsong is learned, and song structure depends on auditory processing at every life stage (Konishi 2004; Woolley 2008). In addition to the extraction of social information from the environment via song perception, song learning requires auditory memory and feedback of self-generated sounds during song practice (Sakata and Yazaki-Sugiyama, Chap. 2). Songbirds that are deprived of hearing adult song as juveniles or are deafened at some point during development sing highly abnormal songs as adults (Brainard and Doupe 2000; Konishi 2004). Adult maintenance of normal song output also requires auditory feedback as deafened adults gradually lose their songs (Nordeen and Nordeen 1992; Woolley and Rubel 1997). The lifelong reliance on auditory processing for normal singing indicates that understanding song learning and production requires understanding the structure and function of the auditory pallium.

5.3 Functions

Vocal signals contain rich information about the signaler, including information about its species, individual identity, location, and motivational state. Receivers can use the information present in vocal signals to make decisions about social behaviors, including whether to attack or mate. How the auditory system extracts information from vocal signals and uses this information to guide social decision-making is a fundamental question in animal behavior and neuroscience.

5.3.1 Species Recognition

One way in which auditory processing guides behavior is by directing birds, including juveniles, to the songs of their own species. Changes in heart rate (Dooling and Searcy 1980), begging behavior (Nelson and Marler 1993), and movement (Stripling et al. 2003) serve as measures of arousal and indicate that songbirds discriminate between conspecific and heterospecific vocalizations. Comparisons of these measures during the playback of different species' songs suggests that a bird's arousal increases most when exposed to conspecific song. Thus auditory preferences for conspecific song likely guide song learning. For example, young male zebra finches actively worked for playback of conspecific song over other songs during song learning (Adret 1993; Braaten and Reynolds 1999).

Early auditory preferences also guide a bird's selection of song material to copy (Nelson 2000). Birds can learn heterospecific song from interactions with heterospecific adults during development (Immelmann 1969; Woolley and Moore 2011) and from audio presentation of heterospecific song (Baptista and Petrinovich 1984; Petrinovich and Baptista 1987). Moreover, when birds of some species copy heterospecific song, they produce renditions of song that are as accurate as those produced by birds copying their own species song. However, given a choice of templates, juveniles preferentially copy their own species' songs over other songs (Marler 1970; Marler and Peters 1977). This selectivity occurs even when basic hearing sensitivity and song spectra are similar between species (Dooling and Searcy 1980; Okanoya and Dooling 1987). The findings that juveniles can copy heterospecific song but preferentially copy conspecific song indicate that song learning biases are not due to motor constraints. Instead, auditory mechanisms appear sensitive to the acoustic features that distinguish conspecific song from other sounds in the environment.

5.3.2 Individual Recognition and Auditory Memory

Complex social relationships, including those that require repeated interactions among the same individuals, benefit from the ability to remember social partners. Individual recognition reduces aggression, promotes cooperation, and stabilizes long-term social relationships (Tibbetts and Dale 2007). Songbirds interact in a number of behavioral contexts for which there is an advantage to being able to identify individuals, and there has been substantial interest in understanding the role of song in individual recognition in those contexts (e.g., Stripling et al. 2003; Dai et al. 2018). Three of the contexts that have been studied best are territoriality, mate recognition, and tutor song memorization.

5.3.2.1 Territoriality: Recognizing Territory Neighbors

Song is used in territorial interactions in a number of songbird species. In particular, male songbirds often compete for breeding territories and use song to advertise their presence and defend their occupation of a territory (Catchpole and Slater 2008; Bradbury and Vehrencamp 2011). Males in adjacent territories will interact in bouts of singing and counter-singing to establish territory boundaries and, ultimately, a relatively stable social order (Beecher et al. 2000; Catchpole and Slater 2008). Novel males or challengers singing at the edge of a territory will initially provoke an aggressive response, which can include singing, counter-singing, and physical interactions (Brooks and Falls 1975; Catchpole and Slater 2008). However, as the contested boundary is resolved, male aggression decreases such that a song broadcast from a consistent location no longer provokes an attack ("dear enemy effect", Fisher 1954; Temeles 1994). If either the song or the location of the song changes, aggression will be reinstated (Ydenberg et al. 1988; Beecher and Brenowitz 2005). These changes in aggressive behavior imply that territorial males are able to remember and integrate information about the location and identity of other males based on their songs.

5.3.2.2 Mate's Song Recognition

Species that form durable, long-lasting pair bonds, including monogamous species, require perceptual mechanisms for recognizing individuals. Individual recognition of a mate based on acoustic cues has been shown in a range of biparental bird species, including gannets (*Sula bassana*), laughing gulls (*Lams atricilla*), least terns (*Sterna albifrons*), eastern silvereyes (*Zosterops lateralis*), and zebra finches (Beer 1971; White 1971; Miller 1979a; Moseley 1979; Robertson 1996). In zebra finches, males and females form life-long, socially monogamous pairs (Zann 1996), and females show strong preferences for the song of their mate relative to the songs of unfamiliar conspecifics (Clayton 1988; Woolley and Doupe 2008). These data indicate that females, like males, form stable auditory memories of song that can be used to identify individuals (Woolley and Doupe 2008).

5.3.2.3 Tutor Song Memorization

Song experience during development organizes long-term perception: male and female adults remember and show behavioral preferences for songs they encountered as juveniles (Miller 1979a; Riebel 2009). Both males and females memorize the songs of their fathers or tutors during development, and these memories persist into adulthood (Miller 1979b; Clayton 1988). These song memories are then used as acquired templates for sensorimotor learning. While the tutor song memory is important for song development in birds that learn to sing, its significance for females in species in which females do not learn to sing (e.g., zebra finches) is less clear. A female may sexually imprint on her father's song, and this auditory learning can influence attraction to particular song features or to regional dialects, thereby sculpting mate choice decisions (Riebel 2009). Taken together, these data highlight the importance of auditory learning and the ability for both male and female songbirds to memorize the songs of particular individuals for use in social interactions.

5.3.3 Song Preference and Mate Choice

Changes in the performance of particular song features can provide information about the social context or motivational state of the signaler (Sakata and Vehrencamp 2012; Podos and Sung, Chap. 9). For example, male canaries (*Serinus canaria*) increase the number of "sexy syllables" (broadband, two-note syllables produced at a high repetition rate) in their songs when singing to females relative to when singing alone (Fig. 5.2A) (Vallet and Kreutzer 1995), and songs incorporating more of these syllables are preferred by females (Fig. 5.2B, C) (Vallet et al. 1998). Similarly, male zebra finches produce songs that are longer, faster, and more stereotyped when males are courting females compared to when they sing in isolation (Fig. 5.2D, E) (Kao and Brainard 2006; Sossinka and Böhner 1980). Female zebra finches

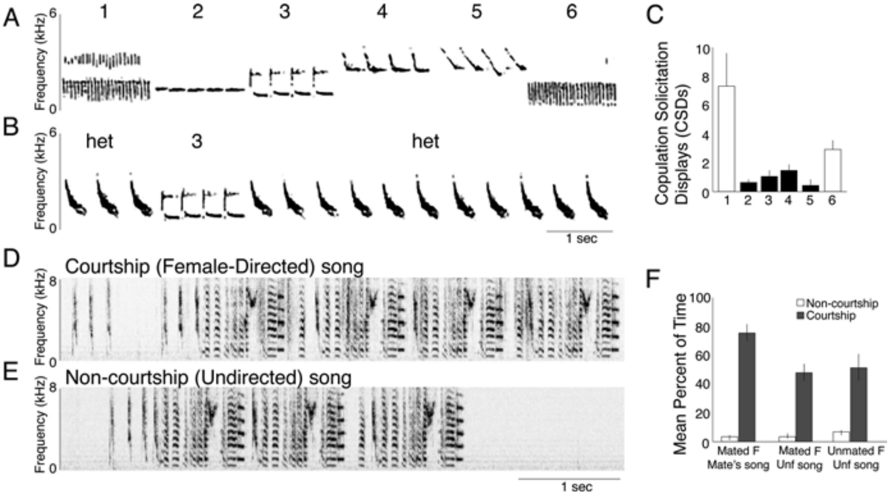


Fig. 5.2 Courtship song preferences in female canaries (*Serinus canaria*) and zebra finches (*Taeniopygia guttata)*. (A) Examples of six different canary song types, including two examples of broadband, rapidly trilled "sexy syllables" (#1 and #6). (B) Vallet and Kreutzer quantified the number of copulation solicitation displays (*CSDs*) in response to canary song types (#3 in this figure) embedded in a greenfinch song (*het*). (C) CSD responses of female canaries to greenfinch songs embedded with each of the six canary song types shown in (A). Female canaries performed significantly more CSDs to greenfinch songs embedded with sexy syllables (#1 and #6; *white bars*) than with other song phrases (*black bars*). (D) In zebra finches, the courtship song contains the same complement of syllables but differs in song performance from the noncourtship song. Courtship songs are longer, faster, and more stereotyped than noncourtship songs (E). (F) Both mated females (*Mated F*) and unmated females (*Unmated F*) prefer the courtship song (*gray bars*) to the noncourtship song (*white bars*). Moreover, females prefer the courtship song even when it is from an unfamiliar singer (*Unf song*). (adapted with permission from Vallet and Kreutzer 1995 and Woolley and Doupe 2008)

generally prefer courtship songs to noncourtship songs, even when the singer is unfamiliar (Fig. 5.2F). Moreover, the strength of the courtship song preference is correlated with the degree of difference in measures of pitch stereotypy or spectral entropy (Woolley and Doupe 2008; Chen et al. 2017). Thus, females attend to and prefer particular vocal characteristics of songs.

5.4 Organization of the Avian Auditory Pallium

Vocal communication is dependent on the ability of receivers to acquire information from acoustic communication signals. The diversity of social tasks for which songbirds use acoustic signals and the evolutionary conservation of auditory circuitry make the songbird an excellent model system for investigating how the auditory system extracts information from vocal sounds to impact social development and communication skills. Although the avian pallium is not laminated like the mammalian neocortex, recent studies of circuit organization, neuron types, gene expression, and physiological response properties demonstrated that there are parallels in the organization and function of the avian pallium and the mammalian neocortex (Karten 2013; Calabrese and Woolley 2015). This detailed knowledge of the neuroanatomy and regional organization provides a critical framework for understanding circuit function as it relates to communication behavior.

5.4.1 Neuroanatomy

The avian auditory pallium, located in the caudal forebrain, contains six major regions organized into contiguous fields of neurons. The regions are heavily interconnected, but they are distinguished by their projections, cell morphology, gene expression, and physiological response properties (Wang et al. 2010; Elliott and Theunissen 2011).

The avian auditory pallium consists of both primary and secondary auditory regions akin to primary and secondary regions of the mammalian auditory cortex. The *primary auditory regions* include Field L (made up of L1, L2a, L2b, and L3) and the caudolateral mesopallium (CLM) (all abbreviations appear in Table 5.1). The two *secondary auditory regions* are the caudomedial mesopallium (CMM) and the caudomedial nidopallium (NCM; see Fig. 5.3). Field L and the adjacent CLM form a layered structure in which the layers correspond to different regions: L1 and CLM are superficial regions, L2a and L2b are intermediate regions, and L3 is the deepest region. Direct input from the auditory thalamus (nucleus ovoidalis) arrives primarily in the intermediate region L2, and different subregions of the auditory thalamus innervate L2a versus L2b (Vates et al. 1996). Neurons in L2 project to the more superficial regions L1 and CLM, to the proximal edge of the deeper region L3, and to a secondary auditory region, the NCM. The superficial CLM connects with

Table 5.1 Abbreviations

AIV	Ventral portion of the intermediate arcopallium
Av	Avalanche
CM	Caudal mesopallium
CLM	Caudolateral mesopallium
CMM	Caudomedial mesopallium
CSD	Copulation solicitation display
HVC	Used as proper name for vocal motor nucleus in the nidopallium
NC	Caudal nidopallium
NCM	Caudomedial nidopallium
NIf	Nucleus interfacialis of the nidopallium
RA	Robust nucleus of the arcopallium
STRF	Spectrotemporal receptive field

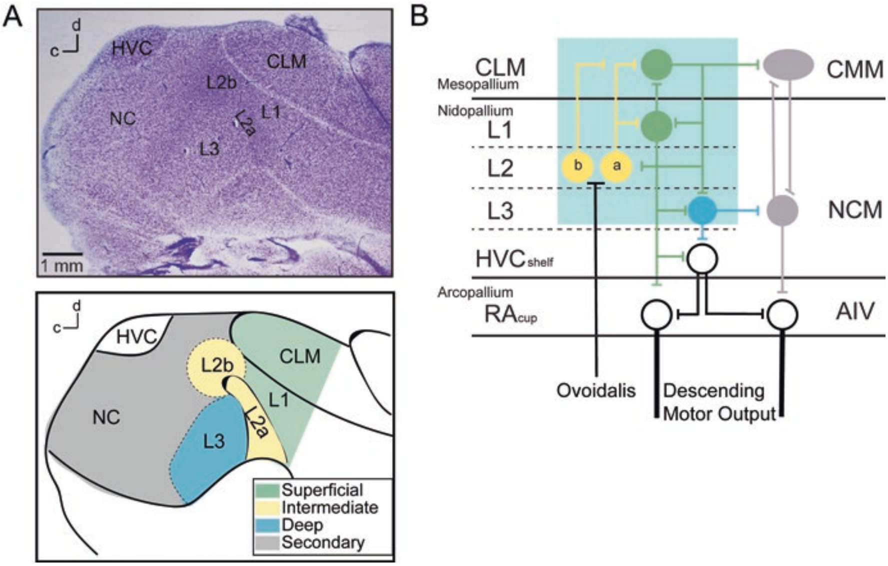


Fig. 5.3 Circuitry of the auditory pallium. (**A**) *Top*: Nissl-stained image of a parasagittal section of the auditory pallium showing cell bodies (*purple stain*) and lamina (*white*). Regions of the primary auditory pallium (*CLM*, *Field L* including *L1*, *L2a*, *L2b*, and *L3*), the secondary auditory area *NC* and the sensorimotor region *HVC* are labeled. *Bottom*: Drawing of the same section, with colors corresponding to the laminar regions of auditory pallium to illustrate the laminar organization. Moving from rostrodorsal to ventrocaudal, *CLM* and *L1* are in the superficial region (*green*); *L2a* and *L2b* are in the intermediate region (*yellow*); *L3* is in the deep region (*blue*); and *NC* is the secondary auditory pallium (*gray*) (*d*, dorsal; *c*, caudal). (**B**) Circuit diagram of the auditory pallium. Colors correspond to those in (A) to indicate superficial (*green*), intermediate (*yellow*), deep (*blue*), and secondary (*gray*) regions. The diagram outlines inputs from the thalamic nucleus *ovoidalis* (*black*); local connections within the auditory pallium, including the primary auditory pallium (*large blue square*) and secondary auditory pallium (*gray*); and outputs to the arcopallium and sensorimotor and motor regions (*black*). *AIV*, ventral portion of the intermediate arcopallium; *CLM*, caudolateral mesopallium; *CMM*, caudomedial mesopallium; *HVC* and *HVCshelf*, used as proper names; *L*, subdivisions of Field L (*L1*, *L2a*, *L2b*, *L3*); *NC*, caudal nidopallium; *NCM*, caudomedial nidopallium; *RAcup*, cup portion of the robust nucleus of the arcopallium (adapted with permission from Calabrese and Woolley, 2015)

multiple auditory regions, including reciprocal connections with each of the subregions of Field L and medial projections to CMM. In addition, the CLM sends projections out to sensorimotor regions important for song production, including HVC (used as a proper name), HVC-shelf, the robust nucleus of the arcopallium (RA) cup, and the nucleus interfacialis of the nidopallium (NIf) (Vates et al. 1996; Bauer et al. 2008). Moreover, within the CLM is a song-selective subregion known as nucleus Avalanche (Av) with bidirectional connections to both HVC and NIf. In contrast, the CMM is heavily interconnected with both the NCM and the CLM with few projections that leave the auditory system (Vates et al. 1996). Finally, the NCM connects most extensively with the CMM and intermediate arcopallium (AIV) and less extensively with other regions of the caudal nidopallium (e.g., caudoventral nidopallium) (Atoji and Wild 2009; Mandelblat-Cerf et al. 2014).

Because these auditory regions lack the thinly laminated structure observed in the mammalian auditory cortex (Fig. 5.3A), there have been various hypotheses regarding the homology of avian and mammalian auditory systems (Karten 1969; Striedter 1997). However, studies of region-specific gene expression (Dugas-Ford et al. 2012), neuron types, and microcircuitry (Wang et al. 2010; Calabrese and Woolley 2015) support potential homologies between regions within the avian auditory pallium and specific layers in the mammalian auditory cortex. In particular, the different regions of the avian auditory pallium appear to be organized in a manner similar to the cortical layers in mammals (Fig. 5.3B). For example, genetic markers that identify thalamo-recipient cortical layer 4 neurons in mammals are expressed in the thalamo-recipient regions of the avian auditory pallium L2a and L2b (Dugas-Ford et al. 2012). Moreover, like the mammalian auditory cortex, the avian auditory pallium is organized into columns with neurons and axons restricted to a column while traversing the multiple regions of the pallium (Wang et al. 2010). Taken together, these data emphasize the impressive similarity between the avian auditory pallium and the mammalian auditory cortex.

5.4.2 Selectivity and Receptive Fields across the Auditory Pallium

The similarities in connectivity within the mammalian sensory cortex and the avian auditory pallium are paralleled by functional similarities. Neurons in the intermediate region have the shortest first spike latencies, and neurons in the secondary region NCM have the longest first spike latencies (Calabrese and Woolley 2015). These latency differences reflect the information processing hierarchy in the pallial circuit and mirror differences in first spike latencies across mammalian cortical layers (Atencio et al. 2009). As in the mammalian cortical circuit, response selectivity and the sparseness of population responses increase at each processing stage of the songbird auditory pallium (Fig. 5.4) (Schneider and Woolley 2013; Calabrese and Woolley 2015). Related to response selectivity, the receptive fields of individual neurons progressively increase in complexity along the hierarchy (Moore and Woolley 2019). Connectivity between putative excitatory and putative inhibitory neurons also differs by region: the connectivity patterns in intermediate, superficial, and deep regions of the songbird auditory pallium (Calabrese and Woolley 2015) map onto connectivity patterns between the same cell types in intermediate, superficial, and deep layers of the mammalian cortex (Hansen et al. 2012; Harris and Mrsic-Flogel 2013). Thus, comparable information-coding strategies of single neurons and neuronal populations in avian pallial regions and mammalian cortical layers suggest that birds and mammals have parallel, possibly homologous, auditory processing circuits.

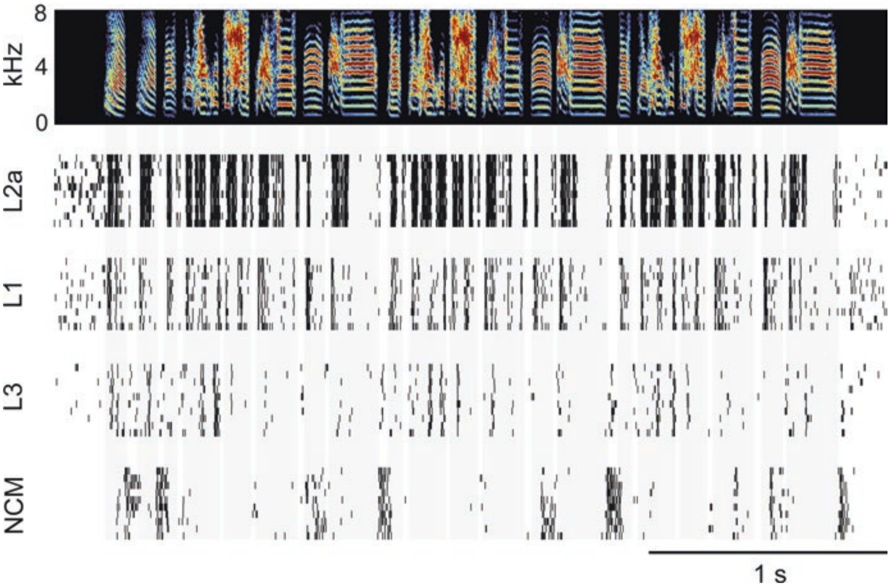


Fig. 5.4 Auditory responses across the auditory pallium. Raster plots of the responses of single neurons to the same song (*spectrogram shown, top*). Raster plots are organized from thalamo-recipient (*L2a*) to secondary (*NCM*) regions. Moving from L2a to NCM, the activity becomes more sparse and selective (i.e., responds to a smaller range of sounds in the song). Each *tick mark* represents a spike of a neuron, and each *row of the raster plot* summarizes the response of a neuron to a single presentation of the sound

5.4.2.1 Primary Auditory Pallium

The regions of the primary auditory pallium, including Field L and CLM, are tonotopically organized. Studies using pure tone stimuli have found that Field L (Zaretsky and Konishi 1976; Heil and Scheich 1991) and CLM display regions of isofrequency contours (Müller and Scheich 1985; Müller and Leppelsack 1985) and have identified multiple subcenters or tonotopic gradients within those areas. In addition, within each region there appears to be a mediolateral gradient of spectral tuning (Kim and Doupe 2011).

Beyond the simple tonotopic organization, studies using more complex sounds and reverse correlation techniques found a basic set of spectrotemporal receptive fields (STRFs), which depict the acoustic features that drive a neuron to fire. Mapping STRFs revealed region-dependent variation in spectral and temporal tuning (Hose et al. 1987; Nagel and Doupe 2008). In particular, the thalamo-recipient region L2 contains neurons with the simplest receptive fields (Nagel et al. 2011; Kim and Doupe 2011), while both deep and superficial regions have more complex receptive fields (L1, L3, CLM).

Regional differences in the firing rate and selectivity of neurons correlate with the differences in receptive fields (Nagel and Doupe 2008; Calabrese and Woolley

2015). In general, single neurons in Field L and CLM fire in response to the presentation of most conspecific songs, and spikes are reliably time-locked to specific acoustic features in a sound (Fig. 5.4). One neuron will produce distinct spike trains in response to acoustically different songs because the acoustic features that match the neuron's receptive field occur at different points in each song. Because receptive fields differ, the same sound evokes different responses from each neuron. In addition, receptive field complexity determines a neuron's response selectivity.

Song selectivity is often measured as the fraction of presented songs that do not evoke a response from a given neuron (Schneider and Woolley 2013; Calabrese and Woolley 2015). On average, neurons in the intermediate region L2 are significantly less selective (i.e., a smaller fraction of sounds do not evoke a response) than those at successive processing stages (Meliza and Margoliash 2012; Calabrese and Woolley 2015). Superficial-region and deep-region neurons (those in L1, L3, and CLM) produce more selective song responses with lower firing rates than do L2 neurons (Calabrese and Woolley 2015; Moore and Woolley 2019). Finally, while neurons throughout the primary auditory pallium may respond strongly to tones and modulated noise, they respond more robustly to song than to other sounds (Theunissen et al. 2004). Together, these region-specific differences indicate that tuning complexity and response selectivity increase along the primary pallial pathway.

5.4.2.2 Secondary Auditory Pallium

Neurons in the secondary regions, CMM and NCM, have more complex response properties than neurons in primary auditory areas. In particular, unlike Field L neurons, neurons in the NCM show wider and more multipeaked tuning. They are also driven less strongly by tones or noise-like sounds; instead, neurons in the NCM respond to more complex auditory features and exhibit pronounced selectivity for particular songs (Schneider and Woolley 2013; Yanagihara and Yazaki-Sugiyama 2016). For example, in European starlings (*Sturnus vulgaris*), CMM neurons responded selectively to learned auditory objects versus unlearned auditory objects (Gentner and Margoliash 2003; Jeanne et al. 2011). Similarly, in zebra finches, NCM neurons preferentially responded to the tutor song and/or the bird's own song following sensory learning during development (Phan et al. 2006; Yanagihara and Yazaki-Sugiyama 2016).

In secondary sensory areas of both mammals and birds, the classical STRFs failed to accurately describe observed responses to natural stimuli (Theunissen et al. 2000; Machens et al. 2004). For example, in songbird secondary auditory regions, STRF models can explain, at most, 30% of a neuron's response to a stimulus (Sen et al. 2001). These data highlight the challenges inherent in modeling responses to sensory stimuli in regions beyond primary auditory pallium. In particular, standard linear models do not capture nonlinear tuning properties and, therefore, do not yield accurate receptive field estimates for upstream neurons. For example, approaches that incorporated information about the probability of sounds, rather than just the spike-evoking acoustic features, improved model predictions of neural responses

(Gill et al. 2008; Lu and Vicario 2017). These and other alternative approaches to strictly linear models will provide novel paths forward for measuring the receptive fields of neurons in these regions (see Sect. 5.8.4) (Gill et al. 2008; Lu and Vicario 2017).

Indeed, there are several factors beyond the selective responses to auditory objects or features that strongly modulate the activity of neurons in secondary auditory regions: stimulus history covering multiple syllables (Schneider and Woolley 2013; Lu and Vicario 2017), the acoustic environment (Terleph et al. 2008; Yang and Vicario 2015), and the behavioral salience of songs (Gobes et al. 2010). For example, repeated playback of a single song led to adaptation of both the electrophysiological (Phan et al. 2006) and immediate early gene responses (Mello et al. 1995) in the NCM (reviewed in Dong and Clayton 2009). However, the acoustic context in which the repetitions occur can affect the response (Kruse 2004; Lu and Vicario 2017). For example, if a repeated stimulus is played in a novel or unexpected context (e.g., playback of a familiar zebra finch song is unexpectedly embedded within a series of canary songs), responses to the familiar song can be enhanced (Lu and Vicario 2017). This enhanced response rapidly returns to the adapted rate when the stimulus is again played in a familiar context (Lu and Vicario 2017). Thus, NCM neuron responses not only provide a read-out of the auditory properties of a stimulus but also encode the probability of sounds or sound transitions. NCM neurons may even generalize probabilities across categories, predicting not only the probability of one auditory object based on its repetition but also the expectation for an entire class of sounds (Lu and Vicario 2017).

While the challenges in characterizing the receptive fields of secondary auditory neurons have led to an incomplete description of the tuning in these regions, there does appear to be variation in the tuning and response properties of neurons across the secondary auditory pallium. For example, there appeared to be topographic differences within the NCM in the degree to which neurons habituate in response to repeated stimuli: dorsal and caudal regions of the nucleus showed greater habituation than rostral or ventral regions (Chew et al. 1995; Mello et al. 1995). Taken together, these data hint at the potential for topographic compartmentalization of function that would help to explain the ability of secondary auditory neurons to perform somewhat disparate tasks, for example, the invariant coding required for auditory scene processing (Sect. 5.5) versus the rapid, stimulus-specific habituation associated with auditory memory (Sect. 5.6).

These findings highlight that, like the mammalian auditory cortex, the avian auditory pallium is organized into a discrete hierarchy of interconnected areas. As one moves from primary to secondary regions, which then project to motor and sensorimotor regions, firing becomes sparser and more selective, and linear models of receptive fields become poorer at estimating actual responses. The hierarchical transformation of song coding in the songbird auditory pallium is similar to transformations in sensory representations in other systems (Graham and Field 2007; Harris and Mrsic-Flogel 2013). In addition, auditory neurons higher up the hierarchy differentially respond to the acoustic context in which sounds are embedded. These changes in firing, receptive fields, and selectivity are functionally significant.

As discussed in the upcoming sections, the emergence of sparse-firing neurons with greater selectivity contributes to a number of important abilities and behaviors, including processing complex auditory scenes (Sect. 5.5), memory formation and individual recognition (Sect. 5.6), and song preference and mate selection (Sect. 5.7). Moreover, these varied functions may themselves be important in elucidating the topographic organization of the secondary auditory pallium (Sect. 5.8.1).

5.5 Invariant Coding Pulls Signals out of Noise

The ability to attend to target sounds, such as a communication signal, in a complex acoustic background is critical for receivers (Bregman 1994; Bee and Micheyl 2008). In songbirds, individuals are able to identify particular songs occurring within complex acoustic scenes such as noise (Dent et al. 2009) and song choruses (Schneider and Woolley 2013). Solving this “cocktail party problem” may depend on the differences in neural firing between neurons in Field L and neurons in secondary auditory regions, in particular, the emergent sparse coding of sounds by NCM neurons (Moore et al. 2013; Schneider and Woolley 2013).

As described in Sect. 5.4, the auditory coding of complex sounds like birdsong dramatically transforms between the thalamo-recipient and higher pallial regions (Nagel and Doupe 2008; Woolley et al. 2009). Early in the cortical processing pathway, neurons respond nonselectively (i.e., each neuron responds to a high proportion of songs) and with many spikes throughout the stimulus because their receptive fields are linear and driven by simple acoustic features found in many complex sounds (Nagel and Doupe 2008; Woolley et al. 2009). This coding scheme results in a dense and redundant neural representation of a song or a chorus of multiple birds’ songs. However, the coding of songs transforms between L2 and subsequent regions where firing is more selective and sparse because these neurons have nonlinear receptive fields, which display responses that vary depending on a variety of factors, including recent history (Sect. 5.4.2). At the highest levels of the auditory pallium, single neuron responses are selective and characterized by few, highly reliable spikes in response to a song (Fig. 5.4). Because responses are so sparse, each neuron produces a highly distinct response pattern to each song, if it responds to the song at all. Higher cortical regions, therefore, represent songs in a sparse spiking code distributed across multiple neurons.

Selective and sparse neural coding may facilitate the coding of target sounds, for example, individual songs in complex scenes (song choruses). The coding of songs within complex scenes requires neurons to fire consistently over multiple presentations. Importantly, as discussed previously, there appears to be some topography within the NCM in how neurons respond to multiple presentations of the same song. In particular, while neurons in the dorsal and caudal NCM habituate following repeated playback of a song, sparse-coding neurons in the rostral NCM produce highly precise song responses; the temporal patterns of their responses are almost identical over multiple presentations of the same song (Schneider and Woolley 2013). Selective, sparse, and precise coding may facilitate the recognition of individual songs because selectivity is inversely correlated with the strength of responses to background sounds. For this reason, sparse-coding NCM neurons have been studied as potentially providing a neural mechanism to solve the cocktail party problem (Moore et al. 2013; Schneider and Woolley 2013).

Specifically, sparse-coding NCM neurons produce very similar responses to one song presented alone and to that same song presented in combination with background sounds (e.g., chorus, songs, noise). These responses are referred to as *background invariant* and have the potential to accurately represent a target vocalization in an acoustic scene composed of vocalizations from many others. Schneider and Woolley (2013) tested the relationship between NCM neuron responses and behavioral recognition of target songs that were presented with varying levels of background choruses. The signal (song) to noise (chorus) ratios of acoustic scenes were varied while birds completed song recognition tasks, which revealed the signal-to-noise ratios that permitted correct identification of target songs in those acoustic scenes. In the same birds, sparse-coding neurons in the rostral NCM produced the same sparse responses to songs alone and to those songs embedded in acoustic scenes as if the background choruses were absent. While consistent firing patterns were observed at signal-to-noise ratios that permitted behavioral recognition, these same neurons stopped firing at signal-to-noise ratios that were too low for behavioral identification of a target song. Those results demonstrated that the responses of sparse-coding NCM neurons parallel perceptual recognition of target songs in acoustic scenes, providing a potential neural solution to the cocktail party problem.

5.6 Secondary Auditory Pallium as a Potential Substrate for Song Memory

Given the importance of memorizing song for vocal learning and individual recognition, songbirds offer a compelling model for understanding the encoding of auditory memory. Indeed, one of the greatest challenges in songbird research has been to identify the neural site (s) in which the tutor song template is stored. The secondary auditory area NCM has been of particular interest. Early studies argued that the habituation of neural activity in NCM to repeated presentations of the same song was indicative of a song memory trace in the NCM (Chew et al. 1995; Phan et al. 2006). The habituation of auditory responses is specific to song, as there is no habituation to tones, implying that the changes in activity are not a consequence of general adaptation of the auditory system to repeated stimuli, but the changes in activity could be related to the encoding of a song memory. In the following sections, the data supporting the role of the NCM in auditory memory are described and some remaining questions regarding the contribution of NCM to auditory memory formation are considered.

5.6.1 Song Memory Formation in Adulthood

Unlike the noise invariant responses of neurons in the rostral portion of the NCM, the largest response of dorsal and caudal NCM neurons to an individual song occurs with the first playback, and responses to the same song decrease over repeated playbacks (Chew et al. 1995; Mello et al. 1995). The degree to which there is habituation of firing or immediate early gene expression over presentations of the same song depends on both the number of consecutive playbacks the bird originally experienced as well as the duration of time between playbacks (Mello et al. 1995). For example, responses to a song are decreased only slightly following ten consecutive playbacks, but responses are almost completely abolished following 200 consecutive playbacks (Kruse et al. 2000).

The habituation of the neural response in both immediate early gene expression and electrophysiology was proposed to represent a memory trace (Chew et al. 1995; Phan et al. 2006). For example, male zebra finches exhibit song recognition learning after passive song playback (Stripling et al. 2003; Dai et al. 2018). The time course of song recognition learning parallels the time course of changes in neural activity in response to repeated song playback (Mello et al. 1995; Stripling et al. 2003). On both the behavioral and neural levels, memory lasts at least a day and, in some cases, can be long lasting (Miller 1979a,b). For example, female zebra finches show strong preferences for the song of their mate even after weeks of separation from their mate (Woolley and Doupe 2008), and such lasting preferences for familiar song require an enduring memory trace. Consistent with a role for the NCM in long-term song memory, immediate early gene expression in the NCM of females is lower in response to hearing their mate's song than in response to hearing the songs of unfamiliar males up to several weeks after separation from mates (Woolley and Doupe 2008).

5.6.2 Memory of Tutor Song

One of the longest lasting auditory memories in songbirds is that of the tutor song (see also Sakata and Woolley, Chap. 1). Emberezine sparrows, such as swamp sparrows (*Melospiza georgiana*), provide a particularly striking example of the endurance of tutor song memory. Juvenile swamp sparrows memorize the song of their tutor in the late summer or fall; however, they only begin practicing to produce those songs in the following spring, months after they were exposed to their tutor song (Marler and Peters 1981). This indicates that there must be an enduring trace of the tutor's song that allows for accurate song imitation in these birds.

Activity in the auditory forebrain has been implicated in both the formation of tutor song memory and the adult recall of tutor song. Both adult male and female zebra finches prefer the tutor song over unfamiliar songs, and lesions of the NCM significantly reduced the strength of the tutor song preference (Gobes and Bolhuis 2007). Moreover, in adult male zebra finches, the fraction of tutor song that is copied is correlated with immediate early gene expression in the NCM in response to the tutor song (Bolhuis and Gahr 2006). Tutor song playback also differentially increased the expression of immediate early genes, such as EGR1, in the CMM of adult female zebra finches (Terpstra et al. 2006). This differential response to tutor song was also observed in juvenile zebra finches: EGR1 responses in both the CMM and NCM were greater for tutor song than for novel song (Bolhuis and Gahr 2006; Gobes et al. 2009). Taken together, these data support a potential role of the NCM and CMM in storing tutor song memory.

Further evidence for a role of the auditory forebrain in tutor song memory comes from experiments manipulating the molecular pathways that regulate the expression of EGR1 (see London, Chap. 8). Specifically, the gene product ERK is part of a molecular pathway critical for memory formation that lies upstream to EGR1 (London and Clayton 2008). In a series of elegant experiments, London and Clayton demonstrated that blocking the ERK pathway during developmental song tutoring leads to poor imitation of the tutor song. The effect does not appear to be a consequence of the disruption of hearing or sensorimotor practice; the effect specifically results from interfering with song memorization. While the infusion of the ERK inhibitor affected EGR1 induction in both the NCM and CMM and, thus, prevented the specific attribution to NCM or CMM, these data provide compelling support for the role of secondary auditory regions in tutor song memory formation.

Taken together, these studies indicate that activity in the NCM and CMM often parallels behavioral measures of learning and memory. However, detailed understanding of the coding properties of these regions remains incomplete, in part, because of variation in the approaches used. For example, analyses of immediate early gene expression have been pivotal in establishing that there is molecular habituation and the extent of this habituation can vary across auditory regions (NCM versus CMM). However, the limited range of stimuli used in immediate early gene studies and the absence of expression of immediate early genes, like EGR1, in primary auditory pallium has hindered the use of these methods in providing a more complete understanding of song memory formation (but see Horita et al. 2010; Horita et al. 2012).

Lesion and manipulation studies have been significant in demonstrating the importance of the auditory forebrain for particular memory tasks, but the ability to discretely affect a single neural locus remains a challenge as does controlling for manipulations that affect sound processing versus memory. Finally, while electrophysiological approaches enable comparisons within a single neuron across a broader array of stimuli and provide needed insight into how auditory memories are encoded, more studies that couple neurophysiological recordings with behavior are necessary to better understand memory coding.

5.7 Neural Mechanisms of Song Preference and Mate Selection

Across a diversity of songbird species, male song serves to attract females (Andersson 1994; Catchpole and Slater 2008). Both field and laboratory studies have found that song can lead females to approach a male or a speaker (Eriksson and Wallin 1986; Woolley and Doupe 2008). Similar studies have found that females will call back in response to hearing songs (Dunning et al. 2014; Chen et al. 2017) and will perform operant tasks (e.g., perch hopping, string pulling) to hear playback of song (Riebel 2009; Schubloom and Woolley 2016). Females show preferences for particular song categories: preference for conspecific over heterospecific songs (Searcy and Brenowitz 1988; Riebel 2009) and for courtship over noncourtship songs (Vallet and Kreutzer 1995; Woolley and Doupe 2008). Such categorical preferences are generally shared across females and are often correlated with particular song features. For example, female zebra finches prefer songs with less variability in pitch across syllable renditions and less within-syllable spectral entropy (Woolley and Doupe 2008; Chen et al. 2017). Thus, behavioral responses to song have been widely used to assess female song preferences that can ultimately affect female mate choice (Riebel 2009).

One approach to study the neural basis of song preference has been to use neural tuning to uncover song features that influence female preferences for song. To this end, a number of studies have measured behavioral responses to songs that differ in a particular feature space and then played those songs back to assess whether particular regions of the auditory forebrain showed differential expression of immediate early genes in response to songs that differ in particular features (Leitner et al. 2005; Woolley and Doupe 2008). For example, EGR1 expression in CMM is increased in response to salient or preferred songs, including courtship song in zebra finches (Woolley and Doupe 2008; Chen et al. 2017), and EGR1 expression in CMM increased in response to songs with sexy-syllables in canaries (Leitner et al. 2005). Both studies raised the possibility that the CMM is involved in discriminating song quality or salience.

However, one challenge has been deciphering whether differential neural responses reflect differences in preferences or are simply a result of differences in acoustic features between preferred and unpreferred stimuli. In the case of the CMM, additional studies have found instances in which the expression of EGR1 in the CMM was uncoupled from behavioral preference for calls (Gobes et al. 2009) and for some songs (Chen et al. 2017; Van Ruijssevelt et al. 2018). For example, unlike their normally reared counterparts, females reared without developmental song exposure (song naïve) do not consistently prefer courtship song over noncourtship song. However, similar to normally reared females, EGR1 in the CMM also increased in song-naïve females in response to courtship song compared to noncourtship song (Chen et al. 2017).

In another study, there was not only a disconnect between the behavioral preferences and neural responses in CMM, but the nature of the neural response provided insight into the features that may be attended to by CMM. In particular, Van Ruijssevelt et al. (2018) measured neural (BOLD fMRI) and behavioral responses to courtship and noncourtship song and to stimuli that manipulated temporal and spectral features of song. The manipulated stimuli contained the characteristic temporal features of courtship song, and BOLD responses in the CMM clustered stimuli on the basis of those temporal acoustic features; behaviorally, the birds differentiated between the manipulated stimuli and the unmanipulated courtship song (Van Ruijssevelt et al. 2018). The temporal structure of song also affected EGR1 expression in the CMM (Lampen et al. 2014), and temporal cues were more important than spectral cues in single-unit auditory responses in a target of CMM, the sensorimotor nucleus HVC (Theunissen and Doupe 1998). Together, these data raise the possibility that the CMM and its targets are biased toward temporal information (Woolley and Rubel 1999; Woolley et al. 2005).

Activity in the NCM also is correlated with song preferences. For example, EGR1 expression in the NCM, but not the CMM, of female starlings was higher following playback of long songs, which females generally prefer, versus short songs (Gentner et al. 2001). Similarly, EGR1 expression also was higher in female zebra finches in response to the preferred courtship song compared to the less preferred noncourtship song (Chen et al. 2017); however, this difference is modulated by familiarity. Whereas EGR1 differences in the NCM were observed between unfamiliar courtship and noncourtship songs (Chen et al. 2017), EGR1 expression in the NCM did not differ between familiar courtship and noncourtship songs (Woolley and Doupe 2008).

Finally, both behavioral preferences and neural responses are shaped by social and acoustic experience during development. For example, female song sparrows prefered the dialect in which they were reared over the dialect of their genetic parents (Hernandez and MacDougall-Shackleton 2004), and female zebra finches prefered the songs of the subspecies with whom they were reared over their genetic parents (Clayton 1990). Moreover, females reared without developmental song exposure showed atypical song preferences as adults. Unlike normally reared females, song-naïve female zebra finches preferred the songs of isolate males (who lack multiple acoustic features of learned song) and had significantly fewer dendritic spines per unit length in the NCM compared to normally reared females (Lauay et al. 2004). Similarly, song-naïve females showed aberrant song preferences and no difference in EGR1 expression in the NCM for courtship versus noncourtship songs (Chen et al. 2017). Whereas electrophysiological studies have indicated that the early acoustic environment had subtle but significant effects on the responses of neurons in Field L of females (Hauber et al. 2013), characterization of the degree to which the responses of neurons in the secondary auditory regions are shaped by developmental song exposure will provide needed insight into the mechanisms by which social and auditory experience shape song preference.

Thus, female songbirds show preferences for particular acoustic features and auditory objects. These preferences go beyond just the ability to discriminate sounds and can be shaped by auditory and social experiences both during development and in adulthood. Activity in both the CMM and NCM has been associated with different aspects of song preferences, though further work is necessary to better delineate the circuitry involved in song preference decisions and mate choice.

5.8 Future Directions

Neuroanatomical tracing (Wang et al. 2010), gene expression (Dugas-Ford et al. 2012), and targeted electrophysiological recordings (Kim and Doupe 2011; Calabrese and Woolley 2015) have resulted in a detailed understanding of the connectivity and coding properties of the avian auditory system and have facilitated comparisons to the mammalian auditory cortex. While these approaches have led to greater recognition and appreciation of the similarities in structure and function of auditory systems across species as well as better depictions of general principals of auditory coding, much remains to be discovered about the organization and function of songbird auditory circuits.

5.8.1 Mapping of Secondary Auditory Areas

Electrophysiological mapping of Field L and CLM have revealed the detailed structure of spectral and temporal response properties both within and between regions; however, adopting similar approaches in the secondary regions, including the NCM and CMM, has been more challenging. Responses of NCM and CMM neurons are sparse, selective, plastic, and highly nonlinear; thus, activity in these regions is poorly characterized by linear models such as STRFs (Meliza and Margoliash 2012; Schneider and Woolley 2013). In starlings, there are facilitative and suppressive interactions between song notes on the spiking responses of neurons in the CMM, thereby making it difficult to predict CMM neuron responses to songs by analyzing responses to single notes presented individually (Meliza and Margoliash 2012). Similarly, EGR1 responses in the secondary auditory pallium to whole canary song were not re-created in the responses to individual components that make up the song (Ribeiro et al. 1998). While neurons in these regions have a propensity to show parallels between neural and behavioral responses to stimuli, the way in which they encode information remains elusive. For example, many of the brain-behavior correlations rely on comparing responses to pairs or small numbers of behaviorally relevant stimuli (e.g., conspecific versus heterospecific, familiar versus unfailiar, courtship versus noncourtship). Studies employing broader stimulus sets will be critical to gain insight into more general principles of NCM and CMM neural responses and, ultimately, into how characteristics of vocal signals are processed and used to guide behavior.

5.8.2 *Catecholamines in the Auditory Cortex Shape Behavior*

Forebrain auditory areas receive neuromodulatory inputs that can affect their activity and plasticity (see Remage-Healey, Chap. 6). Inputs from catecholamines, in particular, may provide a mechanism for translating auditory experience into changes in brain and behavior. Dopaminergic and noradrenergic neurons in the midbrain and hindbrain in songbirds respond to salient or preferred stimuli, indicating to the brain which stimuli are important (Fields et al. 2007; Sara and Bouret 2012). For example, cFOS expression in dopaminergic neurons in the caudal ventral tegmental area of female songbirds was higher following playback of the preferred courtship song than the noncourtship song (Barr and Woolley 2018). Similarly, in juvenile male zebra finches, noradrenergic neurons and dopaminergic neurons expressed more EGR1 in response to tutoring methods that lead to more robust vocal learning (Chen et al. 2016).

Pairing stimulation of midbrain dopaminergic neurons or hindbrain noradrenergic neurons with playback of a tone drove plastic changes to the tonotopic representation of sounds in the mammalian primary auditory cortex (Bao et al. 2001; Martins and Froemke 2015). There is potential for a similar role of catecholamines in shaping neural responses in the avian auditory pallium and behavior. For example, the NCM receives substantial catecholaminergic projections (Van Ruijssevelt et al. 2018), and NCM responses to song can be modulated by norepinephrine in zebra finches (e.g., Velho et al. 2012; Ikeda et al. 2015). Further, decreasing norepinephrine levels can attenuate the rate of auditory learning and discrimination (Velho et al. 2012), reduce copulation solicitation displays to sexually stimulating songs, and reduce responses of forebrain auditory regions to conspecific song (Appeltants et al. 2002; Lynch and Ball 2008). However, while catecholaminergic inputs are well-positioned to modulate plasticity and firing of forebrain auditory regions, little is known about the mechanisms of these effects in songbirds. Studies of how dopamine and norepinephrine affect the response properties of different cell types in the NCM and CMM will be critical for understanding how these neuromodulators contribute to auditory tasks by shaping auditory preferences for or altering cortical representation of an auditory stimulus.

5.8.3 *Development*

Hearing species-typical song during development critically shapes the auditory system of both male and female songbirds. Birds form long-lasting memories of those song exemplars and use them in learning to produce their own songs or in guiding social decisions. Hearing song during development also appears to shape auditory responses to song. For example, the song-evoked firing rates of Field L neurons were significantly higher in zebra finches reared and tutored by conspecific adults than in zebra finches reared and tutored by Bengalese finches (Woolley et al. 2010), and Field L neurons were more selective for conspecific song over simple sounds (such as tone pips) in normally reared zebra finches compared to zebra finches reared in white noise (Amin et al. 2013). While there has been reasonable demonstration that not hearing song during development alters some neural response properties and behavior (Woolley 2012), future work will have to uncover how developmental song exposure influences tuning properties or neural selectivity across the auditory cortex.

5.8.4 Quantifying Nonlinear Responses of Neurons in Secondary Auditory Areas

As discussed previously (Sect. 5.4.2.2; Sect. 5.8.1) the responses of many pallial neurons, particularly those in secondary areas, are poorly characterized by strictly linear models of stimulus-response relationships such as the simplest STRF model. The use of linear-nonlinear models of sensory tuning can improve predictions of auditory responses to complex sounds (e.g., songs) (Calabrese et al. 2011). These models combine the linear filter (STRF) with nonlinear functions designed to capture neural response properties, such as spike threshold and dependence on recent spike history, to predict responses to complex sounds more accurately. While the linear-nonlinear model represents an improvement over the linear model alone, quantifying neural tuning that explains the nonlinearities in responses to natural sounds, including song, will require far more sophisticated models. For example, models that include synaptic depression have accounted for nonlinear modulations in tuning in the mammalian auditory cortex during natural sound processing (David et al. 2009). New approaches that factor in behavioral state and auditory learning will be particularly important for progress in understanding how the auditory pallium encodes and decodes song.

5.9 Conclusions

Songbirds use learned songs to convey diverse information about an individual's species, sex, identity, motivation, or social context. Studies of behavior highlight the abilities of receivers to extract information from song and use this information for song learning, territorial interactions, individual recognition, and mate choice. Investigations of the neural mechanisms underlying these auditory abilities have provided substantial information about the organization, structure, and response properties of neurons in the primary auditory pallium and the potential role of those neurons in shaping behavior. Future research focused on gaining further insight into the roles of developmental experience and neuromodulators and generating improved methods to describe and understand the nonlinear response properties of secondary auditory areas will provide needed insights into the neural basis of auditory learning, coding, memory, and perception.

Compliance with Ethics Requirements Sarah C. Woolley declares that she has no conflict of interest. Sarah M. N. Woolley declares that she has no conflict of interest.

References

Adret P (1993) Operant conditioning, song learning and imprinting to taped song in the zebra finch. Anim Behav 46:149–159. https://doi.org/10.1006/anbe.1993.1170

Amin N, Gastpar M, Theunissen FE (2013) Selective and efficient neural coding of communication signals depends on early acoustic and social environment. PLoS One 8:e61417. https://doi.org/10.1371/journal.pone.0061417

Andersson MB (1994) Sexual selection. Princeton University Press

Appeltants D, Del Negro C, Balthazart J (2002) Noradrenergic control of auditory information processing in female canaries. Behav Brain Res 133:221–235. https://doi.org/10.1016/S0166-4328(02)00005-0

Atencio CA, Sharpee TO, Schreiner CE (2009) Hierarchical computation in the canonical auditory cortical circuit. Proc Natl Acad Sci pnas.0908383106. doi: https://doi.org/10.1073/pnas.0908383106

Atoji Y, Wild JM (2009) Afferent and efferent projections of the central caudal nidopallium in the pigeon (Columba livia). J Comp Neurol 517:350–370. https://doi.org/10.1002/cne.22146

Bao S, Chan VT, Merzenich MM (2001) Cortical remodelling induced by activity of ventral tegmental dopamine neurons. Nature 412:79. https://doi.org/10.1038/35083586

Baptista LF, Petrinovich L (1984) Social interaction, sensitive phases and the song template hypothesis in the white-crowned sparrow. Anim Behav 32:172–181. https://doi.org/10.1016/S0003-3472(84)80335-8

Barr HJ, Woolley SC (2018) Developmental auditory exposure shapes responses of catecholaminergic neurons to socially-modulated song. Sci Rep 8:11717. https://doi.org/10.1038/s41598-018-30039-y

Bauer EE, Coleman MJ, Roberts TF, Roy A, Prather JF, Mooney R (2008) A synaptic basis for auditory-vocal integration in the songbird. J Neurosci 28:1509–1522. https://doi.org/10.1523/JNEUROSCI.3838-07.2008

Bee MA, Micheyl C (2008) The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? J Comp Psychol 122:235–251. https://doi.org/10.1037/0735-7036.122.3.235

Beecher MD, Brenowitz EA (2005) Functional aspects of song learning in songbirds. Trends Ecol Evol 20:143–149. https://doi.org/10.1016/j.tree.2005.01.004

Beecher MD, Campbell SE, Burt JM, Hill CE, Nordby JC (2000) Song-type matching between neighbouring song sparrows. Anim Behav 59:21–27. https://doi.org/10.1006/anbe.1999.1276

Beer CG (1971) Individual recognition of voice in the social behavior of birds. In: Lehrman DS, Hinde RA, Shaw E (eds) Advances in the study of behavior. Academic, pp 27–74

Bolhuis JJ, Gahr M (2006) Neural mechanisms of birdsong memory. Nat Rev Neurosci 7:347–357. https://doi.org/10.1038/nrn1904

Braaten RF, Reynolds K (1999) Auditory preference for conspecific song in isolation-reared zebra finches. Anim Behav 58:105–111. https://doi.org/10.1006/anbe.1999.1134

Bradbury JW, Vehrencamp SL (2011) Principles of animal communication. 2nd. Sunderland, Massachusetts: Sinauer

Brainard MS, Doupe AJ (2000) Auditory feedback in learning and maintenance of vocal behaviour. Nat Rev Neurosci 1:31–40. https://doi.org/10.1038/35036205

Bregman AS (1994) Auditory scene analysis: the perceptual Organization of Sound. MIT Press

Brooks RJ, Falls JB (1975) Individual recognition by song in white-throated sparrows. I. Discrimination of songs of neighbors and strangers. Can J Zool 53:879–888. https://doi.org/10.1139/z75-101

Calabrese A, Schumacher JW, Schneider DM, Paninski L, Woolley SMN (2011) A generalized linear model for estimating Spectrotemporal receptive Fields from responses to natural sounds. PLoS One 6:e16104. https://doi.org/10.1371/journal.pone.0016104

Calabrese A, Woolley SMN (2015) Coding principles of the canonical cortical microcircuit in the avian brain. Proc Natl Acad Sci 112:3517–3522. https://doi.org/10.1073/pnas.1408545112

Catchpole CK, Slater PJB (2008) Bird song: biological themes and variations. Cambridge University Press

Chen Y, Clark O, Woolley SC (2017) Courtship song preferences in female zebra finches are shaped by developmental auditory experience. Proc R Soc B Biol Sci 284:20170054. https://doi.org/10.1098/rspb.2017.0054

Chen Y, Matheson LE, Sakata JT (2016) Mechanisms underlying the social enhancement of vocal learning in songbirds. Proc Natl Acad Sci 113:6641–6646. https://doi.org/10.1073/pnas.1522306113

Chew SJ, Mello C, Nottebohm F, Jarvis E, Vicario DS (1995) Decrements in auditory responses to a repeated conspecific song are long-lasting and require two periods of protein synthesis in the songbird forebrain. Proc Natl Acad Sci 92:3406–3410. https://doi.org/10.1073/pnas.92.8.3406

Clayton NS (1990) Subspecies recognition and song learning in zebra finches. Anim Behav 40:1009–1017. https://doi.org/10.1016/S0003-3472(05)80169-1

Clayton NS (1988) Song discrimination learning in zebra finches. Anim Behav 36:1016–1024. https://doi.org/10.1016/S0003-3472(88)80061-7

Dai JB, Chen Y, Sakata JT (2018) EGR-1 expression in catecholamine-synthesizing neurons reflects auditory learning and correlates with responses in auditory processing areas. Neuroscience 379:415–427. https://doi.org/10.1016/j.neuroscience.2018.03.032

David SV, Mesgarani N, Fritz JB, Shamma SA (2009) Rapid synaptic depression explains nonlinear modulation of Spectro-temporal tuning in primary auditory cortex by natural stimuli. J Neurosci 29:3374–3386. https://doi.org/10.1523/JNEUROSCI.5249-08.2009

Dent ML, McClaine EM, Best V, Ozmeral E, Narayan R, Gallun FJ, Sen K, Shinn-Cunningham BG (2009) Spatial unmasking of birdsong in zebra finches (Taeniopygia guttata) and budgerigars (Melopsittacus undulatus). J Comp Psychol 123:357–367. https://doi.org/10.1037/a0016898

Dong S, Clayton DF (2009) Habituation in songbirds. Neurobiol Learn Mem 92:183–188. https://doi.org/10.1016/j.nlm.2008.09.009

Dooling R, Searcy M (1980) Early perceptual selectivity in the swamp sparrow. Dev Psychobiol 13:499–506. https://doi.org/10.1002/dev.420130508

Dugas-Ford J, Rowell JJ, Ragsdale CW (2012) Cell-type homologies and the origins of the neocortex. Proc Natl Acad Sci 109:16974–16979. https://doi.org/10.1073/pnas.1204773109

Dunning JL, Pant S, Bass A, Coburn Z, Prather JF (2014) Mate choice in adult female Bengalese finches: females express consistent preferences for individual males and prefer female-directed song performances. PLoS One 9:e89438. https://doi.org/10.1371/journal.pone.0089438

Elliott TM, Theunissen FE (2011) The avian auditory pallium. In: Winer JA, Schreiner CE (eds) The auditory cortex. Springer US, Boston, MA, pp 429–442

Eriksson D, Wallin L (1986) Male bird song attracts females — a field experiment. Behav Ecol Sociobiol 19:297–299. https://doi.org/10.1007/BF00300645

Fields HL, Hjelmstad GO, Margolis EB, Nicola SM (2007) Ventral tegmental area neurons in learned appetitive behavior and positive reinforcement. Annu Rev Neurosci 30:289–316. https://doi.org/10.1146/annurev.neuro.30.051606.094341

Fisher J (1954) Evolution and bird sociality. Evol Process:71–83

Gentner TQ, Hulse SH, Duffy D, Ball GF (2001) Response biases in auditory forebrain regions of female songbirds following exposure to sexually relevant variation in male song. J Neurobiol 46:48–58. https://doi.org/10.1002/1097-4695(200101)46:1<48::AID-NEU5>3.0.CO;2-3

Gentner TQ, Margoliash D (2003) Neuronal populations and single cells representing learned auditory objects. Nature 424:669–674. https://doi.org/10.1038/nature01731

Gill P, Woolley SMN, Fremouw T, Theunissen FE (2008) What's that sound? Auditory area CLM encodes stimulus surprise, not intensity or intensity changes. J Neurophysiol 99:2809–2820. https://doi.org/10.1152/jn.01270.2007

Gobes SMH, Bolhuis JJ (2007) Birdsong memory: a neural dissociation between song recognition and production. Curr Biol 17:789–793. https://doi.org/10.1016/j.cub.2007.03.059

Gobes SMH, ter Haar SM, Vignal C, Vergne AL, Mathevon N, Bolhuis JJ (2009) Differential responsiveness in brain and behavior to sexually dimorphic long calls in male and female zebra finches. J Comp Neurol 516:312–320. https://doi.org/10.1002/cne.22113

Gobes SMH, Zandbergen MA, Bolhuis JJ (2010) Memory in the making: localized brain activation related to song learning in young songbirds. Proc R Soc Lond B Biol Sci 277:3343–3351. https://doi.org/10.1098/rspb.2010.0870

Graham DJ, Field DJ (2007) Efficient coding of natural images. New Encyclopedia of Neuroscience 1

Hansen BJ, Chelaru MI, Dragoi V (2012) Correlated variability in laminar cortical circuits. Neuron 76:590–602. https://doi.org/10.1016/j.neuron.2012.08.029

Harris KD, Mrsic-Flogel TD (2013) Cortical connectivity and sensory coding. Nature 503:51–58. https://doi.org/10.1038/nature12654

Hauber ME, Woolley SMN, Cassey P, Theunissen FE (2013) Experience dependence of neural responses to different classes of male songs in the primary auditory forebrain of female songbirds. Behav Brain Res 243:184–190. https://doi.org/10.1016/j.bbr.2013.01.007

Heil P, Scheich H (1991) Functional organization of the avian auditory cortex analogue. I. Topographic representation of isointensity bandwidth. Brain Res 539:110–120. https://doi.org/10.1016/0006-8993(91)90692-O

Hernandez AM, MacDougall-Shackleton SA (2004) Effects of early song experience on song preferences and song control and auditory brain regions in female house finches (Carpodacus mexicanus). J Neurobiol 59:247–258. https://doi.org/10.1002/neu.10312

Horita H, Kobayashi M, Liu W, Oka K, Jarvis ED, Wada K (2012) Specialized motor-driven dusp1 expression in the song Systems of Multiple Lineages of vocal learning birds. PLoS One 7:e42173. https://doi.org/10.1371/journal.pone.0042173

Horita H, Wada K, Rivas M, Hara E, Jarvis ED (2010) The dusp1 immediate early gene is regulated by natural stimuli predominantly in sensory input neurons. J Comp Neurol NA-NA:NA. https://doi.org/10.1002/cne.22370

Hose B, Langner G, Scheich H (1987) Topographic representation of periodicities in the forebrain of the mynah bird: one map for pitch and rhythm? Brain Res 422:367–373. https://doi.org/10.1016/0006-8993(87)90946-2

Ikeda MZ, Jeon SD, Cowell RA, Remage-Healey L (2015) Norepinephrine modulates coding of complex vocalizations in the songbird auditory cortex independent of local Neuroestrogen synthesis. J Neurosci 35:9356–9368. https://doi.org/10.1523/JNEUROSCI.4445-14.2015

Immelmann K (1969) Song development in the zebra finch and other estrildid finches. Bird Vocalizations:61–77

Jeanne JM, Thompson JV, Sharpee TO, Gentner TQ (2011) Emergence of learned categorical representations within an auditory forebrain circuit. J Neurosci 31:2595–2606. https://doi.org/10.1523/JNEUROSCI.3930-10.2011

Kao MH, Brainard MS (2006) Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. J Neurophysiol 96:1441–1455. https://doi.org/10.1152/jn.01138.2005

Karten HJ (1969) The Organization of the Avian telencephalon and some speculations on the phylogeny of the Amniote telencephalon∗. Ann N Y Acad Sci 167:164–179. https://doi.org/10.1111/j.1749-6632.1969.tb20442.x

Karten HJ (2013) Neocortical evolution: neuronal circuits Arise independently of lamination. Curr Biol 23:R12–R15. https://doi.org/10.1016/j.cub.2012.11.013

Kim G, Doupe A (2011) Organized representation of Spectrotemporal features in songbird auditory forebrain. J Neurosci 31:16977–16990. https://doi.org/10.1523/JNEUROSCI.2003-11.2011

Konishi M (2004) The role of auditory feedback in birdsong. Ann N Y Acad Sci 1016:463–475. https://doi.org/10.1196/annals.1298.010

Kruse A (2004) Context-specific habituation of the zenk gene response to song in adult zebra finches. Neurobiol Learn Mem 82:99–108. https://doi.org/10.1016/j.nlm.2004.05.001

Kruse AA, Stripling R, Clayton DF (2000) Minimal experience required for immediate-early gene induction in Zebra finch Neostriatum. Neurobiol Learn Mem 74:179–184. https://doi.org/10.1006/nlme.2000.3968

Lampen J, Jones K, McAuley JD, Chang S-E, Wade J (2014) Arrhythmic song exposure increases ZENK expression in auditory cortical areas and nucleus Taeniae of the adult Zebra finch. PLoS One 9:e108841. https://doi.org/10.1371/journal.pone.0108841

Lauay C, Gerlach NM, Adkins-Regan E, DeVoogd TJ (2004) Female zebra finches require early song exposure to prefer high-quality song as adults. Anim Behav 68:1249–1255. https://doi.org/10.1016/j.anbehav.2003.12.025

Leitner S, Voigt C, Metzdorf R, Catchpole CK (2005) Immediate early gene (ZENK, arc) expression in the auditory forebrain of female canaries varies in response to male song quality. J Neurobiol 64:275–284. https://doi.org/10.1002/neu.20135

London SE, Clayton DF (2008) Functional identification of sensory mechanisms required for developmental song learning. Nat Neurosci 11:579–586. https://doi.org/10.1038/nn.2103

Lu K, Vicario DS (2017) Familiar but unexpected: effects of sound context statistics on auditory responses in the songbird forebrain. J Neurosci 37:12006–12017. https://doi.org/10.1523/JNEUROSCI.5722-12.2017

Lynch KS, Ball GF (2008) Noradrenergic deficits Alter processing of communication signals in female songbirds. Brain Behav Evol 72:207–214. https://doi.org/10.1159/000157357

Machens CK, Wehr MS, Zador AM (2004) Linearity of cortical receptive Fields measured with natural sounds. J Neurosci 24:1089–1100. https://doi.org/10.1523/JNEUROSCI.4445-03.2004

Mandelblat-Cerf Y, Las L, Denisenko N, Fee MS (2014) A role for descending auditory cortical projections in songbird vocal learning. eLife. https://elifesciences.org/articles/02152.

Marler P (1970) A comparative approach to vocal learning: song development in white-crowned sparrows. J Comp Physiol Psychol 71:1–25. https://doi.org/10.1037/h0029144

Marler P, Peters S (1977) Selective vocal learning in a sparrow. Science 198:519–521. https://doi.org/10.1126/science.198.4316.519

Marler P, Peters S (1981) Sparrows learn adult song and more from memory. Science 213:780–782. https://doi.org/10.1126/science.213.4509.780

Martins ARO, Froemke RC (2015) Coordinated forms of noradrenergic plasticity in the locus coeruleus and primary auditory cortex. Nat Neurosci, 18:1483. https://doi.org/10.1038/nn.4090

Meliza CD, Margoliash D (2012) Emergence of selectivity and tolerance in the avian auditory cortex. J Neurosci 32:15158–15168. https://doi.org/10.1523/JNEUROSCI.0845-12.2012

Mello C, Nottebohm F, Clayton D (1995) Repeated exposure to one song leads to a rapid and persistent decline in an immediate early gene's response to that song in zebra finch telencephalon. J Neurosci 15:6919–6925. https://doi.org/10.1523/JNEUROSCI.15-10-06919.1995

Miller DB (1979a) The acoustic basis of mate recognition by female zebra finches. Anim Behav 27:376–380

Miller DB (1979b) Long-term recognition of father's song by female zebra finches. Nature 280:389–391. https://doi.org/10.1038/280389a0

Moore RC, Lee T, Theunissen FE (2013) Noise-invariant neurons in the avian auditory cortex: hearing the song in noise. PLoS Comput Biol 9:e1002942. https://doi.org/10.1371/journal.pcbi.1002942

Moore JM, Woolley SM (2019) Emergent tuning for learned vocalizations in auditory cortex. Nat Neurosci 22:1469–76. https://doi.org/10.1038/s41593-019-0458-4

Moseley LJ (1979) Individual auditory recognition in the least tern (Sterna albifrons). Auk Ornithol Adv 96:31–39. https://doi.org/10.1093/auk/96.1.31

Müller CM, Leppelsack H-J (1985) Feature extraction and tonotopic organization in the avian auditory forebrain. Exp Brain Res 59:587–599. https://doi.org/10.1007/BF00261351

Müller SC, Scheich H (1985) Functional organization of the avian auditory field L. J Comp Physiol A 156:1–12. https://doi.org/10.1007/BF00610661

Nagel K, Kim G, McLendon H, Doupe A (2011) A bird brain's view of auditory processing and perception. Hear Res 273:123–133. https://doi.org/10.1016/j.heares.2010.08.008

Nagel KI, Doupe AJ (2008) Organizing principles of spectro-temporal encoding in the avian primary auditory area Field L. Neuron 58:938–955. https://doi.org/10.1016/j.neuron.2008.04.028

Nelson DA (2000) A preference for own-subspecies' song guides vocal learning in a song bird. Proc Natl Acad Sci 97:13348–13353. https://doi.org/10.1073/pnas.240457797

Nelson DA, Marler P (1993) Innate recognition of song in white-crowned sparrows: a role in selective vocal learning? Anim Behav 46:806–808. https://doi.org/10.1006/anbe.1993.1258

Nordeen KW, Nordeen EJ (1992) Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. Behav Neural Biol 57:58–66. https://doi.org/10.1016/0163-1047(92)90757-U

Okanoya K, Dooling RJ (1987) Hearing in passerine and psittacine birds: a comparative study of absolute and masked auditory thresholds. J Comp Psychol 101:7–15. https://doi.org/10.1037/0735-7036.101.1.7

Petrinovich L, Baptista LF (1987) Song development in the white-crowned sparrow: modification of learned song. Anim Behav 35:961–974. https://doi.org/10.1016/S0003-3472(87)80153-7

Phan ML, Pytte CL, Vicario DS (2006) Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds. Proc Natl Acad Sci 103:1088–1093. https://doi.org/10.1073/pnas.0510136103

Ribeiro S, Cecchi GA, Magnasco MO, Mello CV (1998) Toward a song code: evidence for a syllabic representation in the canary brain. Neuron 21:359–371. https://doi.org/10.1016/S0896-6273(00)80545-0

Richner H (2016) Interval singing links to phenotypic quality in a songbird. Proc Natl Acad Sci 113:12763–12767. https://doi.org/10.1073/pnas.1610062113

Riebel K (2009) Chapter 6 song and female mate choice in Zebra finches: a review. In: Advances in the study of behavior. Academic, pp 197–238. https://doi.org/10.1016/S0065-3454(09)40006-8

Robertson BC (1996) Vocal mate recognition in a monogamous, flock-forming bird, the silvereye, Zosterops lateralis. Anim Behav 51:303–311. https://doi.org/10.1006/anbe.1996.0030

Sakata JT, Vehrencamp SL (2012) Integrating perspectives on vocal performance and consistency. J Exp Biol 215:201–209. https://doi.org/10.1242/jeb.056911

Sara SJ, Bouret S (2012) Orienting and reorienting: the locus Coeruleus mediates cognition through arousal. Neuron 76:130–141. https://doi.org/10.1016/j.neuron.2012.09.011

Schneider DM, Woolley SMN (2013) Sparse and background-invariant coding of vocalizations in auditory scenes. Neuron 79:141–152. https://doi.org/10.1016/j.neuron.2013.04.038

Schubloom HE, Woolley SC (2016) Variation in social relationships relates to song preferences and EGR1 expression in a female songbird. Dev Neurobiol 76:1029–1040. https://doi.org/10.1002/dneu.22373

Searcy WA, Brenowitz EA (1988) Sexual differences in species recognition of avian song. Nature 332:152–154. https://doi.org/10.1038/332152a0

Sen K, Theunissen FE, Doupe AJ (2001) Feature analysis of natural sounds in the songbird auditory forebrain. J Neurophysiol 86:1445–1458. https://doi.org/10.1152/jn.2001.86.3.1445

Sossinka R, Böhner J (1980) Song types in the zebra finch poephila guttata castanotis1. Z Für Tierpsychol 53:123–132. https://doi.org/10.1111/j.1439-0310.1980.tb01044.x

Striedter GF (1997) The telencephalon of tetrapods in evolution; pp. 205–213. Brain Behav Evol 49:205–213. https://doi.org/10.1159/000112992

Stripling R, Milewski L, Kruse AA, Clayton DF (2003) Rapidly learned song-discrimination without behavioral reinforcement in adult male zebra finches (Taeniopygia guttata). Neurobiol Learn Mem 79:41–50. https://doi.org/10.1016/S1074-7427(02)00005-9

Temeles EJ (1994) The role of neighbours in territorial systems: when are they "dear enemies". Anim Behav 47:339–350. https://doi.org/10.1006/anbe.1994.1047

Terleph TA, Lu K, Vicario DS (2008) Response properties of the auditory telencephalon in songbirds change with recent experience and season. PLoS One 3:e2854. https://doi.org/10.1371/journal.pone.0002854

Terpstra NJ, Bolhuis JJ, Riebel K, van der BJMM, den B-VAM (2006) Localized brain activation specific to auditory memory in a female songbird. J Comp Neurol 494:784–791. https://doi.org/10.1002/cne.20831

Theunissen FE, Amin N, Shaevitz SS, Woolley SMN, Fremouw T, Hauber ME (2004) Song selectivity in the song system and in the auditory forebrain. Ann N Y Acad Sci 1016:222–245. https://doi.org/10.1196/annals.1298.023

Theunissen FE, Doupe AJ (1998) Temporal and spectral sensitivity of complex auditory neurons in the nucleus HVc of male zebra finches. J Neurosci 18:3786–3802. https://doi.org/10.1523/JNEUROSCI.18-10-03786.1998

Theunissen FE, Sen K, Doupe AJ (2000) Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. J Neurosci 20:2315–2331. https://doi.org/10.1523/JNEUROSCI.20-06-02315.2000

Tibbetts EA, Dale J (2007) Individual recognition: it is good to be different. Trends Ecol Evol 22:529–537. https://doi.org/10.1016/j.tree.2007.09.001

Vallet E, Beme I, Kreutzer M (1998) Two-note syllables in canary songs elicit high levels of sexual display. Anim Behav 55:291–297. https://doi.org/10.1006/anbe.1997.0631

Vallet E, Kreutzer M (1995) Female canaries are sexually responsive to special song phrases. Anim Behav 49:1603–1610. https://doi.org/10.1016/0003-3472(95)90082-9

Van Ruijssevelt L, Chen Y, von Eugen K, Hamaide J, De Groof G, Verhoye M, Güntürkün O, Woolley SC, Van der Linden A (2018) fMRI reveals a novel region for evaluating acoustic information for mate choice in a female songbird. Curr Biol 28:711–721.e6. https://doi.org/10.1016/j.cub.2018.01.048

Vates GE, Broome BM, Mello CV, Nottebohm F (1996) Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches (Taenopygia guttata). J Comp Neurol 366:613–642. https://doi.org/10.1002/(SICI)1096-9861(19960318)366:4<613::AID-CNE5>3.0.CO;2-7

Velho TAF, Lu K, Ribeiro S, Pinaud R, Vicario D, Mello CV (2012) Noradrenergic control of gene expression and long-term neuronal adaptation evoked by learned vocalizations in songbirds. PLoS One 7:e36276. https://doi.org/10.1371/journal.pone.0036276

Wang Y, Brzozowska-Prechtl A, Karten HJ (2010) Laminar and columnar auditory cortex in avian brain. Proc Natl Acad Sci 107:12676–12681. https://doi.org/10.1073/pnas.1006645107

White SJ (1971) Selective responsiveness by the gannet (Sula bassana) to played-back class. Anim Behav 19:125–131. https://doi.org/10.1016/S0003-3472(71)80146-X

Woolley SC, Doupe AJ (2008) Social context–induced song variation affects female behavior and gene expression. PLoS Biol 6:e62. https://doi.org/10.1371/journal.pbio.0060062

Woolley SMN (2012) Early experience shapes vocal neural coding and perception in songbirds. Dev Psychobiol 54:612–631. https://doi.org/10.1002/dev.21014

Woolley SMN (2008) Auditory feedback and singing in adult birds. In: Zeigler HP, Marler P (eds) Neuroscience of birdsong. Cambridge University Press, Cambridge, pp 228–239

Woolley SMN, Moore JM (2011) Coevolution in communication senders and receivers: vocal behavior and auditory processing in multiple songbird species. Ann N Y Acad Sci 1225:155–165. https://doi.org/10.1111/j.1749-6632.2011.05989.x

Woolley SMN, Rubel EW (1997) Bengalese finches Lonchura Striata domestica depend upon auditory feedback for the maintenance of adult song. J Neurosci 17:6380–6390. https://doi.org/10.1523/JNEUROSCI.17-16-06380.1997

Woolley SMN, Rubel EW (1999) High-frequency auditory feedback is not required for adult song maintenance in bengalese finches. J Neurosci 19:358–371. https://doi.org/10.1523/JNEUROSCI.19-01-00358.1999

Woolley SMN, Fremouw TE, Hsu A, Theunissen FE (2005) Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. Nat Neurosci 8:1371–1379. https://doi.org/10.1038/nn1536

Woolley SMN, Gill PR, Fremouw T, Theunissen FE (2009) Functional groups in the avian auditory system. J Neurosci 29:2780–2793. https://doi.org/10.1523/JNEUROSCI.2042-08.2009

Woolley SMN, Hauber ME, Theunissen FE (2010) Developmental experience alters information coding in auditory midbrain and forebrain neurons. Dev Neurobiol 70:235–252. https://doi.org/10.1002/dneu.20783

Yanagihara S, Yazaki-Sugiyama Y (2016) Auditory experience-dependent cortical circuit shaping for memory formation in bird song learning. Nat Commun 7:11946. https://doi.org/10.1038/ncomms11946

Yang LM, Vicario DS (2015) Exposure to a novel stimulus environment alters patterns of lateralization in avian auditory cortex. Neuroscience 285:107–118. https://doi.org/10.1016/j.neuroscience.2014.10.022

Ydenberg RC, Giraldeau LA, Falls JB (1988) Neighbours, strangers, and the asymmetric war of attrition. Anim Behav 36:343–347. https://doi.org/10.1016/S0003-3472(88)80004-6

Zann RA (1996) The Zebra finch: a synthesis of Field and laboratory studies. Oxford University Press

Zaretsky MD, Konishi M (1976) Tonotopic organization in the avian telencephalon. Brain Res 111:167–171. https://doi.org/10.1016/0006-8993(76)91058-1

Chapter 6
Hormonal Regulation of Avian Auditory Processing

Luke Remage-Healey

Abstract This chapter explores the current understanding of the hormonal regulation of auditory function in songbirds by focusing on three themes. The first section is an overview of seasonal changes in the auditory pathway that are regulated by hormones. Next, the concept of the songbird brain as both a source and a target of neuromodulatory steroid hormones is discussed in the context of auditory function. Finally, the way that hormones interact with classical neurotransmitter systems (the biogenic amines dopamine, norepinephrine, and serotonin) to modulate auditory processing is presented. Reflecting on the sum total of these studies, understanding of the hormonal regulation of auditory function in songbirds has progressed considerably in the past few decades. More broadly, the field of songbird neuroethology has been continually propelled by an integrative perspective that examines the development, evolution, and hormonal modulation of neural circuits for song production, learning, and processing. This holistic approach to songbird neuroethology research, inspired by Niko Tinbergen and Peter Marler, will continue to be important as an increasing number of tools become available to explore the brain and behavior of songbirds.

Keywords Auditory lobule · CM · Dopamine · Estrogens · Field L · Song learning · NCM · Neuroestrogens · Neuromodulator · Neurophysiology · Norephinephrine · Song selectivity

---

L. Remage-Healey (✉)
Department of Psychological and Brain Sciences, University of Massachusetts Amherst,
Amherst, MA, USA
e-mail: healey@cns.umass.edu

J. T. Sakata et al. (eds.), *The Neuroethology of Birdsong*, Springer Handbook
of Auditory Research 71, https://doi.org/10.1007/978-3-030-34683-6_6

6.1 Introduction

The study of song production, learning, and perception in oscine songbirds has led to extraordinary advances in neuroethology. Songbirds are the premiere neuroethological system for understanding vocal learning and neuroplasticity (see Sakata and Yazaki-Sugiyama, Chap. 2; Murphy, Lawley, Smith, and Prather, Chap. 3). The current state of progress owes its origins to formative neuroethological approaches and insights made in the 1950s and 1960s by William Thorpe, Peter Marler, and their contemporaries. Their visionary work set the major research trajectories for the field of songbird neuroethology. The classic framework outlined by Niko Tinbergen as “the four questions” (Tinbergen 1951) emphasized integrating the study of *neural and hormonal mechanisms*, *adaptive significance, ontogeny,* and *phylogeny* (Adkins-Regan 2005; Marler 2008). The field of songbird neuroethology has since matured and diversified impressively along these key levels of analysis, yet currently these levels are at best only partially integrated. Peter Marler clearly articulated the benefits of the integrative approach throughout his distinguished career (Marler 2008). The true richness in the neuroethology of birdsong continues to be revealed when findings are integrated across these key levels of analysis, leading to deep and lasting insights (Sakata and Vehrencamp 2012; Hofmann et al. 2016).

Social behavior in songbirds is intimately linked to the production, learning, and sensory processing of vocalizations. The auditory system of songbirds is a network of interconnected regions that methodically unpack increasingly complex song features as the signals progress from the cochlear nucleus up through the midbrain, thalamus, and cortex (Fig. 6.1) (e.g., Woolley and Woolley, Chap. 5; Elie and Theunissen, Chap. 7). A central role for hormones in regulating the songbird brain has been evident since the foundational times of Peter Marler’s lab at Rockefeller University (reviewed in Marler 2008). Studies of sex differences and seasonal plasticity in the brain and the hormonal regulation of vocal behavior and territoriality have repeatedly emphasized the power of steroid hormones in shaping the songbird brain and regulating behavioral flexibility. This chapter reviews some of the recent work on how the songbird auditory system is regulated by hormones, including those hormones that are produced within the brain itself. As this field continues to mature, integrating the perspective of the songbird brain as a neurohormonal source/target with orthogonal and complimentary studies of neural mechanisms, ontogeny, adaptive value, and species-level comparisons will continue to be important. Together, this integration will provide the best path forward to achieve a deep and rich understanding of the neuroethology of birdsong.

6.1.1 Hormones, Singing, and Behavioral Responses

The study of the hormonal regulation of birdsong has been guided by the observation of a relationship between breeding and singing in seasonally breeding songbirds. The temporal correspondence between elevated circulating testosterone levels

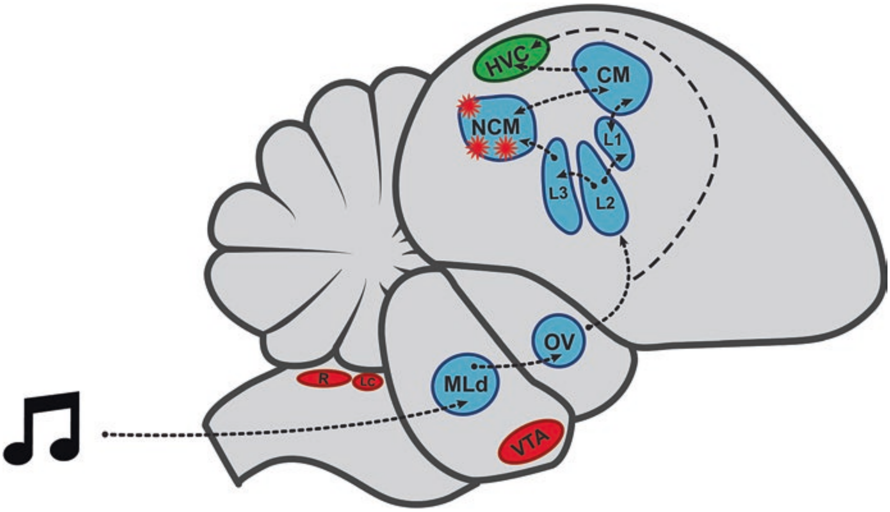


Fig. 6.1 Schematic of the songbird ascending auditory pathway with relevant neuromodulatory inputs to auditory and sensorimotor areas. The regions dedicated to auditory processing are shown in *blue* and neuromodulatory areas are in *red* (sensorimotor *HVC* is shown in *green*). Auditory stimuli are first processed in the cochlear nucleus and brainstem (not shown), followed by the midbrain *MLd* (dorsal lateral nucleus of the mesencephalon) and the thalamic *Ov* (ovoidalis). Akin to the mammalian ascending auditory pathway, the avian thalamus sends a primary projection into the cortical nucleus Field L2 (subdivisions *L1, L2, L3* shown), and signals are further processed in the reciprocally connected secondary cortical regions, the caudomedial nidopallium (*NCM*) and caudal mesopallium (*CM*) (see Woolley and Woolley, Chap. 5; Elie and Theunissen, Chap. 7). Auditory signals reach the sensorimotor *HVC* by way of *CM* and other sources (not pictured). Also not depicted are the descending forebrain basal ganglia and song motor circuit pathways that guide song learning and song production, respectively. The midbrain ventral tegmental area (*VTA*) provides dopaminergic modulatory input to the forebrain, while the brainstem raphe (*R*) and locus coeruleus (*LC*) provide serotonergic and noradrenergic modulatory inputs to the forebrain, respectively; the *red stars* in NCM indicate a substantial production of neuroestrogens that have local neuromodulatory effects. (adapted in part from Vahaba and Remage-Healey 2018)

in males and the onset of song production (Zigmond et al. 1973; Marler et al. 1987) and territoriality (Wingfield et al. 1987; Marler et al. 1988) was established at about the same time that the neural substrates for singing were first unveiled (Arnold et al. 1976; Nottebohm and Arnold 1976). This was a fortuitous coincidence, as it has become clear that testosterone and its metabolites, such as estrogens (Ball et al. 2002; Meitzen et al. 2007b), are essential for the motor control of song by neurons in the forebrain (Smith et al. 1997; Brenowitz and Lent 2002) and for the function of other neural and neuromuscular loci (Brenowitz 2004; Alward et al. 2013).

As findings accumulated in multiple songbird species regarding the role of androgens in song motor control, interest also grew in the way the songbird auditory system is similarly regulated by steroid hormones. These investigations have revealed that, just like the motor pathway, the auditory pathway of songbirds is regulated by testosterone and its metabolites at all levels, from the peripheral auditory hair cells up through the brainstem, midbrain, and forebrain (Caras and

Remage-Healey 2016). The pleiotropic regulation of the songbird auditory pathway by hormones at multiple timescales is the subject of interest in this chapter (Fig. 6.1).

In examining the receiver side of songbird communication, a strong link has been established between auditory-dependent behaviors and circulating levels of estrogens, particularly in females. Plasma estrogens are elevated in females during nest building and egg laying (Wingfield and Farner 1976) and in response to the continued presence of a live male (Tchernichovski et al. 1998). Playback of conspecific songs can also increase fecundity in female canaries (Bentley 2000). The resultant elevated levels of circulating estrogens likely have multiple actions on the reproductive axis, the brain, and behavior (Maney and Rodriguez-Saltos 2016).

The behavioral responses of female songbirds to song playback are regulated by estrogens. When female songbirds were treated with implants containing crystalline 17-beta-estradiol (E_2, a potent estrogen), they often responded behaviorally to song playback by adopting a characteristic posture, known as a copulation solicitation display (Maney et al. 1997). Females treated with estrogens also approached nest boxes broadcasting song (Gentner and Hulse 2000) and exhibited high levels of motivation in an operant task in which they pressed a key to hear song playback (Maney and Rodriguez-Saltos 2016). Clearly, the auditory system is highly sensitive to reproductive state in female songbirds as they listen to—and base mating decisions on—the performance aspects of the songs they hear (see Woolley and Woolley, Chap. 5; Podos and Sung, Chap. 9).

6.1.2 *The Songbird Brain as a Steroid Powerhouse*

Pioneering work in the 1970s (Reddy et al. 1973; Naftolin et al. 1975), revealed that the central nervous system of vertebrates contains the enzymatic machinery to produce its own supply of steroid hormones. The enzyme aromatase catalyzes the local production of estrogens, is expressed at high levels in the hypothalamus, and is consequently associated with reproduction in most vertebrate classes (Schlinger and Brenowitz 2008). While neurosteroid production is common among vertebrates, neuroestrogen production is especially pronounced in the brains of teleost fishes (Forlano et al. 2001) and birds (Balthazart and Ball 2012; Schlinger and Remage-Healey 2012). For example, in addition to aromatase expression in the hypothalamus, songbirds also have elevated expression of aromatase in the forebrain, in particular in the ascending auditory pathway and hippocampus (Saldanha et al. 2013; Vahaba and Remage-Healey 2015). Contrary to the common assumption that estrogens are female-typical hormones, the brain in both male and female songbirds has substantial capacity for estrogen synthesis and action. Aromatase is expressed not only in the cell bodies of neurons of the songbird auditory forebrain but also at their presynaptic terminals (Peterson et al. 2005). This subcellular localization at precise synaptic sites is consistent with the emerging picture of estrogens as genuine neuromodulators of auditory processing, in conjunction with the more long-term reproductive role of estrogens derived from peripheral sources (see Sect. 6.3) (Remage-Healey 2014a). In fact, the songbird forebrain itself contains all of the enzymes necessary to synthesize estrogens de novo from the primary precursor molecule cholesterol (London et al. 2006), and estradiol levels in the plasma of males actually originate in the brain itself (Schlinger and Arnold 1992). As described in the sections that follow, recent work shows that this enzymatic capacity has multiple implications for brain-derived *neuroestrogens* in modulating auditory gain and coding in the songbird brain.

6.1.3 Steroid Receptors in the Auditory Pathway

The neuroanatomy of the ascending auditory pathway in songbirds (Fig. 6.1) is described in detail in other chapters in this volume (e.g., Woolley and Woolley, Chap. 5; Elie and Theunissen, Chap. 7). However, to set the stage for the rest of this chapter, it is important to briefly describe the expression of steroid receptors in the auditory pathway. Classical nuclear androgen receptors are found throughout the song control system as well as in the auditory forebrain (Metzdorf et al. 1999; Ball et al. 2003). Similarly, classical nuclear estrogen receptors (ERalpha and ERbeta) are found at many sites in the ascending auditory pathway (Caras and Remage-Healey 2016). In both males and females, the classical ERalpha is found in auditory hair cells (within the inner ear itself) along with aromatase (Noirot et al. 2009), indicating that seasonal, reproductive, and local modulatory signaling events can all potentially shape auditory processing directly at the site of peripheral encoding. Key regions, such as the auditory cortical area (caudomedial nidopallium, NCM) and the sensorimotor nucleus HVC (letters used as a proper name), exhibit expression of both androgen and estrogen receptors (Bernard et al. 1999; Metzdorf et al. 1999), and their expression levels can fluctuate seasonally (Fusani et al. 2000). This arrangement likely provides cells in these regions with dynamic sensitivity to androgens, estrogens, and the ratio of androgens to estrogens to regulate important sensory and sensorimotor functions. More recently, a G-protein-coupled membrane estrogen receptor (GPER1) has received a great deal of attention because it can mediate rapid, modulatory effects of estrogens in cancer cell lines and within the brains of mammals (Srivastava and Evans 2013; Crimins et al. 2017) and teleosts (Mangiamele et al. 2017). Interestingly, GPER1 is also found in the songbird brain, in particular within the auditory regions HVC, NCM (all abbreviations appear in Table 6.1), and caudal mesopallium (CM) (Acharya and Veney 2012; Krentzel et al. 2018). Therefore, it is now clear that the songbird ascending auditory pathway is sensitive to both androgens and estrogens at multiple locations from peripheral to central nuclei.

Table 6.1 Abbreviations

ABR	auditory brainstem response
BOS	bird's own song
CM	caudal mesopallium
CMM	caudomedial mesopallium
ER	estrogen receptor (alpha and beta)
GPER1	G-protein-coupled membrane-bound estrogen receptor
HVC	used as a proper name for vocal motor nucleus in the nidopallium
IEG	immediate early gene
LC	locus coeruleus
MLd	dorsal lateral nucleus of the mesencephalon
NCM	caudomedial nidopallium
NE	norepinephrine
R	dorsal raphe
VTA	ventral tegmental area

6.2 Seasonal Plasticity and Hormonal Regulation of Auditory Processing

As discussed in Sect. 6.1.1, there is a strong correspondence between seasonal rhythms and plasticity in the songbird brain. This section describes some of the recent discoveries about the seasonal regulation of songbird auditory brain circuits and behaviors by hormones.

Songbirds are highly sensitive to seasonal cues (including photoperiod, temperature, rainfall, and food abundance) to time their breeding events. Migratory and sedentary birds typically show dramatic changes in anatomy and physiology as they transition from the nonbreeding to the breeding season, including changes in steroid hormone secretion from the gonads and season-dependent growth and regression of the gonads (Nottebohm et al. 1987; Schlinger and Brenowitz 2008). As mentioned in Sect. 6.1.2, the songbird brain itself makes its own supply of steroid hormones, and this capacity changes according to season to meet the demands of both breeding and nonbreeding (Soma et al. 1999, 2003). As a result of both peripheral and central changes in steroid levels, songbirds exhibit season-dependent changes in the anatomy and physiology of song system circuits (Brenowitz et al. 1991; Meitzen et al. 2007a).

In the auditory system, the auditory brainstem response (ABR) is a summed evoked potential of thousands of neurons, often recorded noninvasively. An ABR typically peaks at 10–15 ms after the presentation of an auditory stimulus and is used to make inferences about peripheral (brainstem) auditory functions. In a series of studies on a diversity of songbirds, including nuthatches, chickadees, and woodpeckers, Lucas and colleagues (Lucas et al. 2002; Henry and Lucas 2009) showed that the ABR changed according to season. While the direction and nature of these changes are species-specific, the observations indicated that seasonal changes in auditory processing occur as early as the level of the brainstem (Lucas et al. 2002; Henry and Lucas 2009). Subsequent studies in other seasonally breeding songbirds showed that breeding-dependent changes in ABR thresholds and frequency sensitivity were associated with elevations in plasma estradiol levels (Caras et al. 2010; Gall et al. 2013).

The seasonal effects of estrogens have become more evident in the peripheral auditory system of birds (Forlano et al. 2006), fish (Sisneros and Bass 2003; Sisneros et al. 2004), frogs (Arch and Narins 2009), and mammals (Petrulis 2013). Recent attention has focused, therefore, on the seasonal and hormonal modulation of central auditory processing in songbirds. That is, do hormones and/or season regulate the response properties of neurons in the songbird midbrain and forebrain? Caras and colleagues discovered that forebrain neurons in Field L of white-crowned sparrows (*Zonotrichia leucophrys*) had lower detection thresholds and greater dynamic coding ranges during the breeding season than during the nonbreeding season (Caras et al. 2012). Interestingly, the strength of neuronal auditory responses to tones was directly related to plasma estradiol levels in a cell-type specific manner. A follow-up study in Field L and CM showed that the seasonal auditory enhancement was also reflected in the spike timing-based encoding of song, as revealed by a *pattern classifier analysis* of neuronal responses to tones (Caras et al. 2015). Together, these studies demonstrate that peripheral and central auditory responses and encoding are enhanced during the breeding season and are most likely associated with the actions of estrogens.

In parallel with the electrophysiological findings, a similar line of investigation has taken advantage of the neural induction of immediate early genes (IEGs) in response to socially relevant acoustic signals (see also London, Chap. 8). Among vertebrates generally, hormonal regulation of IEG responses to social stimuli has been observed in the midbrain and forebrain of anurans (Lynch and Wilczynski 2008; Chakraborty and Burmeister 2015), teleosts (Okuyama et al. 2011; Maruska et al. 2013), and mammals (Abizaid et al. 2004; Tachikawa et al. 2013). In songbirds, the seasonal and hormonal modulation of IEG induction has focused on the auditory forebrain.

Mello and colleagues originally identified the auditory regions NCM and CM by IEG induction studies (Mello et al. 1992), and the upregulation of IEGs in these regions reflects higher-order processing functions like auditory memorization (Chew et al. 1995; Gobes and Bolhuis 2007) and the processing of social context (Gentner et al. 2001; Woolley and Doupe 2008). The seasonal and hormonal regulation of IEG responses in auditory areas, including NCM, the caudomedial mesopallium (CMM), and the dorsal lateral nucleus of the mesencephalon (MLd) of female white-throated sparrows (*Zonotrichia albicollis*) has been extensively investigated by Maney and colleagues. These studies have taken advantage of the sparrows' inherent seasonal sensitivity to photoperiod. In one study, female sparrows exhibited a near-complete regression of gonadal function during short days, and the levels of circulating estrogens plummeted accordingly. In these females, playback of songs or tones elicited comparably low IEG responses in the auditory forebrain. In a separate set of females treated with estradiol implants, IEG responses to songs as compared to tones were enhanced in the auditory forebrain (Maney et al. 2006). Further studies showed that this selective enhancement of response to songs following estrogen treatment was specific to subregions within the NCM (Sanford et al. 2010), which is consistent with subregion-dependent neurochemical modulation of auditory processing in the songbird auditory forebrain (see Sect. 6.4) (Matragrano et al. 2012a). Studies in other songbird species, such as black-capped chickadees (*Poecile atricapillus*), have reinforced the idea that IEG induction within the auditory forebrain in response to song stimuli changes seasonally (Phillmore et al. 2011), likely via interactions with hormones. Lastly, noninvasive brain imaging studies have shown that estrogen implants can modulate the forebrain auditory responses of female house sparrows (*Passer domesticus*) (Lattin et al. 2017) and that auditory responses change seasonally in starlings (*Sturnus vulgaris*) (De Groof et al. 2013). Therefore, the expression of steroid receptors in the ascending auditory pathway of songbirds is clearly related to the seasonal regulation of singing and processing song.

6.3 Rapid Steroid Modulation of Auditory Processing

The seasonal regulation of auditory function by steroid hormones like estrogens is consistent with the "classical" receptor mechanism of steroid action via intracellular receptors. The first receptors to be described for steroids were intracellular receptors that bind steroids and translocate to the nucleus to alter gene transcription events. The effects on brain and behavior, therefore, were traditionally considered to occur over the timescale of hours, days, and sometimes even weeks following initiation. In the last forty years, however, a more rapid "nonclassical" mode of action of steroids in the brain has become irrefutable. Initially, estradiol was shown to change the firing rate of rat hypothalamic neurons within seconds of application (Kelly et al. 1976), and additional observations of the rapid effects of steroids on brain and behavior soon followed, which were led by studies on rough-skinned newts (*Taricha granulosa*; Boyd and Moore 1990; Orchinik et al. 1991). Continuing investigations over the last twenty years have revealed that all major classes of steroid hormones can act within seconds to minutes in the brain to shape neuronal activity and behavior (Foradori et al. 2007; Vasudevan and Pfaff 2008). This perspective has merged with the concept that estrogens and other steroids can be synthesized rapidly within the brain and, therefore, be considered as genuine neuromodulators (Balthazart and Ball 2006; Remage-Healey 2014b).

Over the last ten years, the songbird auditory forebrain has been investigated to understand the behavioral significance of rapid neuroestrogen synthesis and action. In vivo microdialysis techniques were adapted to measure rapid (~30 min) fluctuations in estradiol levels and to manipulate brain estrogen levels in adult songbirds (Remage-Healey et al. 2008). A detailed description of these methods and their application can be found elsewhere (Ikeda et al. 2014; Caras and Remage-Healey 2016). These studies showed that neuroestrogen levels (i.e., concentrations of

17-beta-estradiol that were made locally within the NCM) in both sexes were elevated during 30 min of song playback but not during playback of duration-matched white noise stimuli (Remage-Healey et al. 2008, 2012). Therefore, the adult songbird NCM is sensitive to socially relevant auditory stimuli via a local elevation in neuroestrogens.

By contrast, the response pattern was completely different in juvenile birds. Namely, in juveniles, neuroestrogen levels were significantly suppressed in NCM during song playback itself and were elevated only during a 60-min silent period after the playback of song stopped (Chao et al. 2014). One hypothesis derived from these observations is that the early song learning period must be "protected" from the plasticity associated with neuroestrogens during periods of high-fidelity song encoding (Vahaba and Remage-Healey 2015; Vahaba et al. 2017). From this work using in vivo microdialysis, it became clear that neuroestrogens dynamically fluctuate in the songbird auditory forebrain in response to social and auditory contexts, and the fluctuations and actions of neuroestrogens are different depending on developmental age.

Given that neuroestrogens are dynamically regulated within NCM, it followed that they likely modulate the activity of neurons both local to and downstream of NCM. The first set of experiments testing this idea showed that estradiol caused a rapid switch from tonic firing to burst firing for individual NCM neurons in adult zebra finches, which is consistent with a transient increase in synaptic drive (Remage-Healey et al. 2010). This was accompanied by a rapid increase in auditory-evoked firing rates of NCM neurons during estradiol treatment that was observed in both adult males and females (Remage-Healey et al. 2010, 2012). Neurons in the NCM of juvenile males are also acutely sensitive to estrogens but in a hemisphere-dependent and age-dependent manner that likely reflects a developmental maturation (Vahaba et al. 2017). The modulation of NCM neuronal activity also has consequences for downstream processing in the song system network. For example, in adult males, estradiol treatment in NCM enhanced response selectivity for the bird's own song (BOS) downstream in HVC (Remage-Healey and Joshi 2012). Therefore, neuroestrogens in NCM act not only to enhance song representations locally within NCM but also to alter the sensorimotor representation of the song motor program in downstream areas. These findings also suggest that, for songbird species that are highly photoperiodic, changes in day length may modulate sensory-driven neuroestrogen synthesis in NCM and shape auditory and sensorimotor representations. Indirect support for this idea comes from studies of adult canaries in which the representation of BOS is seasonally shifted in HVC (Del Negro and Edeline 2002; Del Negro et al. 2005). In sum, the songbird NCM and its downstream targets are rapidly modulated by local actions of estrogens.

The rapid nonclassical regulation of neural function and behavior by steroids is now considered to involve a host of potential mechanisms, including interactions with cell-membrane trafficked versions of the classical nuclear hormone receptors, initiation of intracellular signaling cascades, direct actions at ion channels, and binding to modified membrane receptors like G-protein-coupled receptors (reviewed in Balthazart et al. 2018). Neurons in the songbird NCM express nuclear estrogen receptors (Bernard et al. 1999; Metzdorf et al. 1999) as well as GPER1 (Acharya and Veney 2012; Krentzel et al. 2018). Experiments with the membrane-restricted estrogen conjugate E-6-biotin showed that the rapid effects of 17-beta-estradiol on NCM firing modulation could be accounted for by actions restricted to the neuronal membrane (Remage-Healey et al. 2012) and were not mimicked by selective agonists for the classical nuclear hormone receptors (Remage-Healey et al. 2013). These findings indicated that neuroestrogens can modulate NCM neurons via a membrane-bound receptor.

The membrane-restricted nature of the rapid effects of estrogens have led to recent investigations of the nature of the membrane receptor, including potential interactions with metabotropic glutamate receptors and membrane GPER1 in the avian brain (Seredynski et al. 2015). In the NCM, reverse dialysis of a GPER1 receptor antagonist decreased auditory-evoked responses and coding accuracy of narrow-spiking neurons in adult male zebra finches but had no effect in adult females (Krentzel et al. 2018). Those findings also revealed a sex difference in the song-evoked auditory coding in NCM, in which neurons in females had higher evoked-firing rates than NCM neurons in males. This difference was unexpected due to a lack of gross morphological differences in the NCM of males versus females. Together, these results showed that acute neuroestrogen actions can account for (or compensate for) this sex difference in the auditory physiology of NCM neurons. Despite these recent advances, we currently lack a complete picture of the receptor mechanisms that can account for rapid effects of estrogens in the songbird brain (Heimovics et al. 2015).

In parallel, a companion set of studies confirmed that suppressing the endogenous synthesis of neuroestrogens within NCM had consequences for auditory coding and behavior. Reverse microdialysis of fadrozole, a selective inhibitor of the estrogen synthesis enzyme (aromatase), caused a rapid suppression of the auditory-evoked bursting of NCM neurons in adult males and a rapid suppression of auditory-evoked firing rates of NCM neurons in adult females (Remage-Healey et al. 2010, 2012). Similar treatments also suppressed the representation of BOS downstream in HVC (Remage-Healey and Joshi 2012) and within a sensorimotor nucleus that passes information between NCM and HVC (Pawlisch and Remage-Healey 2015). Importantly, all of these observations with aromatase inhibition were directly opposite to those observed with 17-beta-estradiol treatments, providing definitive evidence that neuroestrogens precisely and rapidly tune the firing state of NCM neurons within a dynamic, modulatory range. Further independent confirmation of these findings was obtained using fMRI in adult starlings, in which acute aromatase inhibition dampened auditory responses in the left NCM region (De Groof et al. 2017).

At the level of singing and auditory-dependent behaviors, there have been only a few tests of a role for neuroestrogen synthesis and action in the NCM of songbirds. In phonotaxis and/or operant behavioral tests, zebra finches expressed a preference to hear playbacks of a familiar song (e.g., their own song or a song of their tutor) versus an unfamiliar song (Riebel et al. 2002; Holveck and Riebel 2007). In awake, behaving adult male zebra finches, reverse microdialysis of the aromatase inhibitor fadrozole into NCM caused a significant impairment in the behavioral preference for familiar songs (Remage-Healey et al. 2010). Future tests are required to determine whether this behavioral result is due to an impairment in auditory discrimination, motivation, ranging, and/or other computations involved in phonotaxis. Together with electrophysiological data, these findings indicate that neuroestrogen fluctuations have functional significance for songbirds as they hear and interpret signals in their natural environment.

6.4 Hormone Regulation of Auditory Function Via Monoaminergic Signaling

The auditory forebrain of songbirds, like the mammalian auditory cortex, receives dense monoaminergic innervation from key nuclei (Fig. 6.1): the ventral tegmental area (VTA) in the case of dopaminergic inputs, the locus coeruleus (LC) in the case of noradrenergic inputs, and the dorsal raphe (R) in the case of serotonergic inputs. Therefore, the modulation of auditory functions by hormones, in many cases, is likely to involve co-regulation of these important neurotransmitter systems to shape stimulus salience, valence, coding, and reward (Maney and Rodriguez-Saltos 2016).

The experience of hearing song impacts a variety of monoaminergic systems in the songbird brain. Catecholaminergic neurons (e.g., dopamine neurons in the VTA and noradrenergic neurons in the LC) are also hormone sensitive in songbirds (Barclay and Harding 1988, 1990), so it is reasonable to infer that the monoamine and hormone signals interact. Song playback leads to the induction of IEG expression in both the VTA and LC of songbirds, and these responses were up-regulated by treatment with implants containing estradiol (LeBlanc et al. 2007; Kabelik et al. 2011). Song stimulation also can directly drive cathecholamine responses in the auditory forebrain of songbirds (Castelino and Schmidt 2010; Matragrano et al. 2012b). In addition, in the NCM and CM, the seasonal elevation in estrogens also increases auditory-evoked activity in serotonin neurons in female white-throated sparrows (Matragrano et al. 2012a).

In some cases, the actions of monoamines that shape auditory function in songbirds may or may not involve explicit actions of hormones. Norepinephrine (NE) can shape the IEG responses to song and the long-term memory function of neurons in NCM (Lynch et al. 2012; Velho et al. 2012), and NE can also modulate the activity of premotor neurons such as those in HVC and RA (Cardin and Schmidt 2004; Sizemore and Perkel 2008). Moreover, pharmacological inhibition of adrenergic signaling can impair hormone-dependent and auditory-dependent behaviors and IEG induction (Vyas et al. 2008; Pawlisch and Riters 2010). Finally, at the level of electrophysiology, NE acutely enhanced the auditory gain and coding accuracy of neurons in NCM (Ikeda et al. 2015; Lee et al. 2018), and this appeared to be independent of the local synthesis of estrogens within NCM itself (Ikeda et al. 2015). In summary, the above work shows that, in some cases, the hormonal regulation of auditory responses in songbirds involves a direct interaction with monoaminergic systems, which can support the animal's attending and assigning value to an auditory stimulus (Maney and Rodriguez-Saltos 2016). Clearly, continued investigation of the interaction between hormones and monoaminergic signaling in the songbird brain is warranted.

6.5 Summary and Conclusions

This review focuses on the central idea that hormones regulate auditory function in songbirds. The evidence has accumulated to such an extent that we can draw several broad conclusions from this work. First, steroid hormones, in particular estrogens, can shape the auditory responses to song throughout the auditory pathway across a variety of species. The seasonal up-regulation of estrogens in females increases auditory responses in terms of both behavior and brain circuits. Second, the male and female songbird brain also has the capacity for central neuroestrogen synthesis, which has a modulatory role in acute auditory responses. Third, like many other vertebrate classes, the processing of sensory information in the songbird auditory circuit is regulated by monoaminergic signaling, which most likely serves to encode stimulus valence and salience. With some exceptions, the auditory regulation of dopaminergic, noradrenergic, and serotoninergic signaling is also shaped by the actions of hormones. Most broadly, the field of songbird neuroethology has been pushed forward by an integrative perspective that examines how the development and function of neural circuits for song production, learning, and processing are modulated by hormones. This holistic view of songbird neuroethology research will continue to be important as new tools become available to further study and understand the brain and behavior of songbirds.

Acknowledgements Preparation of this chapter was supported in part by NIH R01NS082179 and NSF IOS 1354906.

Compliance with Ethics Requirements Luke Remage-Healey declares that he has no conflict of interest.

References

Abizaid A, Mezei G, Horvath TL (2004) Estradiol enhances light-induced expression of transcription factors in the SCN. Brain Res 1010(1–2):35–44

Acharya KD, Veney SL (2012) Characterization of the G-protein-coupled membrane-bound estrogen receptor GPR30 in the zebra finch brain reveals a sex difference in gene and protein expression. Dev Neurobiol 72(11):1433–1446

Adkins-Regan E (2005) Hormones and animal social behavior. Princeton University Press, Princeton

Alward BA, Balthazart J, Ball GF (2013) Differential effects of global versus local testosterone on singing behavior and its underlying neural substrate. Proc Natl Acad Sci U S A 110(48):19573–19578

Arch VS, Narins PM (2009) Sexual hearing: the influence of sex hormones on acoustic communication in frogs. Hear Res 252(1–2):15–20

Arnold AP, Nottebohm F, Pfaff DW (1976) Hormone concentrating cells in vocal control and other areas of brain of zebra finch (*Poephila-Guttata*). J Comp Neurol 165(4):487–511

Ball GF, Riters LV, Balthazart J (2002) Neuroendocrinology of song behavior and avian brain plasticity: multiple sites of action of sex steroid hormones. Front Neuroendocrinol 23(2):137–178

Ball, G. F., Castelino, C. B., Maney, D. L., Appeltants, D., & Balthazart, J. (2003). The activation of birdsong by testosterone - Multiple sites of action and role of ascending catecholamine projections *Steroids and the Nervous System,* 1007, 211–231

Balthazart J, Ball GF (2006) Is brain estradiol a hormone or a neurotransmitter? Trends in Neurosci 29(5):241–249

Balthazart J, Ball GF (eds) (2012) Brain aromatase, estrogens, and behavior. Oxford University Press, Oxford

Balthazart J, Choleris E, Remage-Healey L (2018) Steroids and the brain: 50years of research, conceptual shifts and the ascent of non-classical and membrane-initiated actions. Horm Behav 99:1–8

Barclay SR, Harding CF (1988) Androstenedione modulation of monoamine levels and turnover in hypothalamic and vocal control nuclei in the male zebra finch: steroid effects on brain monoamines. Brain Res 459(2):333–343

Barclay SR, Harding CF (1990) Differential modulation of monoamine levels and turnover rates by estrogen and/or androgen in hypothalamic and vocal control nuclei of male zebra finches. Brain Res 523(2):251–262

Bentley G (2000) Stimulatory effects on the reproductive axis in female songbirds by conspecific and heterospecific male song. Horm Behav 37(3):179–189

Bernard DJ, Bentley GE, Balthazart J, Turek FW, Ball GF (1999) Androgen receptor, estrogen receptor alpha, and estrogen receptor beta show distinct patterns of expression in forebrain song control nuclei of European starlings. Endocrinol 140(10):4633–4643

Boyd SK, Moore FL (1990) Evidence for gaba involvement in stress-induced inhibition of male amphibian sexual-behavior. Horm Behav 24(1):128–138

Brenowitz EA (2004) Plasticity of the adult avian song control system. Behavioral Neurobiology of Birdsong 1016:560–585

Brenowitz EA, Lent K (2002) Act locally and think globally: intracerebral testosterone implants induce seasonal-like growth of adult avian song control circuits. Proc Natl Acad Sci 99(19):12421–12426

Brenowitz EA, Nalls B, Wingfield JC, Kroodsma DE (1991) Seasonal-changes in avian song nuclei without seasonal-changes in song repertoire. J Neurosci 11(5):1367–1374

Caras ML, Remage-Healey L (2016) Modulation of peripheral and central auditory processing by estrogens in birds. In: Bass AH (ed) Hearing and hormones, vol 57. Springer, Heidelberg, pp 77–100

Caras ML, Brenowitz E, Rubel EW (2010) Peripheral auditory processing changes seasonally in Gambel's white-crowned sparrow. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 196(8):581–599

Caras ML, O'Brien M, Brenowitz EA, Rubel EW (2012) Estradiol selectively enhances auditory function in avian forebrain neurons. J Neurosci 32(49):17597–17611

Caras ML, Sen K, Rubel EW, Brenowitz EA (2015) Seasonal plasticity of precise spike timing in the avian auditory system. J Neurosci 35(8):3431–3445

Cardin JA, Schmidt MF (2004) Noradrenergic inputs mediate state dependence of auditory responses in the avian song system. J Neurosci 24(35):7745–7753

Castelino CB, Schmidt MF (2010) What birdsong can teach us about the central noradrenergic system. J Chem Neuroanat 39(2):96–111

Chakraborty, M., & Burmeister, S. S. (2015). Effects of estradiol on neural responses to social signals in female tungara frogs. J Exp Biol*,* 218(Pt 22), 3671–3677

Chao A, Paon A, Remage-Healey L (2014) Dynamic variation in forebrain estradiol levels during song learning. Dev Neurobiol

Chew SJ, Mello C, Nottebohm F, Jarvis E, Vicario DS (1995) Decrements in auditory responses to a repeated conspecific song are long-lasting and require 2 periods of protein-synthesis in the songbird forebrain. Proc Natl Acad Sci 92(8):3406–3410

Crimins JL, Wang AC, Yuk F, Puri R, Janssen WGM, Hara Y, Rapp PR, Morrison JH (2017) Diverse synaptic distributions of G protein-coupled estrogen receptor 1 in monkey prefrontal cortex with aging and menopause. Cereb Cortex 27(3):2022–2033

De Groof G, Poirier C, George I, Hausberger M, Van der Linden A (2013) Functional changes between seasons in the male songbird auditory forebrain. Front Behav Neurosci 7:196

De Groof G, Balthazart J, Cornil CA, Van der Linden A (2017) Topography and lateralized effect of acute aromatase inhibition on auditory processing in a seasonal songbird. J Neurosci 37(16):4243–4254

Del Negro C, Edeline JM (2002) Sex and season influence the proportion of thin spike cells in the canary HVc. Neuroreport 13(16):2005–2009

Del Negro C, Lehongre K, Edeline JM (2005) Selectivity of canary HVC neurons for the bird's own song: modulation by photoperiodic conditions. J Neurosci 25(20):4952–4963

Foradori C, Weiser M, Handa R (2007) Non-genomic actions of androgens. Front Neuroendocrinol

Forlano PM, Deitcher DL, Myers DA, Bass AH (2001) Anatomical distribution and cellular basis for high levels of aromatase activity in the brain of teleost fish: aromatase enzyme and mRNA expression identify glia as source. J Neurosci 21(22):8943–8955

Forlano PM, Schlinger BA, Bass AH (2006) Brain aromatase: new lessons from non-mammalian model systems. Front Neuroendocrinol 27(3):247–274

Fusani L, Van't Hof T, Hutchison JB, Gahr M (2000) Seasonal expression of androgen receptors, estrogen receptors, and aromatase in the canary brain in relation to circulating androgens and estrogens. J Neurobiol 43(3):254–268

Gall MD, Salameh TS, Lucas JR (2013) Songbird frequency selectivity and temporal resolution vary with sex and season. Proc Biol Sci 280(1751):20122296

Gentner TQ, Hulse SH (2000) Female European starling preference and choice for variation in conspecific male song. Anim Behav 59(2):443–458

Gentner TQ, Hulse SH, Duffy D, Ball GF (2001) Response biases in auditory forebrain regions of female songbirds following exposure to sexually relevant variation in male song. J Neurobiol 46(1):48–58

Gobes SMH, Bolhuis JJ (2007) Birdsong memory: a neural dissociation between song recognition and production. Curr Biol 17(9):789–793

Heimovics SA, Ferris JK, Soma KK (2015) Non-invasive administration of 17beta-estradiol rapidly increases aggressive behavior in non-breeding, but not breeding, male song sparrows. Horm Behav 69:31–38

Henry KS, Lucas JR (2009) Vocally correlated seasonal auditory variation in the house sparrow (*Passer domesticus*). J Exp Biol 212(23):3817–3822

Hofmann HA, Renn SC, Rubenstein DR (2016) Introduction to symposium: new frontiers in the integrative study of animal behavior: nothing in neuroscience makes sense except in the light of behavior. Int Comp Biol 56(6):1192–1196

Holveck MJ, Riebel K (2007) Preferred songs predict preferred males: consistency and repeatability of zebra finch females across three test contexts. Anim Behav 74:297–309

Ikeda M, Rensel MA, Schlinger BA, Remage-Healey L (2014) *In vivo* detection of fluctuating brain steroid levels in zebra finches. *Cold Spring Harb Protoc*

Ikeda MZ, Jeon SD, Cowell RA, Remage-Healey L (2015) Norepinephrine modulates coding of complex vocalizations in the songbird auditory cortex independent of local neuroestrogen synthesis. J Neurosci 35(25):9356–9368

Kabelik D, Schrock SE, Ayres LC, Goodson JL (2011) Estrogenic regulation of dopaminergic neurons in the opportunistically breeding zebra finch. Gen Comp Endocrinol 173(1):96–104

Kelly MJ, Moss RL, Dudley CA (1976) Differential sensitivity of preoptic-septal neurons to microelectrophoresed estrogen during the estrous cycle. Brain Res 114(1):152–157

Krentzel AA, Macedo-Lima M, Ikeda MZ, Remage-Healey L (2018) A membrane G-protein coupled estrogen receptor is necessary but not sufficient for sex-differences in zebra finch auditory coding. Endocrinology 159(3):1360–1376

Lattin CR, Stabile FA, Carson RE (2017) Estradiol modulates neural response to conspecific and heterospecific song in female house sparrows: an in vivo positron emission tomography study. PLoS One 12(8):e0182875

LeBlanc MM, Goode CT, MacDougall-Shackleton EA, Maney DL (2007) Estradiol modulates brainstem catecholaminergic cell groups and projections to the auditory forebrain in a female songbird. Brain Res 1171:93–103

Lee V, Pawlisch BA, Macedo-Lima M, Remage-Healey L (2018) Norepinephrine enhances song responsiveness and encoding in the auditory forebrain of male zebra finches. J Neurophysiol 119(1):209–220

London SE, Monks DA, Wade J, Schlinger BA (2006) Widespread capacity for steroid synthesis in the avian brain and song system. Endocrinol 147(12):5975–5987

Lucas JR, Freeberg TM, Krishnan A, Long GR (2002) A comparative study of avian auditory brainstem responses: correlations with phylogeny and vocal complexity, and seasonal effects. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 188(11–12):981–992

Lynch KS, Wilczynski W (2008) Reproductive hormones modify reception of species-typical communication signals in a female anuran. Brain Behav Evol 71(2):143–150

Lynch KS, Diekamp B, Ball GF (2012) Colocalization of immediate early genes in catecholamine cells after song exposure in female zebra finches (*Taeniopygia guttata*). Brain Behav Evol 79(4):252–260

Maney DL, Rodriguez-Saltos CA (2016) Hormones and the incentive salience of birdsong. In: Bass AH (ed) Hearing and hormones, vol 57. Springer, Heidelberg, pp 101–132

Maney DL, Richardson RD, Wingfield JC (1997) Central administration of chicken gonadotropin-releasing hormone-II enhances courtship behavior in a female sparrow. Horm Behav 32(1):11–18

Maney DL, Cho E, Goode CT (2006) Estrogen-dependent selectivity of genomic responses to birdsong. Eur J Neurosci 23(6):1523–1529

Mangiamele LA, Gomez JR, Curtis NJ, Thompson RR (2017) GPER/GPR30, a membrane estrogen receptor, is expressed in the brain and retina of a social fish (*Carassius auratus*) and colocalizes with isotocin. J Comp Neurol 525(2):252–270

Marler P (2008) Birdsong and monkey talk: an ethological journey. In: Zeigler HP, Marler P (eds) Neuroscience of birdsong. Cambridge University Press, New York, pp 449–462

Marler P, Peters S, Wingfield J (1987) Correlations between song acquisition, song production, and plasma-levels of testosterone and estradiol in sparrows. J Neurobiol 18(6):531–548

Marler P, Peters S, Ball GF, Dufty AM, Wingfield JC (1988) The role of sex steroids in the acquisition and production of birdsong. Nature 336(6201):770–772

Maruska, K. P., Becker, L., Neboori, A., & Fernald, R. D. (2013). Social descent with territory loss causes rapid behavioral, endocrine and transcriptional changes in the brain. J Exp Biol*,* 216(Pt 19), 3656–3666

Matragrano LL, Sanford SE, Salvante KG, Beaulieu M, Sockman KW, Maney DL (2012a) Estradiol-dependent modulation of serotonergic markers in auditory areas of a seasonally breeding songbird. Behav Neurosci 126(1):110–122

Matragrano LL, Beaulieu M, Phillip JO, Rae AI, Sanford SE, Sockman KW, Maney DL (2012b) Rapid effects of hearing song on catecholaminergic activity in the songbird auditory pathway. PLoS One 7(6):e39388

Meitzen J, Perkel DJ, Brenowitz EA (2007a) Seasonal changes in intrinsic electrophysiological activity of song control neurons in wild song sparrows. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 193(6):677–683

Meitzen J, Moore IT, Lent K, Brenowitz EA, Perkel DJ (2007b) Steroid hormones act transsynaptically within the forebrain to regulate neuronal phenotype and song stereotypy. J Neurosci 27(44):12045–12057

Mello CV, Vicario DS, Clayton DF (1992) Song presentation induces gene expression in the songbird forebrain. Proc Natl Acad Sci U S A 89(15):6818–6822

Metzdorf R, Gahr M, Fusani L (1999) Distribution of aromatase, estrogen receptor, and androgen receptor mRNA in the forebrain of songbirds and nonsongbirds. J Comp Neurol 407(1):115–129

Naftolin, F., Ryan, K. J., Davies, I. J., Reddy, V. V., Flores, F., Petro, Z., Kuhn, M., White, R. J., Takaoka, Y., & Wolin, L. (1975). The formation of estrogens by central neuroendocrine tissues. *Recent Progress in Hormone Research,* Proceedings of the 1996 Conference*, Vol* 52*,* 31, 295–319

Noirot IC, Adler HJ, Cornil CA, Harada N, Dooling RJ, Balthazart J, Ball GF (2009) Presence of aromatase and estrogen receptor alpha in the inner ear of zebra finches. Hear Res 252(1–2):49–55

Nottebohm F, Arnold A (1976) Sexual dimorphism in vocal control areas of the songbird brain. Science 194(4261):211–213

Nottebohm F, Nottebohm ME, Crane LA, Wingfield JC (1987) Seasonal-changes in gonadal hormone levels of adult male canaries and their relation to song. Behav Neural Biol 47(2):197–211

Okuyama T, Suehiro Y, Imada H, Shimada A, Naruse K, Takeda H, Kubo T, Takeuchi H (2011) Induction of c-fos transcription in the medaka brain (*Oryzias latipes*) in response to mating stimuli. Biochem Biophys Res Commun 404(1):453–457

Orchinik M, Murray TF, Moore FL (1991) A corticosteroid receptor in neuronal membranes. Science 252(5014):1848–1851

Pawlisch BA, Remage-Healey L (2015) Neuroestrogen signaling in the songbird auditory cortex propagates into a sensorimotor network via an ‘interface’ nucleus. Neuroscience 284:522–535

Pawlisch BA, Riters LV (2010) Selective behavioral responses to male song are affected by the dopamine agonist GBR-12909 in female european starlings (*Sturnus vulgaris*). Brain Res 1353:113–124

Peterson RS, Yarram L, Schlinger BA, Saldanha CJ (2005) Aromatase is pre-synaptic and sexually dimorphic in the adult zebra finch brain. Proc Roy Soc B-Biol Sci 272(1576):2089–2096

Petrulis A (2013) Chemosignals, hormones and mammalian reproduction. Horm Behav 63(5):723–741

Phillmore LS, Veysey AS, Roach SP (2011) Zenk expression in auditory regions changes with breeding condition in male black-capped chickadees (*Poecile atricapillus*). Behav Brain Res 225(2):464–472

Reddy VV, Naftolin F, Ryan KJ (1973) Aromatization in the central nervous system of rabbits: effects of castration and hormone treatment. Endocrinology 92(2):589–594

Remage-Healey L (2014a) Frank Beach award winner: steroids as neuromodulators of brain circuits and behavior. Horm Behav 66(3):552–560

Remage-Healey, L. (2014b)

Remage-Healey L, Joshi NR (2012) Changing neuroestrogens within the auditory forebrain rapidly transform stimulus selectivity in a downstream sensorimotor nucleus. J Neurosci 32(24):8231–8241

Remage-Healey L, Maidment NT, Schlinger BA (2008) Forebrain steroid levels fluctuate rapidly during social interactions. Nat Neurosci 11(11):1327–1334

Remage-Healey L, Coleman MJ, Oyama RK, Schlinger BA (2010) Brain estrogens rapidly strengthen auditory encoding and guide song preference in a songbird. Proc Natl Acad Sci U S A 107(8):3852–3857

Remage-Healey L, Dong SM, Chao A, Schlinger BA (2012) Sex-specific, rapid neuroestrogen fluctuations and neurophysiological actions in the songbird auditory forebrain. J Neurophysiol 107(6):1621–1631

Remage-Healey L, Jeon SD, Joshi NR (2013) Recent evidence for rapid synthesis and action of oestrogens during auditory processing in a songbird. J Neuroendocrinol 25(11):1024–1031

Riebel K, Smallegange IM, Terpstra NJ, Bolhuis JJ (2002) Sexual equality in zebra finch song preference: evidence for a dissociation between song recognition and production learning. Proc Roy Soc B-Biol Sci 269(1492):729–733

Sakata, J. T., & Vehrencamp, S. L. (2012). Integrating perspectives on vocal performance and consistency. J Exp Biol*,* 215(Pt 2), 201–209

Saldanha CJ, Remage-Healey L, Schlinger BA (2013) Neuroanatomical distribution of aromatase in birds: cellular and subcellular analyses. In: Balthazart GBJ (ed) Brain aromatase, estrogens and behavior. Oxford, UK, Oxford, pp 100–114

Sanford SE, Lange HS, Maney DL (2010) Topography of estradiol-modulated genomic responses in the songbird auditory forebrain. Dev Neurobiol 70(2):73–86

Schlinger BA, Arnold AP (1992) Circulating estrogens in a male songbird originate in the brain. Proc Natl Acad Sci U S A 89(16):7650–7653

Schlinger B, Brenowitz EA (2008) Neural and hormonal control of birdsong. In: Pfaff DW (ed) *Hormones, Brain and Behavior*, vol 2. Elsevier, pp 897–941

Schlinger BA, Remage-Healey L (2012) Neurosteroidogenesis: insights from studies of songbirds. J Neuroendocrinol 24(1):16–21

Seredynski AL, Balthazart J, Ball GF, Cornil CA (2015) Estrogen receptor beta activation rapidly modulates male sexual motivation through the transactivation of metabotropic glutamate receptor 1a. J Neurosci 35(38):13110–13123

Sisneros JA, Bass AH (2003) Seasonal plasticity of peripheral auditory frequency sensitivity. J Neurosci 23(3):1049–1058

Sisneros JA, Forlano PM, Deitcher DL, Bass AH (2004) Steroid-dependent auditory plasticity leads to adaptive coupling of sender and receiver. Science 305(5682):404–407

Sizemore M, Perkel DJ (2008) Noradrenergic and GABAB receptor activation differentially modulate inputs to the premotor nucleus RA in zebra finches. J Neurophysiol 100(1):8–18

Smith GT, Brenowitz EA, Beecher MD, Wingfield JC (1997) Seasonal changes in testosterone, neural attributes of song control nuclei, and song structure in wild songbirds. J Neurosci 17(15):6001–6010

Soma KK, Schlinger BA, Wingfield JC, Saldanha CJ (2003) Brain aromatase, 5 alpha-reductase, and 5 beta-reductase change seasonally in wild male song sparrows: relationship to aggressive and sexual behavior. J Neurobiol 56(3):209–221

Soma KK, Bindra RK, Gee J, Wingfield JC, Schlinger BA (1999) Androgen-metabolizing enzymes show region-specific changes across the breeding season in the brain of a wild songbird. J Neurobiol 41(2):176–188

Srivastava DP, Evans PD (2013) G-protein oestrogen receptor 1: trials and tribulations of a membrane oestrogen receptor. J Neuroendocrinol 25(11):1219–1230

Tachikawa KS, Yoshihara Y, Kuroda KO (2013) Behavioral transition from attack to parenting in male mice: a crucial role of the vomeronasal system. J Neurosci 33(12):5120–5126

Tchernichovski O, Schwabl H, Nottebohm F (1998) Context determines the sex appeal of male zebra finch song. Anim Behav 55:1003–1010

Tinbergen N (1951) The study of instinct. Clarendon Press, Oxford Eng

Vahaba DM, Remage-Healey L (2015) Brain estrogen production and the encoding of recent experience. Curr Op Behav Sci 6:148–153

Vahaba DM, Remage-Healey L (2018) Neuroestrogens rapidly shape auditory circuits to support communication learning and perception: Evidence from songbirds. Horm Behav 104:77–87. https://doi.org/10.1016/j.yhbeh.2018.03.007

Vahaba DM, Macedo-Lima M, Remage-Healey L (2017) Sensory coding and sensitivity to local estrogens shift during critical period milestones in the auditory cortex of male songbirds. eNeuro 4(6)

Vasudevan N, Pfaff DW (2008) Non-genomic actions of estrogens and their interaction with genomic actions in the brain. Front Neuroendocrinol 29(2):238–257

Velho TA, Lu K, Ribeiro S, Pinaud R, Vicario D, Mello CV (2012) Noradrenergic control of gene expression and long-term neuronal adaptation evoked by learned vocalizations in songbirds. PLoS One 7(5):e36276

Vyas A, Harding C, McGowan J, Snare R, Bogdan D (2008) Noradrenergic neurotoxin, N-(2-chloroethyl)-N-ethyl-2-bromobenzylamine hydrochloride (DSP-4), treatment eliminates estrogenic effects on song responsiveness in female zebra finches (*Taeniopygia guttata*). Behav Neurosci 122(5):1148–1157

Wingfield JC, Farner DS (1976) Plasma Lh and sex steroids in whitecrowned sparrow, *Zonotrichia-Leucophrys-Pugetensis*. Amer Zool 16(2):257–257

Wingfield JC, Ball GF, Dufty AM, Hegner RE, Ramenofsky M (1987) Testosterone and aggression in birds. Amer Sci 75(6):602–608

Woolley SC, Doupe AJ (2008) Social context-induced song variation affects female behavior and gene expression. PLoS Biol 6(3):e62

Zigmond RE, Notteboh F, Pfaff DW (1973) Androgen-concentrating cells in midbrain of a songbird. Science 179(4077):1005–1007

Chapter 7
The Neuroethology of Vocal Communication in Songbirds: Production and Perception of a Call Repertoire

Julie E. Elie and Frédéric E. Theunissen

> *"Bird calls are the neglected orphans of avian behavioral neurobiology."*
>
> – Marler 2004.

Abstract Oscines learn to produce a complex vocalization, the song, which they copy from a conspecific as young birds. The song is an attractive and conspicuous acoustic signal with striking spectral and temporal complexity. The oscine song copying behavior is also remarkable because vocal imitation is a relatively rare ability in vertebrates and because none of the nonavian species can outperform the best oscine mimics. Studies of the neurobiology of song learning have unraveled many of the mechanisms involved in this impressive vocal behavior. Song, however, is only one of the many vocalizations that are produced by oscines. The vocal repertoire of oscines is impressive not only because of the number of vocalizations produced but also because of the flexible production and usage of these sounds. This chapter reviews the vocal behavior of oscines in the framework of animal communication and examines the mechanisms underlying the production and perception of all vocalization types. The chapter also reviews how the auditory system and vocal and social brain networks might be connected to generate appropriate responses to communication calls and song. As a whole, this chapter argues that studies of the mechanisms underlying song learning and also the mechanisms underlying call plasticity, production, and perception are critical for understanding the neuroethology

---

J. E. Elie (✉)
Department of Bioengineering, University of California, Berkeley, CA, USA
e-mail: julie.elie@gmail.com

F. E. Theunissen
Departments of Psychology, Integrative Biology, and Helen Wills Neuroscience Institute,
University of California, Berkeley, CA, USA
e-mail: theunissen@berkeley.edu

J. T. Sakata et al. (eds.), *The Neuroethology of Birdsong*, Springer Handbook
of Auditory Research 71, https://doi.org/10.1007/978-3-030-34683-6_7

of vocal communication in oscines. Embracing the complexity of the vocal communication system of oscines will enhance our understanding of the brain areas that, until now, have mostly been studied in the context of song imitation.

Keywords Animal communication · Auditory categories · Auditory cortex · Auditory memory · Auditory objects · Neural invariance · Song imitation · Vocal learning · Vocal plasticity

7.1 Introduction

Research on the song-copying behavior of songbirds and its neural underpinnings is one of the success stories of neuroethology and the principal subject of this volume. Songbirds (oscines) learn to produce a copy of a tutor song by *imitation* and/or *selection of innate templates* (Marler 1997). Songbirds can not only imitate the song of a particular conspecific but, in at least several hundred oscine species, this vocal imitation ability is relatively unconstrained, yielding the remarkable capacity of *mimicry*: the imitation of vocalizations produced by heterospecifics, human speech or man-made sounds, or other natural sounds (Goller and Shizuka 2018). Although various forms of vocal plasticity are described in mammals (Janik and Slater 1997), vocal imitation has been shown in only a few mammalian clades: elephants (Stoeger and Manger 2014), marine mammals (Janik 2014), bats (Knörnschild et al. 2010), and humans. In nonhuman mammals, it is fair to say that vocal imitation is relatively rudimentary compared to what can be observed in champion songbird mimics such as the North American mockingbird, *Mimus polyglottos* (Gammon and Altizer 2011) or the Australian superb lyrebird, *Menura novaehollandiae* (Dalziell and Magrath 2012).

The chapters in this volume summarize the impressive extent of research into the neurobiology of song imitation and control. The motor programs used to produce song, along with circuits that provide sensory feedback and the reinforcement that guide vocal development have been investigated (see Sakata and Yazaki-Sugiyama, Chap. 2; Murphy, Lawley, Smith, and Prather, Chap. 3). Not only does that research provide potentially useful insights, by analogy, into human brain circuits and into genes involved in speech production and learning (see London, Chap. 8), but more generally, songbird research has yielded multiple fundamental findings in neuroscience (see Sakata and Woolley, Chap. 1).

Most of the classical research in songbirds, however, has focused on the neural mechanisms underlying a precise behavior: the imitation and control of a species-typical, stereotyped vocalization, the *song*. In these neurobiological studies, the function of the song as a communication signal has been ignored for the most part. Which neural mechanisms support how the emitter chooses whether and what to sing and which neural computations are necessary for the correct interpretation of this song and for the choice of the adaptive response by receivers are questions that have rarely been addressed. Yet, oscines sing in various contexts: as a territorial signal, as a sexual display, or both (see Podos and Sung, Chap. 9). To react appropriately, receivers must recognize the signal and interpret the meaning of the song.

Song can be interpreted as an aggressive signal. For example, a western song sparrow (*Melospiza melodia*) might not only counter-sing in response to hearing a neighbor's or an unfamiliar male's song but could repeatedly match the type of song to show his aggressive intents in a *song duel* (Searcy and Beecher 2009). Alternatively, song can be interpreted as an attractive affiliative signal. For example, a female yellowhammer (*Emberiza citronella*) might reply to a male singing in a familiar dialect with a copulation solicitation display (Baker et al. 1987). The brain networks that interpret the song signal and subsequently activate the brain regions for aggressive, avoidance, or sexual behaviors are unknown for the most part (see Woolley and Woolley, Chap. 5). A related question is: To what extent are the circuits that are involved in the recognition of song as a communication signal similar to those that are used by young birds for memorizing the songs of their tutor? Moreover, do the sensory-motor circuits that produce appropriate behavioral responses (e.g., counter-singing) overlap with the neural pathways involved in song learning and the initiation of spontaneous song production? One can already see that by placing the study of song production, perception, and learning back in the context of vocal communication, many outstanding questions with regard to the neuroethology of song arise.

However, song is only one of the many vocalization types produced by songbirds for communication. Indeed, songbirds possess a rich repertoire of vocalizations that they use to communicate states (e.g., stress), intents (e.g., territorial defense), needs (e.g., food), or desires (e.g., soliciting mate, partnership) (see Marler 2004). For instance, pair-bonded zebra finches (*Taeniopygia guttata*) exchange a series of loud contact calls when one is in need of the other, but they remain out of sight. By alternating their vocalizations in a call/call-back sequence they establish an acoustic contact. Songbirds, therefore, must recognize the song signal to distinguish conspecific song from heterospecific vocalizations or environmental sounds and also, probably much more frequently, recognize and categorize all the vocalizations in their species' repertoire. Given that, it is important to reveal whether the transformation of the entire repertoire of conspecific sounds into meaning is mediated by the same high-level auditory areas and association areas as those involved in the recognition of song as a conspecific signal. Similarly, to what extent are the circuits mediating behavioral responses to song different from those mediating responses to other vocalization types? Is the decision to produce call-back responses to contact calls generated in separate motor circuits than the decision involved in counter-singing? Are sequences of calls generated in separate motor circuits from those involved in the generation of song syllable sequences?

Songbird vocalizations also carry information about the identity or group membership of the signaler. In many oscine species, song dialects are characteristic of local groups and are used for recognition and mate choice by males and females (Slabbekoorn and Smith 2002). In other species, such as the Australian zebra finch (*Taeniopygia guttata*), songs have a strong individual signature and can be used to recognize one's mate (Miller 1979a; Woolley and Doupe 2008), father (Miller 1979b) and peers (Honarmand et al. 2015). However, individual recognition by vocalizations is not restricted to song; distance calls (Vignal et al. 2004) and begging calls (Ligout et al. 2016) are also used for individual recognition in juveniles and adults. Both the recognition of a song dialect and the recognition of individuals based on calls or song require the formation and use of auditory memories.

With regard to song copying, an auditory memory (*template*) of the tutor's song must be created by juvenile birds prior to the formation of the motor memory of the bird's own song (Gobes and Bolhuis 2007); this is particularly true for species for which the sensory phase is distinct from the sensory-motor phase (Beecher and Brenowitz 2005; see Sakata and Woolley, Chap. 1). An important question in this regard is whether auditory memories of the vocalizations (calls and songs) of particular groups or individuals are distinct from the auditory memory of the tutor song used for song copying?

While most of these questions remain unanswered, this chapter reviews the research on the perception and production of all vocalizations in a songbird's repertoire within the context of vocal communication. Then that body of research is discussed in relation to the more classical neurobiological studies of song production and perception in the context of song copying. A central objective of this chapter is to examine how the neuroethology of vocal communication, taken in the larger context that includes all vocalizations and all related behaviors, could inform the research on the neural basis of song imitation and vice-versa.

7.2 The Vocal Repertoire of Songbirds

7.2.1 What Is a Vocal Repertoire?

The definition of vocal repertoire for songbirds, as well as social mammals, is the naturalist's attempt to document the range of sounds used for communication and, ultimately, to associate these acoustically distinct vocalizations to specific meanings as reflected by the context and the behaviors of the sender and receiver. The complexity of an oscine's vocal repertoire is species dependent, but a typical repertoire is composed of approximately ten distinct vocalization types in the social species (Elie and Theunissen 2016). The song is one of these vocalization types. Some of the earliest and most complete descriptions of songbird repertoires are for the European chaffinch, *Fringilla coelebs* (Marler 1956) and the Australian zebra finch (Zann 1996). As described by Peter Marler, categorization of vocalizations into types allows one to understand the vocabulary that is available for the "language" of a particular species (Marler 1956).

7.2.2 An Example: The Zebra Finch Repertoire

The vocal repertoire of the zebra finch is typical of songbirds with vocalizations that serve typical core social functions, from courting to alarm signaling. These functional categories are generally conserved, although many oscine species lack a particular vocalization or use vocalizations in different contexts (e.g., some species use a variant of the alarm call, namely a mobbing call, to coordinate attacks on predators) or have additional vocalization categories. The spectrograms shown in Fig. 7.1 illustrate the acoustic variation that is found in the zebra finch repertoire that comprises both affiliative and nonaffiliative vocalization types (Elie and Theunissen 2016).

Affiliative vocalizations include soft and loud contact calls, flight calls, soft nest calls, and the song. The soft contact calls are further separated into *tet* and *stack* calls. These soft contact calls are produced between zebra finches interacting within visual contact and appear to be directed to all group members; the rate of calling communicates the level of activity and is an invitation for social interaction. Zebra finches also produce a loud contact call, the *distance call*, used by visually separated individuals to find their mate or their social group by call-back exchanges. The distance call is sexually dimorphic, has very strong individual signatures, and triggers selective responses among mates in both males and females (Vicario et al. 2001; Vignal et al. 2004); therefore, the distance call clearly advertises the caller identity. Male and female juveniles produce a precursor of the distance call named the *long tonal call* that will mature into an adult distance call (Zann 1985). Although the distance call is thought to be innate in females, in male zebra finches the distance call is learned by imitation from the father (Zann 1990; Slater and Jones 1995). Zebra finches also make a series of acoustically distinct soft calls, *nest* and *whine calls*, that are used in courtship behavior and nest choice. Finally, the male song is also a typical vocalization of precopulatory courtship and, like the male distance call, it is also learned and imitated from the father (Zann 1990).

Nonaffiliative calls in the zebra finch repertoire include alarm calls, aggressive calls, and distress calls. Alarm calls are short vocalizations that can be segregated into two subtypes: the *thuk* and the *tuck*. The thuk is an alarm call produced only by parents and directed to their juveniles and the mate; juveniles respond by refraining from begging and the mate responds by fleeing the nest. In contrast, the tuck is a more general alarm call that is used to alert the entire group of the presence of a potential predator. Aggressive calls (also called the *wsst*) are produced in agonistic encounters between conspecifics. *Distress calls* are produced when zebra finches are attacked by a conspecific, and the distress call can communicate physical or social distress.

The division of vocalizations into discrete types is a useful simplification for beginning to understand the bird's vocabulary. However, using this categorical approach to describe the repertoire often does not fully incorporate the complexity

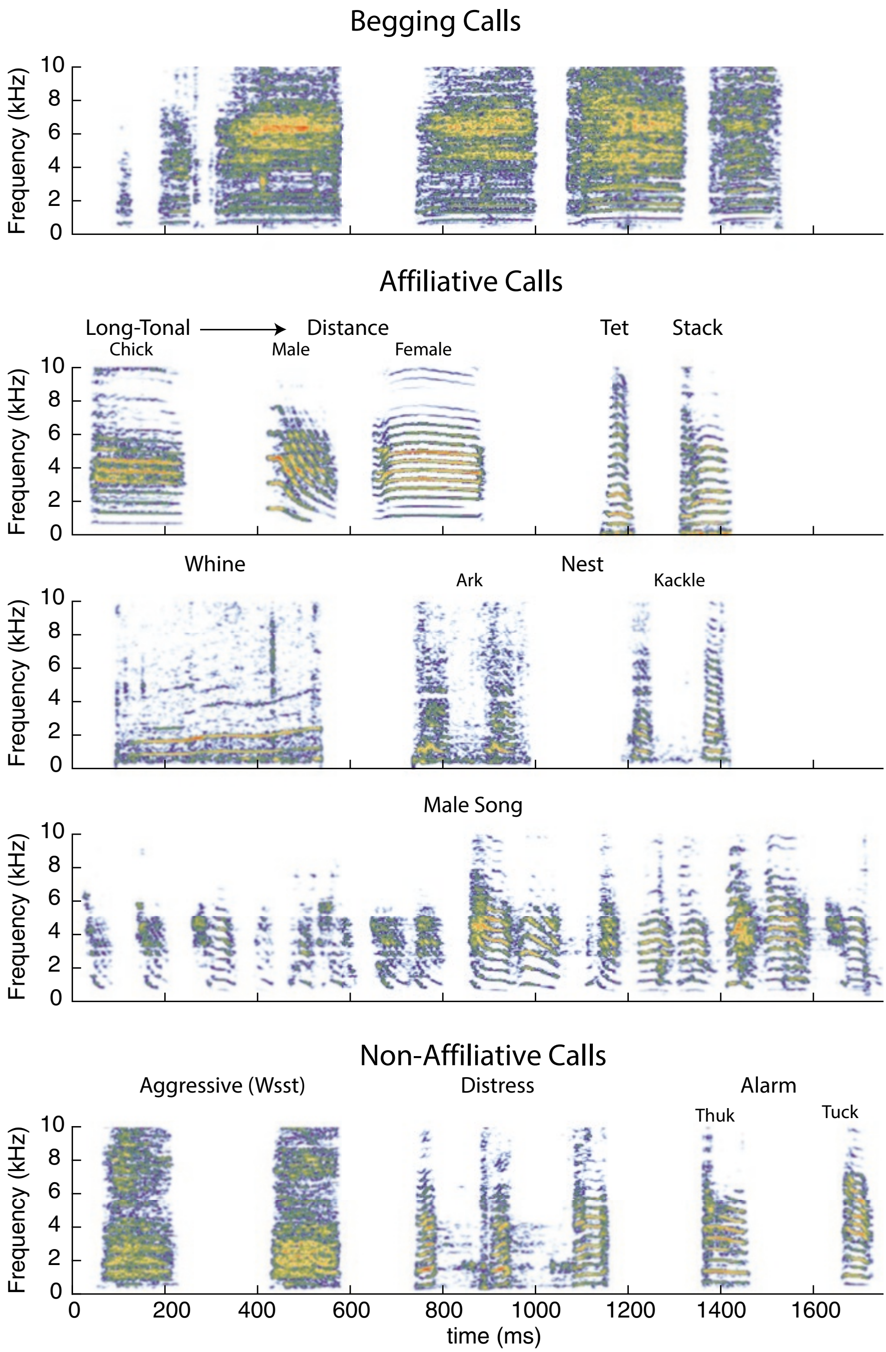


Fig. 7.1 Spectrograms of example calls from the vocal repertoire of zebra finches. The vocalization types can be coarsely segregated into chick begging calls, adult affiliative vocalizations used in social bonding, and adult non-affiliative vocalizations used in aggressive or alarming contexts. (*warm colors* on the spectrograms correspond to time-frequency points of high energy)

that is observed in avian vocal communication systems. The following section considers examples for which the simple categories break down. These examples are instructive for the discussion of the neurobiology of song imitation in the larger context of vocal communication.

7.2.3 A Misleading Distinction: Learned Song Versus Innate Calls

The distinction between learned song versus innate calls, conveniently made as a shorthand, reflects the fact that songbirds that are acoustically isolated or deafened during the typical period of song development will sing a highly abnormal song (called the *isolate song*) while still producing apparently normal calls (Marler 1970). The distinction is problematic for two reasons. One reason is that, unwillingly perhaps, it separates song from all other vocalization types and, by doing so, undervalues the role of song in the complex communication exchanges that are occurring between male and female birds. In the neurobiology of song imitation, song is often simply considered to be a complex sexual display signal (akin to bright plumage) for which acoustic complexity obtained by good imitation is the ultimate goal. In reality, song production is plastic, even in adults. Moreover, within a single species, song can have different meanings. For example, adult zebra finches produce song in two distinct contexts: when actively courting a female, the male addresses a female with a very stereotyped song, the *directed song*; when singing without an obvious song addressee, potentially for practice, advertisement, nest or mate guarding (Zann 1996), the male sings a more variable version of the same song, the *undirected song*. In territorial songbirds that have multiple song types, males can signal their level of aggressive intent to competitors based on the choice of the song type they produce in counter-singing bouts (Searcy and Beecher 2009).

The second issue is that nonsong vocalizations can also be shaped by imitation or vocal learning. For example, the distance call of the male zebra finch is imitated from the father with a developmental time frame and a degree of copying that is very similar to the song (Zann 1985; Slater and Jones 1995). In addition, like the song, most calls also have a developmental trajectory that can be influenced by the environment. For example, male and female zebra finch nestlings modify their begging calls when they are cross-fostered with Bengalese finches (*Lonchura striata domestica*), presumably to maximize the efficiency of their “feed me” signal (Villain et al. 2015). Similar vocal learning is also naturally found in the begging calls of bird species that are brood parasites (Langmore et al. 2008).

7.2.4 The Plasticity and Complexity of Calls

The classification of calls into categories, which is implicit in the nomenclature used to describe vocal repertoires, can also be misleading. It suggests that vocalizations have fixed meanings and that they are produced in an almost reflexive manner that is triggered by specific behavioral contexts (Seyfarth and Cheney 2010). In reality, vocal communication in songbirds relies both on categorical and graded signals (Marler 1967) and can even exhibit high degrees of vocal plasticity, such as phonemic contrast or syntactic properties, which enable the production of multiple meanings by using distinct combinations of the same set of sounds (Engesser et al. 2015; Griesser et al. 2018). The distance call of the zebra finch is an example of a graded signal: birds emit calls with higher spectral means when they are stressed (Perez et al. 2015).

There is also an impressive degree of plasticity in the usage and interpretation of vocalization types in songbirds that is often a result of experience (Marler 1982). The study of the production and interpretation of alarm calls is particularly relevant. Just as in primates, songbirds can produce alarm calls that are predator specific or situation specific. The calls are said to be *functionally referential* because they will elicit the same anti-predator behaviors in the receiver as if the receiver had seen the predator (Gill et al. 2013; Suzuki 2018). Chickadees (*Poecile carolinensis*) use a variable number and acoustically different notes in their "chick-a-dee-dee" that reflect both the emitter's locality and the presence and physical position of a predator (Freeberg 2008). Southern pied babblers (*Turdoides bicolor*) and Japanese tits (*Parus minor*) syntactically combine alarm calls and recruitment calls into a single sequence to encourage group members to participate in predator mobbing (Engesser et al. 2016; Suzuki et al. 2018). Songbirds can also react to the alarm or recruitment calls of heterospecifics. In superb fairy-wrens (*Malurus cyaneus*), this is a socially learned behavior that involves associating a new alarm sound (from the heterospecific) to one that is already known as a reliable alarm call from the conspecific (Potvin et al. 2018). African drongos (*Dicrurus adsimilis*) in the Kalahari use both their species alarm calls and the alarm calls of other species, which they have learned to imitate, to produce false alarms that facilitate food stealing from other species, including meerkats (Flower et al. 2014). This is a striking example that involves both vocal imitation and plasticity in the use of nonlearned calls. Collectively, these studies illustrate the plasticity, complexity, and learning mechanisms that can be engaged not only for song but also for the production and interpretation of many call types in the oscine vocal repertoire.

7.2.5 Cracking the Acoustic Code for Categories and Identity

With the advances of large-scale recordings and machine learning techniques, bioacoustic analyses of repertoires have been used to further test the idea of vocalization categories as well as to explicitly determine the acoustic features that distinguish vocalization types or vocalizers from each other. Identifying the information-bearing features in the acoustics of vocalizations is important because it informs researchers on how the sounds are shaped by vocal systems to produce informative signals. Similarly, at the perceptual level, identifying the distinguishing acoustic features of vocalizations can inform the understanding of how the auditory system can detect and categorize vocalizations to extract their meaning.

Unsupervised clustering shows that zebra finch vocalizations are a combination of acoustically distinct groups and graded signals (Elie and Theunissen 2016). These vocalization types are principally distinguished based on their *pitch saliency*, *duration*, and the *resonant peaks* in their *spectral envelopes* but not by their fundamental frequency. The pitch saliency quantifies the degree of harmonicity in the vocalizations and distinguishes tonal calls (e.g., distance and tet calls) from noisy calls (e.g., begging and aggressive calls). The pitch saliency is controlled by the vocal organ, the syrinx. The shape of the spectral envelope is also partly generated by the syrinx but is further determined by the bird's upper vocal tract and, in particular, by the volume of the oropharyngeal-esophageal cavity (OEC; all abbreviations appear in Table 7.1) (Riede et al. 2013). Duration is controlled by the timing of expirations and inspirations. Thus, coordination of the syrinx, the upper vocal tract, and the respiratory system is essential for producing distinct vocalization types (see Sect. 7.3). Similarly, individual signatures in zebra finch vocalizations cannot be explained solely on the small individual variations in vocal anatomy. Instead, zebra finches produce individualized vocalizations by actively shaping each of the vocalization types in a unique and reliable fashion. Individuality in zebra finch calls, therefore, also requires individual-specific neural activations of the syrinx and upper vocal tract (Elie and Theunissen 2018).

7.2.6 Temporal Complexity in Song and Other Vocalization Types

Song complexity is based on individual-specific acoustic patterns that are found not only at the level of a single syllable but also in the temporal sequencing of syllables (Lehongre et al. 2008). Indeed, one of the most striking distinctions made between song and other vocalizations is the temporal complexity of the song as a sequence of syllables. Both the spectral content and the temporal sequence can be learned with a great range of variability across species in the types of sounds and sequences that are learned from other conspecifics (Brenowitz and Beecher 2005).

Table 7.1 Abbreviations

AFP	anterior forebrain pathway
AH	Anterior hypothalamus
Ail	Lateral arcopallial nucleus
AIV	Ventral portion of the intermediate arcopallium
Av	Avalanche, a ventral region in CLM
CM	Caudal mesopallium (CLM + CMM)
CLM	Caudolateral mesopallium
CMM	Caudomedial mesopallium
DLM	Medial portion of the dorsolateral thalamic nucleus
DM	Dorsomedial nucleus of the intercollicular complex
HVC	Used as proper name for vocal motor nucleus in the nidopallium
ICo	Intercollicular complex
L1; L2a,b; L3	Subdivisions of primary auditory pallium field L
LL	Lateral lemniscus
LMAN	Lateral magnocellular nucleus of the anterior nidopallium
LS	Lateral septum
MLd	Dorsolateral nucleus of the mesencephalon
NA	Nucleus angularis
NCM	Caudomedial nidopallium
NIf	Nucleus interfacialis of the nidopallium
NL	Nucleus laminaris
NM	Nucleus magnocellularis
nXIIts	Tracheosyringeal part of hypoglossal nucleus (nXII)
Ov	Nucleus ovoidalis (medial geniculate body)
OEC	Oropharyngeal-esophageal cavity
PAG	Avian periaqueductal gray
PAm	Nucleus paraambigualus
POA	Preoptic area
POM	Medial preoptic nucleus
RA	Robust nucleus of the arcopallium
RA-cup	Arcopallium ventral and rostral to RA
RAm	Nucleus retroambigualis
RPcvm	Ventromedial region of parvocellular part of reticular formation
SO	Superior olive
Ts	Tracheosyringeal nerve
Uva	Nucleus uvaeformis
VMH	Ventromedial hypothalamus
VMHm	Ventromedial hypothalamic nucleus, medial portion
VTA	Ventral tegmental area
X	Area X, vocal portion of avian basal ganglia

Although the temporal complexity of song is missing, for the most part, in the other vocalization types produced by songbirds, it is not completely absent. The usage of a variable number of notes in the chickadee alarm call to communicate the emitter's position serves as a useful counterexample (Freeberg 2008). Furthermore, in some species, the temporal sequence of calls is highly informative: oscines combine single vocal elements to achieve different meanings by either rearranging meaningless sound units (Engesser et al. 2015) or by combining functionally distinct vocalizations (Engesser et al. 2016; Suzuki et al. 2018). In addition, the timing of contact calls in vocal exchanges obeys specific rules that reflect relationships among birds (Elie et al. 2011; D'Amelio et al. 2017). Given these degrees of temporal organization of calls, one could ask whether the brain areas that control the temporal organization of song are also involved in the timing and sequencing of other vocalization types.

7.2.7 Summary of Vocal Repertoire

Social songbirds use a complex array of vocalizations to communicate with conspecifics, including song and other vocalization types. Song is a distinctive vocalization type because of its acoustic complexity and because it is one of the vocalizations that is reliably imitated in the oscine repertoire. However, other vocalization types, such as the distance call in male zebra finches, are also imitated and, like song, many vocalization types show plasticity in production and usage. This plasticity allows more adaptive (e.g., efficient begging calls) or complex (e.g., referential alarm calls) communicative exchanges between conspecifics and heterospecifics and is often associated with learning during development. Such vocal learning, that does not necessarily implicate vocal imitation, is observed for many vocalizations produced as part of the oscine "language." One is led to ask, therefore, to what extent are the brain regions and neural mechanisms shown to be critical for song imitation and production also involved in learning and producing other vocalizations? And, similarly, to what extent are the brain regions and mechanisms involved in the perception and production of calls also involved in song imitation? Those questions are addressed in Sects. 7.3–7.5.

7.3 Repertoire Production in Songbirds

To investigate the neural control of all vocalizations, it is imperative to understand how sounds are produced and shaped by the peripheral vocal system. Great progress has been made in understanding vocal production in oscines and has resulted in modification of the more vocal-organ-centric view to embrace the importance of all the components of the peripheral vocal system (Podos and Sung, Chap. 9), including the respiratory system and the upper vocal tract (Schmidt and Wild 2014). Indeed, the *source-filter model* that describes how vocal sounds are shaped in mammals also applies to songbirds (Riede and Goller 2010). During controlled expirations, sounds are generated in the avian vocal organ, the syrinx, and then further shaped by the bird's upper vocal tract (Düring and Elemans 2016).

7.3.1 Vocal Apparatus

The vocal apparatus consists of a vocal organ, the respiratory system, and the upper vocal tract. The syrinx is a sophisticated instrument that has two independently controlled pairs of sound-producing membranes: one pair at the end of each bronchus. The name "syrinx" comes from the Greek word for the pan flute, reminding us that birds can produce sounds using two "flutes" (Suthers et al. 1994). The respiratory system generates air pressure pulses that trigger vibration of the syringeal membranes, thereby affecting the loudness, the shape of the temporal envelope, and the temporal patterns of the sounds.

The passerine syrinx is precisely controlled by two pairs of external muscles. In songbirds, there are an additional six pairs of internal muscles. The additional muscles affect the tension of the membranes as well as their location and the aperture at the tracheobronchial junction (Larsen and Goller 2002). This arrangement allows fine control of the syrinx for the generation of a rich set of sounds that vary not only in fundamental frequency but also in pitch saliency and in the shape of the spectral envelope (Sitt et al. 2008).

Finally, the spectral envelope is further shaped by active filtering of the upper vocal tract. For example, in zebra finches, although adjustment of vocal tract length does not seem to play a role in shaping spectral envelopes in song (Daley and Goller 2004), the active modification of the beak gape and of the OEC volume create resonant spectral peaks in the sound spectrum of song syllables and distance calls (Goller et al. 2004; Riede et al. 2013). Acoustic analyses revealed that these resonant spectral peaks are specific to each vocalization type (Elie and Theunissen 2016). Additionally, Simpson and Vicario (1990) described some calls as remaining "normal" following denervation of the syrinx; thus, while the production of vocalizations involves neural control of the respiratory system, the syrinx, and the upper vocal tract, a passive syrinx can also produce vocalizations as long as the respiratory system and upper vocal tract are activated appropriately.

7.3.2 Hindbrain and Midbrain Vocal Circuits

Hindbrain and midbrain circuits for vocal production are shared across all bird species (song learners and others). In particular, the muscles of the syrinx of all birds are driven by motor neurons in the tracheosyringeal part of the hypoglossal nucleus

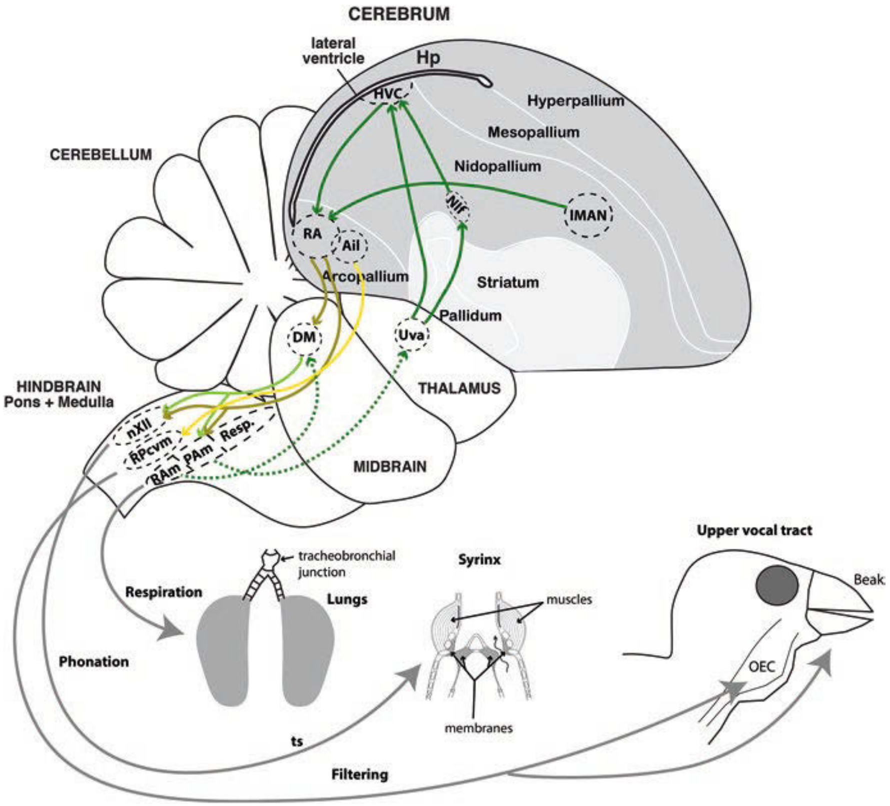


Fig. 7.2 Oscine brain circuits for the production of the vocal repertoire. Vocal production in oscines requires the activation of three systems: respiration, phonation (sound generation by the *syrinx*), and filtering by the upper vocal tract. The three systems are driven by motor neurons found in distinct hindbrain regions. Those motor nuclei receive projections from a midbrain vocal center (*DM*, *light green arrows*) and at least two distinct regions from the arcopallium (*RA* and *Ail*, *olive green* and *yellow arrows*). The schematic also shows the two principal sources of motor input to *RA* from the nidopallium, the song nuclei *HVC* and *LMAN* (*dark green arrows*). Feedback connections that are essential for song production are also shown (*dark green dotted arrows*). Note that anatomical locations are approximate (e.g., *Ail* is lateral to *RA*) The *Pallial* areas are depicted in *dark gray*, the *Pallidum* and *Striatum* in *light gray*, and the *thalamus*, the *midbrain* and *hindbrain* in *white*. See Table 7.1 for abbreviations. (anatomical outlines from Jarvis et al. 2013, relying on references described in the text)

(nXIIts; Fig. 7.2). The upper vocal tract (including the OEC and the beak) is controlled by cranial nerve motor nuclei (e.g., nV, nVII, nIX) that are jointly innervated by the ventromedial region of the parvocellular part of the reticular formation (RPcvm; Fig. 7.2) (Wild and Krützfeldt 2012). Just as in mammals, various respiration centers are found in the pons and medulla (Schmidt and Wild 2014). Of these, neurons in the nucleus paraambigualus (PAm; Fig. 7.2) and nucleus retroambigualis (RAm; Fig. 7.2) are thought to play a prominent role in the coordination of vocalization and respiration through their projections to vocal and respiratory motoneurons.

The PAm seems to play an important role in the timing of vocalizations by providing feedback information to thalamus and forebrain regions involved in call and song initiation (Schmidt and Wild 2014).

The respiratory and syringeal motor nuclei are innervated by a midbrain nucleus known as the dorsomedial nucleus (DM) of the intercollicular complex (ICo) (Wild et al. 1997). The DM seems to play a key role in orchestrating respiration and phonation and also has been postulated to control the production of all vocalizations in nonoscine birds and the production of all of the unlearned/innate calls in oscines. Small electrical stimulations of DM did indeed elicit call-like vocalizations in zebra finches (Vicario and Simpson 1995; Ashmore et al. 2008). Lesioning DM caused significant deficits in calling, such as eliminating the production of distance calls in both male and female Bengalese finches (Fukushima and Aoki 2000). Motor nuclei that drive changes in the OEC, however, do not receive direct projections from DM; they receive projections from a lateral arcopallial nucleus (Ail) through RPcvm (Wild and Krützfeldt 2012). Despite our knowledge of some of the neuroanatomical circuitry, how the three components of the vocal production system are coordinated is unknown thus far.

7.3.3 Pallial Vocal Circuits

Oscines also have evolved a collection of pallial nuclei, referred to as the song nuclei (Nottebohm et al. 1976, Kroodsma and Konishi 1991), that are known to be critical for the imitation and production of song (Sakata and Yazaki-Sugiyama, Chap. 2; Murphy, Lawley, Smith, and Prather, Chap. 3). The song nucleus RA (robust nucleus of the arcopallium) projects directly and indirectly (through DM) to nXIIts and to respiratory nuclei, including RAm and PAm. Such a direct projection from a pallial motor area to the motor hypoglossal nucleus, which controls the syrinx, has been proposed as a specialization in vocal learners that provides fine and plastic control of the vocal organ by the telencephalon (Fitch et al. 2010). The direct projection from RA to nXIIts would be analogous to the direct projection between the laryngeal premotor cortex and brainstem vocal neurons that is found in humans (and great apes) but is absent in other mammalian nonvocal learners such as monkeys (Jarvis 2013; Belyk and Brown 2017). It is noteworthy that direct projections from the pallium to motor nuclei controlling the vocal apparatus are not restricted to vocal learners: RPcvm, which controls the OEC, receives direct pallial input from the Ail in both oscines and nonoscines (Wild and Krützfeldt 2012). This direct telencephalic control of the upper vocal tract may be important for other forms of fine vocal motor control.

Two song nuclei found in the nidopallium project to RA: HVC (used as proper name) and LMAN (lateral magnocellular nucleus of the anterior nidopallium). The HVC is essential for adult song production (Nottebohm et al. 1976; Simpson and Vicario 1990); whereas, LMAN and its upstream inputs are critical for song plasticity but not critical for song production in adults (Bottjer et al. 1984; Aronov et al. 2008). Multiple experiments showed a hierarchical code for motor commands in the two highly connected pallial song nuclei HVC and RA (e.g., Vu et al. 1994; Long and Fee 2008). The HVC neurons that project to RA showed sparse and precise bursts of activity at unique times during a song motif and are thought to code the timing and sequence of song syllables (Hahnloser et al. 2002); neural activity in RA was correlated with spectrotemporal features in song syllables (Vu et al. 1994; Sober et al. 2008). Finally, the nucleus uvaeformis (Uva) is the principal thalamic input onto HVC (Nottebohm et al. 1982) and plays a role in triggering normal song production and in coordinating the feedback from midbrain respiratory, visual, and somatosensory centers (see Sect. 7.5).

7.3.4 Role of Vocal Circuits for the Production of Nonsong Vocalizations

What is the role of motor pallial brain regions in the production of other vocalization types? As mentioned above, male zebra and Bengalese finches, but not females, learn aspects of the distance call through imitation (Zann 1985; Slater and Jones 1995). Early work comparing the effects of lesions of HVC, RA, and the tracheosyringeal nerve (ts) in male and female zebra finches indicated that these regions contribute to the production of the learned aspects of the call. In particular, lesions of Uva (Coleman and Vu 2005), HVC, RA, or the ts affected the acoustic features of the distance call of males but not of females (Simpson and Vicario 1990). Simpson and Vicario (1990) suggested that two parallel pathways are activated simultaneously for distance call production in males: (1) the pathway involving the classical song nuclei that are necessary for the learned components of the vocalizations, and (2) a midbrain pathway that includes DM and controls the unlearned components of vocalizations.

However, subsequent data showed that the parallel pathways model might be an oversimplification. First, electrical stimulation of DM can generate typical distance calls in male Bengalese finches (Fukushima and Aoki 2000) and, thus, the contribution of RA might be to activate male learned motor programs stored, at least to some extent, in DM. Second, neural recordings in freely moving and naturally communicating zebra finches showed that activity in RA is associated not only with the production of learned song syllables or the distance call but also with the production of a soft contact call, the *stack* (Fig. 7.3) (Ter Maat et al. 2014). This premotor activity showed a social context dependency: neural activity in RA was higher when the stack call was produced spontaneously compared to when it was produced in response to a conspecific vocalization. Third, RA is implicated in the adaptive timing of calls during vocal exchanges of tets and stacks in both males and females (Benichov et al. 2016). Thus, the classical song production areas (HVC and RA) are clearly implicated in the plastic production of other vocalization types, controlling both the production of the learned component of the male distance call and the timing of the soft contact calls in males and females.

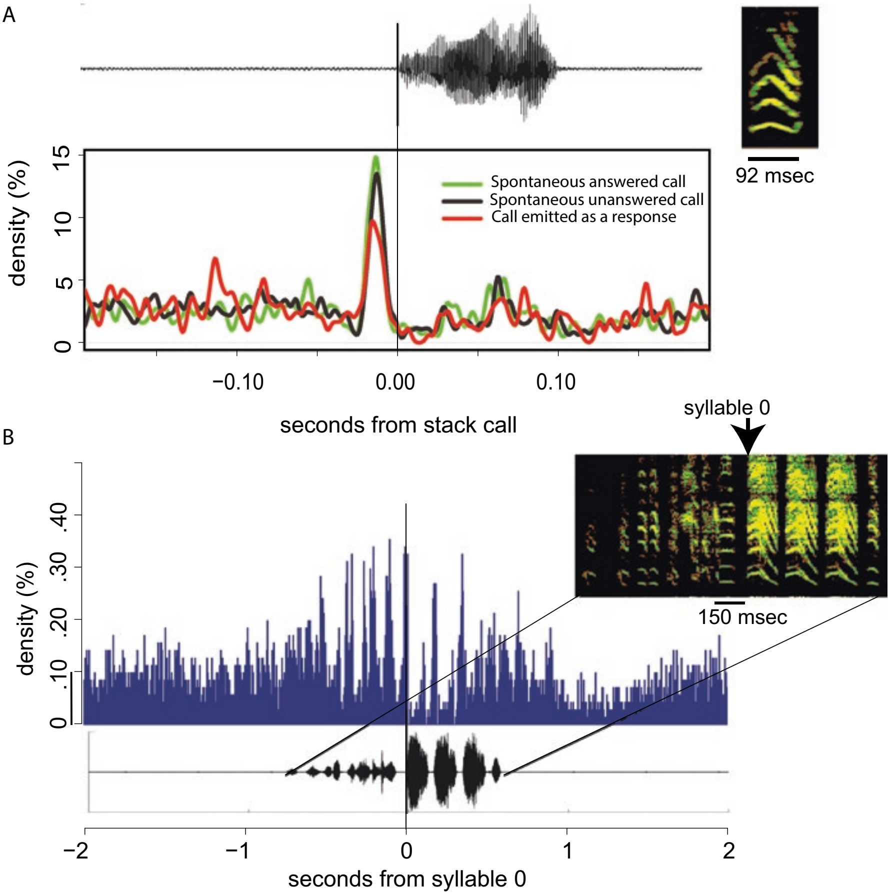


Fig. 7.3 Pre-motor activity in RA in relation to the production of *stack calls* and *song*. The average neural activity of the same single neuron recorded while a male zebra finch produced *stack calls* (A) and *song* (B). The neural activity preceding a *stack call* was depressed when the bird produced the call in response to the mate's call (*red line*) compared to when it spontaneously produced the call (*green and black* lines). In (B) the PSTH is aligned to the first of the syllables that is repeated three times in the song, found in the middle of the song motif, and labelled *syllable 0* (figure modified from Ter Maat et al. 2014)

7.3.5 Summary of Repertoire Production

Just as a binary distinction between song versus unlearned calls is an oversimplification (see Sect. 7.2), the binary distinction between a pallial brain circuit involved in learned song production and a midbrain vocal center involved in the production of all other vocalization types is almost certainly too simple. In addition to imitation of vocalizations, song nuclei also appear to be involved in controlling the timing of vocal responses and to be potentially involved in the selection of the appropriate vocalization during communicative exchanges. A more accurate model, therefore, might involve two or more sensory-motor loops that are engaged in parallel during vocal communication: a midbrain loop involving DM for the production of core vocal motor programs; a pallial loop involving the song nuclei HVC and RA for plastic control of timing and phonation (syrinx); and, hypothetically, a loop involving Ail for the plastic control of the filtering performed by the upper vocal track. Moreover, the range of complex vocal behaviors, as described in previous sections, invites additional in-depth investigations of the role of the song nuclei in other high-level vocal functions, such as the control of the number of particular syllables in the chickadee call or the imitation of alarm calls for deceptive purposes in African drongos (see Sect. 7.2.4).

In addition, worth noting is the potential role of the *anterior forebrain pathway* (AFP; includes the song nuclei LMAN and Area X) for the production or development of nonsong vocalizations, which has not been discussed. These song nuclei are crucial for the normal ontogeny of the song-copying behavior and are implicated in the social context-dependent modulation of song (Hessler and Doupe 1999; Woolley et al. 2014). As of now, there are no studies that have investigated a role for the AFP in the development or social modulation of nonsong vocalizations, for example, in the progression from long tonal call to the learned distance call in young male zebra finches or in the plastic begging calls in brood-parasite species (as described in Sect. 7.2.3).

7.4 Perception of the Repertoire of Songbirds

The perception and interpretation of a vocal communication repertoire as rich as the one observed in songbirds requires many signal processing steps that are performed by the auditory system. Some of these processing steps are needed for the processing of all sounds in the natural soundscape, while other steps are specific to the processing of communication signals. As an initial step, the perception of vocalizations requires an *auditory scene analysis*: vocalizations must be distinguished from background sounds and recognized as specific signals, even if they have been degraded by propagation through the natural environment. In addition, signals from multiple birds vocalizing simultaneously might have to be segregated into distinct signals.

Vocal communication also requires a *sound to meaning transformation* that involves a *categorization task* and a *memory task*. The categorization task is required to group sounds into behaviorally meaningful units, for example, distinguishing conspecific from heterospecific vocalizations or distinguishing the different vocalization types within the conspecific repertoire. The memory task involves the formation and the recall of memories for auditory objects. These auditory memories are clearly essential for behaviors that require the recognition of specific individuals (e.g., mates, friends, or foes). Auditory memories could also shape preferences for conspecifics, for example, auditory memories of song formed during development that shape female mate choice (see Woolley and Woolley, Chap. 5). As discussed previously, juvenile male songbirds also rely on auditory memories of a tutor's song for song imitation. Finally, the categorization and memory tasks can be interrelated as there could be both innate and learned categories: the recognition of song dialects or of alarm calls that become referential through experience are two examples of categories that require auditory learning and memory formation.

7.4.1 Auditory Hindbrain and Midbrain

The oscine ascending auditory system is very similar in anatomical location and in the number of processing stages between the inner ear and the primary pallial/cortical auditory areas to what has been described in other vertebrates (Fig. 7.4). Afferents from auditory hair cells project onto cells in two second-order nuclei in the medulla: nucleus angularis (NA) and nucleus magnocellularis (NM). These two avian nuclei specialize in processing amplitude/spectral information (NA) and timing/frequency information (NM). Neurons from NM project principally and bilaterally to the third-order auditory nucleus in the medulla, the nucleus laminaris (NL). These early processing stages incorporate the initial computations needed for auditory scene analysis (Konishi 2003) and for the extraction of acoustic features that are fundamental to an efficient neural representation of all vocalizations.

As in mammals, there are both direct and indirect projections from NA and NL to the midbrain auditory center, the dorsolateral nucleus of the mesencephalon (MLd; known as the inferior colliculus in mammals). The indirect pathways innervate the superior olive (SO) and the lateral lemniscal (LL) nuclei. The LL pathway plays a role in estimating interaural level differences. Contrary to what has been observed in barn owls, which have excellent sound localization abilities, the pathways that specialize in spatial auditory processing (NL, SO, LL) and in amplitude/spectral processing (NA) are not segregated in the oscine MLd (Krützfeldt et al. 2010). Krützfeldt et al. (2010) suggest that this difference between the two avian clades reflects the difference in relative importance of the *where pathway* (more important for the barn owl) and the *what pathway* (more important for oscines). This anatomical difference, therefore, could mark the first auditory specialization for the processing of vocalizations. Neural recordings in the MLd have shown that single neurons have complex response properties beyond frequency tuning that include tuning to temporal patterns, which is useful for extracting the onset and offset of sound elements. On the one hand, these complex responses are relevant for the representation of specific spectrotemporal features or temporal sequences found in natural vocalizations (Woolley and Portfors 2013). On the other hand, neurons in the oscine MLd also responded well to synthetic sounds with spectrotemporal acoustic structures that match those found in natural sounds (Hsu et al. 2004). Thus,

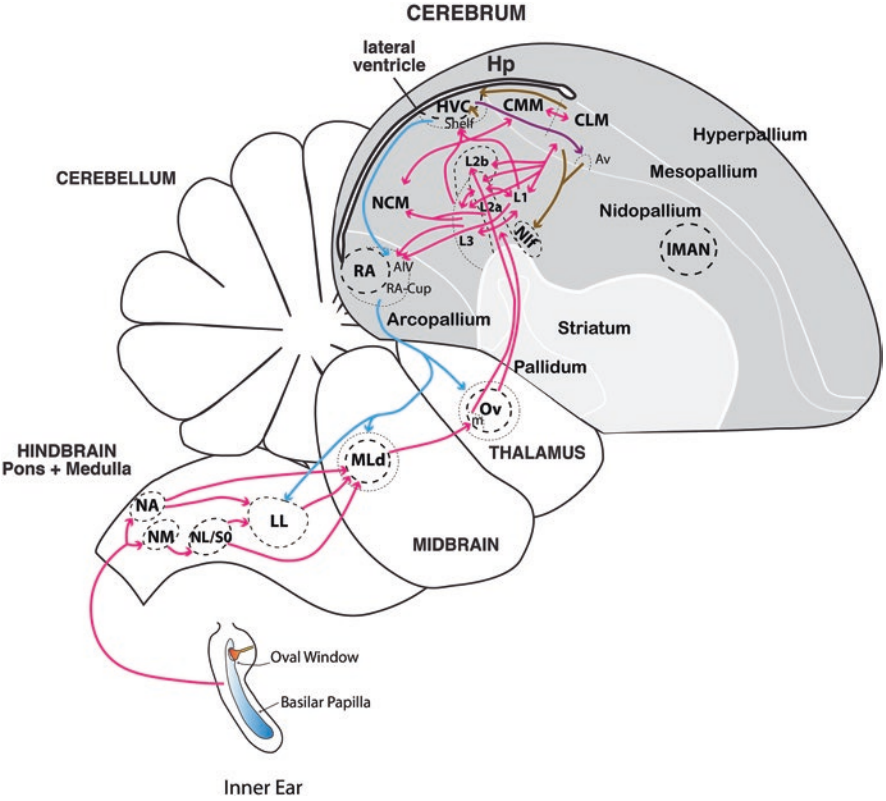


Fig. 7.4 Oscine brain circuits for the perception of the vocal repertoire. The brain regions and connections of the avian auditory system are illustrated. The feedforward and "lateral" connections of the ascending auditory stream are shown in *pink*. The descending auditory pathways are shown in *blue*, and the known connections between auditory areas and song nuclei *HVC* and *NIf* in *brown*. Locations are approximate and the reader should be aware that the medial/lateral dimension is not represented. *L1*, *L2a*, *L2b*, *L3* are the primary auditory areas in the *nidopallium*, and together with the *CLM* constitute the avian primary auditory "cortex." *CMM* and *NCM* are secondary pallial auditory areas. *Av* is embedded in *CLM*, receives a projection from the song nucleus *HVC*, and is shown in *purple* to indicate that it could combine sensory and motor information. The ventral portion of the intermediate acropallium (*AIV*) overlaps with *RA-cup*. The *dashed lines* around *Ov* and *MLd* are the shell regions of these nuclei that receive descending input from *RA-cup*. Neurons in a medial part of the *Ov* complex (*m*) project to L2b. (all abbreviations appear in Table 7.1)

one might conclude that MLd neurons demonstrate tuning for low-level features found in vocalizations (and other natural sounds) but do not exhibit response properties that reflect categorization of vocalization types or memories for auditory objects. However, this assertion needs further investigation with experiments designed to directly test neural invariance and plasticity.

7.4.2 Auditory Thalamus

Neurons in the MLd project to the thalamic auditory nucleus, the nucleus ovoidalis (Ov), which is analogous to the mammalian auditory thalamus, the medial geniculate body. Neurons in Ov, in turn, project to the primary auditory pallium. Although the thalamic sensory nuclei are often thought of as simple relay nuclei, the spectrotemporal tuning of neurons in Ov are more similar in their complexity to the tuning observed in the pallium than to the tuning of the MLd (Amin et al. 2010). Relative to MLd neurons, neurons in Ov show greater tuning for intermediate spectral and temporal modulations: they are able to detect changes in the spectral patterns of sounds over longer periods of time than MLd neurons. Thus, it is clear that Ov plays an additional role in the generation of complex responses that could be beneficial for the processing of vocalizations. The shelf of Ov is also the recipient of the first auditory feedback signal from the pallial regions, but the role of this auditory feedback has not yet been investigated (Mello et al. 1998).

7.4.3 Auditory Pallium

Auditory information from Ov enters the oscine auditory nidopallium through the thalamic recipient auditory areas in L2. Two separate projections have been identified: the core of Ov projects to L2a and a medial region within the Ovoidalis complex projects to L2b. Neurons in L2 project to more dorsal/rostral and ventral/caudal regions of the nidopallium, L1 and L3, respectively, as well as to the caudal lateral region of the mesopallium (CLM) (Vates et al. 1996). This anatomical organization is similar to the one found in the mammalian neocortex, where L2 would be the equivalent of the thalamic recipient layer 4 neurons in primary auditory cortex (A1), and L1, L3, and CLM correspond to other cortical layers (Wang et al. 2010). The avian Field L (L1, L2, L3) and CLM would be analogous to the primary auditory cortex, and thus could be called the Field L/CLM complex. This analogy between the avian primary auditory pallium and the mammalian auditory cortex has important limitations; for example, contrary to what is found in mammals, in oscines there are no direct feedback projections from the Field L/CLM complex to the auditory thalamus nor to the contralateral auditory pallium (Wild and Krützfeldt 2010). Neurons in primary auditory fieldL/CLM complex exhibit heterogenous tuning response properties as assessed by their spectrotemporal receptive fields (STRFs). Similar to receptive fields in vision, STRFs are the filters in the spectrotemporal domain that extract the acoustic properties of the sound to which a given neuron is responding. These tuning properties in the primary auditory Field L/CLM complex can be clustered in 4–5 response types that are efficient at representing distinct perceptual qualities (e.g., rhythm, timbre, pitch) of natural sounds, including vocalizations (Woolley et al. 2009).

The deeper (L3) and superficial (L1 and CLM) areas of the primary auditory pallium project to two secondary auditory areas: the caudomedial nidopallium (NCM) and the caudomedial mesopallium (CMM). Both NCM and CMM are reciprocally connected to each other (Vates et al. 1996) and exhibit multiple high-level response properties that are relevant for the processing of a vocal repertoire in natural conditions (see also Woolley and Woolley, Chap. 5).

7.4.4 Auditory Scene Analysis

The auditory pallium, in particular the secondary auditory area NCM, is involved in extracting vocalizations from noise. One key discovery was of NCM neurons that are responsive to song stimuli but are invariant to background noise (Wang et al. 2007; Schneider and Woolley 2013). Some of this selectivity for signal over noise can be explained by tuning for sharp but long spectral structures that are found primarily in vocalizations (e.g., harmonic stacks) (Moore et al. 2013). Secondary auditory areas like NCM also showed neural correlates of sound source separation when multiple sounds were present (Maddox et al. 2012), making them excellent candidates for detecting and extracting a vocalization signal in natural conditions. Indeed, neurons in CMM, CLM, and NCM encode the vocalizer identity of naturally propagated distance calls even in very low signal-to-noise ratios; these neurons also show a high degree of signal sound-level invariance under constant background noise levels (Mouterde et al. 2017). Thus, the oscine secondary auditory pallium clearly performs auditory scene analysis tasks that are critical for the detection of vocalizations in the natural environment.

7.4.5 Categories for Auditory Objects

The interpretation of vocal signals requires categorization of vocalizations into meaningful units. Categorization of vocal signals also requires neural response invariance but, for this categorization task, responses should be invariant to different renditions of signals grouped according to their unique behavioral significance. Neurons in the auditory pallial areas L3, NCM, CMM, and CLM exhibit such categorical responses. Starlings (*Sturnus vulgaris*) use the characteristic song types of different individuals for individual recognition (Gentner and Hulse 1998; Gentner 2004). Neurons in the auditory pallium of starlings produced similar responses to variants of song types and, by doing so, categorized song variants into their respective song types (Meliza and Margoliash 2012). Similarly, the only study that examined the neural representation of all vocalization types in the repertoire of an oscine, the zebra finch, found categorical neurons that yielded invariant responses to different renditions of the same vocalization type (Fig. 7.5A) (Elie and Theunissen 2015). Both in European starlings and zebra finches, neural selectivity and invariance are positively correlated; thus, categorical neurons tended to be selective for one or two vocalization or song types.

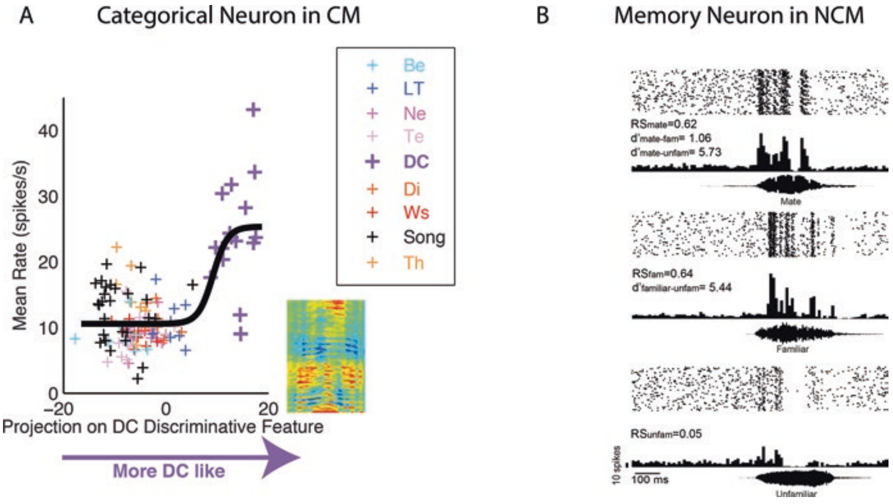


Fig. 7.5 Illustrative examples of neural correlates of auditory memories and categorization. (A) Mean firing rates of a neuron in CM presented with a dozen of renditions from different birds for nine vocalization types (*Be*, begging; *LT*, long tonal; *Ne*, nest; *Te*, tet; *DC*, distance call; *Di*, distress call; *Ws*, Wsst or alarm; *Th*, Thuk). Each "+" corresponds to one stimulus vocalization, and the color codes the vocalization type. The mean firing rate (*y-axis*) is plotted against an acoustical measure of "similarity to distance call" for each vocalization used as a stimulus (*x-axis*). The acoustical measure is obtained by projecting the spectrogram of each vocalization on the weights that one obtains in a logistic regression fitted to separate distance calls from all other vocalizations. (B) Raster plots and spike histograms of a single unit in NCM presented with Distance calls from familiar (*mate* and *fam*) or unfamiliar birds (*unfam*). The response strength index (RS) is the firing rate during the stimulus minus the background, normalized by the sum of these two firing rates. The psychophysical measure of effect size, d', is used to quantify differences in responses obtained for two different stimuli (figures adapted from Elie and Theunissen 2015; Menardy et al. 2012)

Categorical auditory neurons are also found for learned auditory objects. European starlings can learn to categorize songs from multiple individuals according to the reward value associated with these songs in an operant conditioning task. Neurons in CLM and CMM showed categorical responses, both at the single neuron level and at the population level, for learned categories that were only relevant for the behavioral task (Gentner and Margoliash 2003; Jeanne et al. 2011).

7.4.6 Memories for Auditory Objects

The formation of memories for auditory objects in oscines is important for the interpretation of vocalizations, in particular for individual recognition, for the flexible usage of vocalizations in some species (e.g., the African drongo's use of other species' alarm calls), and, of course, for the imitation of vocalizations such as the song and the distance call in zebra finches. There are multiple lines of evidence that implicate neurons in NCM, CMM, or CLM in auditory memories. On the one hand, neurons in CMM and CLM have been implicated in the storage of specific complex vocalizations and higher-level auditory objects with higher levels of information on these learned object categories in CMM. Interestingly though, neural correlates for memories in these CM regions have so far only been demonstrated for rule-based categories in operant conditioning tasks in which European starlings had to memorize arbitrary groups of songs from different individuals according to the reward value associated with those songs (Gentner and Margoliash 2003; Jeanne et al. 2011). On the other hand, neurons in NCM show selective responses that could correspond to auditory memories relevant for natural behaviors. In particular, NCM has been identified as a potential locus of the tutor song memory based on selective immediate-early-gene expression (Gobes et al. 2010) or adaptation properties of single neurons (Phan et al. 2006). Although these experiments clearly show that familiarity with the tutor song changes neurophysiological properties, how these physiological changes could be used for memory recall is not yet understood. More convincingly, a small population of NCM neurons was discovered that was selectively responsive to tutor song syllables or to the bird's own song in tutored birds but not in untutored birds (Yanagihara and Yazaki-Sugiyama 2016).

None of the studies listed in the previous paragraph rule out the possibility that NCM could be a general purpose memory-storage area for communication signals, potentially in both male and female birds. For example, in the study by Yanagihara and Yazaki-Sugiyama (2016), exposure to tutor song also increased the population of NCM neurons that selectively responded to other communication calls, such as the female distance call. Further evidence that NCM might store general auditory memories comes from studies that examine the neural representation of familiar versus unfamiliar distance calls in both male and female zebra finches. Auditory responses to familiar calls in both sexes were larger than those to unfamiliar calls (Fig. 7.5B) (Menardy et al. 2012).

In summary, both CM and NCM show neural correlates of memories for auditory objects. Neurons in NCM show selective auditory responses for familiar vocalizations that could be used for identifying a mate or for storing a sensory copy of the song to imitate. CMM shows selective responses for the perception of higher order objects such as rapidly learned novel sound categories. Although this functional distinction is appealing, the reader should remember that NCM and CMM on the one hand, and CMM and CLM on the other hand, are interconnected and that none of the studies have contrasted the role of CMM/L and NCM in the same memory task.

7.4.7 Summary of Repertoire Perception

Neurophysiological recordings in the auditory system of oscines have revealed neural correlates of three interconnected high-level auditory functions that are required for the processing of vocalizations: auditory scene analysis, auditory object categorization, and auditory memory formation (e.g., memory for tutor song or vocalizations of others). In these auditory areas, it remains unclear whether the circuits for storing the tutor song are different from those used for storing the song and nonsong vocalizations of other familiar conspecifics. Experiments that investigate the necessity of particular circuits for specific vocal communication behaviors, and thus test causal links, are very much needed.

7.5 The Sensory-Motor Integration of Vocal Signals in Songbirds

The complex vocal behavior in oscines requires significant integration between perceptual systems and the motor system. Once vocalizations are recognized and interpreted by the auditory system, the appropriate behavioral response must be elicited. These responses include vocal behaviors, such as call-backs and singing, as well as other behaviors, such as approach, avoidance, copulatory display, or feeding. Vocal production can also be triggered by nonauditory stimuli, including visual stimulation (e.g., song upon sight of a female), stress (distress call when chased by a conspecific), or social isolation (distance call of an isolated bird trying to localize group members). Because of the extensive amount of research into the process of song imitation, a relatively detailed but complex picture of the sensory-motor integration involved in song copying can be provided; however, the putative sensory-motor circuits that could be involved in all aspects of vocal communication and behaviors can only be very coarsely sketched. In this section, we explore the feedback loops engaged for vocal communication, from the perception of a signal to the production of a response, and highlight unexplored questions.

7.5.1 Auditory to Auditory Feedback Systems

Auditory to auditory feedback systems might be required for enhancing the selective processing of vocalizations once they have been categorized by the circuits in the auditory pallium, for instance, for sustaining attention. Auditory feedback from the auditory pallium to the auditory thalamus and auditory midbrain is indirect. Neurons in L1 and L3 project to a nidopallial region just ventral to HVC, namely HVC-shelf (Fig. 7.4). The HVC-shelf, in turn, projects to regions of the arcopallium ventral and rostral to RA called the RA-cup (Mello et al. 1998) and the intermediate arcopallium (AIV) (Figs. 7.4 and 7.6) (Mandelblat-Cerf et al. 2014). These two areas were defined by auditory anterograde and dopaminergic retrograde labeling, respectively. Since they overlap extensively, RA-cup/AIV will be used in this chapter.

The RA-cup/AIV region projects to areas surrounding the auditory thalamus, Ov, and auditory midbrain, MLd. In the midbrain, RA-cup/AIV projecting neurons

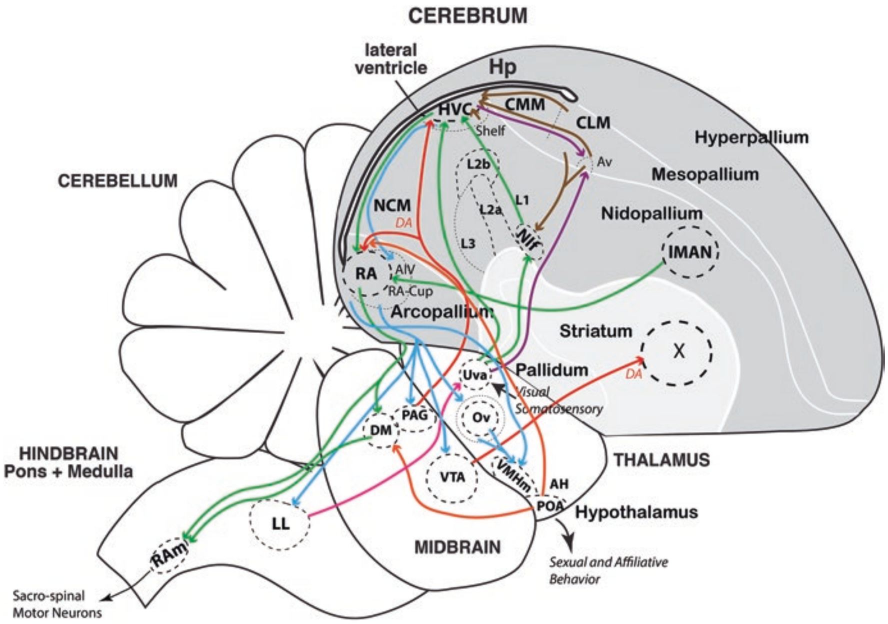


Fig. 7.6 Oscine brain circuits for sensory-motor integration. The brain regions and connections that produce appropriate behaviors in response to vocalizations are illustrated (*blue arrows*, descending auditory pathways; *brown arrows*, auditory to motor connections; *green arrows*, motor to motor connections); also shown are the brain regions and connections that trigger vocalizations in response to appropriate stimuli (*orange and green arrows*, pathways from the social brain network to motor pathways), including other vocalizations (motor pathways triggered by the ascending auditory connections, *pink and green arrows*; triggered by the secondary auditory regions, *brown and green arrows*). The motor to auditory pathways are indicated in *purple*. The dopaminergic pathways are shown in *red* and involve the *VTA* and the *PAG*. The social brain network involves multiple areas found in the preoptic area and hypothalamus: *POA*, *VMHm*, *AH*. Area X (*X*) in the striatum is one of the song nuclei in the *AFP*. (see Figs. 7.2 and 7.4 for other brain structures and Table 7.1 for other abbreviations)

innervate the ICo but not the vocal nucleus DM (Mello et al. 1998). These auditory feedback connections are thought to affect the ascending processing in Ov and MLd, potentially after being integrated with signals from other modalities or other modulatory systems (Durand et al. 1992; Wild 2017). The actual function of these auditory feedback pathways remains unknown (see Figs. 7.4, 7.6).

7.5.2 *Auditory Signals to Vocal Reponses*

The integration of the auditory system with the vocal system can be divided into two processes: (1) the input of the auditory system to the song nuclei for song copying and (2) the input of the auditory system to the song and vocal centers for eliciting song and calls (although the two might overlap). How auditory input, as a guiding signal, enters the song system for vocal copying is a well-researched and complex topic that will be very briefly summarized here. Auditory information is known to affect responses in song nuclei both directly and indirectly, for example, via the dopaminergic system. Some of the connections between the auditory system and the song nuclei are between L1/L3 and HVC through the HVC-shelf (Vates et al. 1996), between CLM and HVC (Shaevitz and Theunissen 2007; Bauer et al. 2008), between CM and NIf (nucleus interfacialis of the nidopalium) (Lewandowski et al. 2013) (Fig. 7.4), and, finally, between the auditory LL and Uva (Coleman et al. 2007) (Fig. 7.6). A small ventral region in CLM known as Avalanche (Av) seems to play a key role in the loops that link auditory areas and the song nuclei. Avalanche not only projects to NIf and HVC, providing auditory input to the song system, but Av also receives input from HVC and Uva (Akutagawa and Konishi 2010). Thus, there are reciprocal feedback connections between high-level auditory areas and the song nuclei, some of which are critical for normal song copying behavior in young birds (Roberts et al. 2017).

In addition, auditory input could indirectly affect processing in song nuclei through two midbrain dopaminergic centers, the ventral tegmental area (VTA) (Hoffmann et al. 2016; Hisey et al. 2018) and the avian periaqueductal gray (PAG) (Tanaka et al., 2018; Hamaguchi and Mooney 2012). The VTA and the PAG receive inputs from the RA-cup/AIV, areas that receive auditory inputs from the HVC-shelf (Fig. 7.6), L1, and L3 (Fig. 7.4). Normal song copying requires both an intact RA-cup/AIV (Mandelblat-Cerf et al. 2014) and a functional dopaminergic VTA projection to Area X in the AFP (Xiao et al. 2018; Hisey et al. 2018). The other dopaminergic input to the song nuclei comes from PAG, which projects to both RA and HVC (Appeltants et al. 2002). The projection from PAG to HVC also is important in determining the storage of the motor program for the copied song and might principally gate the auditory input using social reward signals (Tanaka et al. 2018). Thus, the high-level auditory areas that exhibit neural correlates of auditory memories provide both direct sensory input to the song nuclei and indirect input via the dopaminergic system.

This wealth of research allows researchers studying the neurobiology of song imitation to formulate more specific hypotheses but also reveals the complexity of the system (see Sakata and Yazaki-Sugiyama, Chap. 2). To give a broad picture, one could hypothesize that the auditory memory of the tutor song is stored in NCM or Av during the initial sensory-motor phase of the song imitation behavior. This auditory memory, triggered by the appropriate auditory feedback, would then be used during the sensory-motor phase to gate the activity in HVC (and/or Uva, PAG) and also selectively stimulate RA-cup/AIV and the downstream dopaminergic center, VTA. The gating and dopaminergic signals would then be used to learn the appropriate motor program for that target song. This motor program would be stored in HVC and RA. This is a scenario that is consistent with the neural mechanisms implicated in song copying that have been described so far and that were briefly reviewed here, but other comprehensive working hypotheses could be formulated. The details of these working hypotheses would have to be specified and tested with both additional experiments and computational models that could generate the observed neural responses and behavioral output.

How other learned vocalizations, such as the male distance call in the zebra finch, are processed by these same sensory-motor loops or sensory-dopaminergic-motor loops is unknown. Furthermore, other social calls (e.g., all contact calls in the zebra finch) elicit call-back behaviors that can be individual specific and context specific, and they could require the recall of auditory memories. As mentioned previously, RA could have a role in the modulation of the timing of calls. Since RA receives input from HVC, auditory memories for contact calls from individual vocalizers potentially stored in NCM could also trigger these specific responses in RA via HVC. Also of interest would be further investigations of the putative auditory input to Ail both for song and nonsong vocalization production. Finally, nonsong vocalizations can be directly elicited by activation of DM independently of RA, but auditory input to DM, excluding the projection from RA (Wild 1994a), has not been well characterized. Auditory input to DM could come from the auditory midbrain nucleus MLd or from the auditory pallial regions that project to RA-cup/AIV, and from there to ICo, and finally to DM. Recording auditory activity in response to calls or songs perception in DM before and after inactivation of RA would help to elucidate the role of DM in the sensory-motor loop. The presence of auditory-evoked activity in DM would further support the role of DM in controlling call-back behavior for nonsong vocalizations; on the other hand, if auditory-evoked activity is not found in DM, the role of RA in controlling the timing of all vocalizations would be strengthened.

7.5.3 Auditory Signals to Nonvocal Social Responses

In most oscines, vocal communication plays an important role in many social behaviors. Therefore, one expects to observe significant bidirectional connectivity between high-level auditory areas, the social behavior network, and the dopaminergic reward systems. The social behavior network shares many similarities across all vertebrates and includes brain areas such as the extended amygdala, the preoptic area (POA), the anterior hypothalamus (AH), the lateral septum (LS), and the ventromedial hypothalamus (VMH) (Goodson 2005). The networks that link the auditory pallium to these brain regions are not well known but are important for understanding mechanisms underlying social and vocal interactions. The only documented connections between the auditory pallium and the social behavior network are projections to the medial part of the ventromedial hypothalamic nucleus (VMHm), a brain region implicated in the control of reproductive behavior (Wild 2017). The ventral region of NCM, the medial arcopallium, Ov, and Ov-shelf all project to VMHm. How Ov, a key auditory center (see Sect. 7.4.2), could be involved in detecting attractive songs or calls and in triggering appropriate affiliative or sexual behaviors is unknown. Similarly, neural responses in NCM are selective for the vocalizations of familiar individuals (see 7.4.6), but how this neural activity relates to social behaviors or song learning remains unclear.

Multiple lines of research also implicate RA and RA-cup/AIV for triggering social behaviors. First, RA-cup/AIV is the principal hub between the descending auditory pathways and the dopaminergic system (VTA and PAG). Because all affiliative vocalizations trigger approach behaviors, the auditory perception of such signals could potentially affect the activity of the VTA and the PAG. The activation of VTA or PAG via RA-cup/AIV might be useful, therefore, not only for the reinforcement of song copying but also for other social behaviors such as mate bonding. The PAG in mammals controls diverse social behaviors, including approach and sexual behaviors and fight-or-flight behaviors (Sewards and Sewards 2003). Studies have shown that the avian PAG and ICo are analogous to the mammalian PAG, albeit with a slightly different anatomical organization (Kingsbury et al. 2011). Thus, vocal signals that arrive in these areas from RA-cup/AIV could potentially trigger a range of social behaviors, including escape upon hearing an alarm or aggressive call. RA has also directly been implicated in triggering social behavior. The projection from RA to RAm controls not only respiratory motor neurons but also sacrospinal motor neurons needed for copulatory behaviors in males and females (Wild and Botelho 2015). In female zebra finches, projections from CM to RA and to the caudal striatum could also be involved in mediating the preference for attractive familiar songs (Dunning et al. 2018). Therefore, RA and RA-cup/AIV could be involved in both males and females for the activation of multiple social behaviors (song and call responses, approach, bonding, or copulation) after being triggered by high-level auditory areas that recognized the social signal (e.g., a mate's song or contact call).

7.5.4 Nonauditory Social Signal to Vocal Responses

Nonvocal social cues can also elicit song and calls. For example, a male zebra finch will produce a directed song upon seeing a desired female zebra finch, and this social modulation of the type of song produced is mediated, in part, by the VTA (Hara et al. 2007). The thalamic song nucleus Uva relays visual and somatosensory information to HVC and NIf (Wild 1994b), suggesting that Uva could also be important for eliciting the production of courtship song from visual stimuli. Nuclei from the social behavior network are integrators of perceived social behaviors and are also good candidates for triggering the vocal motor program in response to nonvocal social stimuli. In starlings, the medial preoptic nucleus (POM), a region known to be involved in sexual motivation, projects to DM and to a region in the arcopallium that borders the dorsomedial boundary of RA (Riters and Alger 2004). Infusions of testosterone into POM of castrated canaries (*Senirius canaria*) increase the production of song and, indirectly, the size of song nuclei (Alward et al. 2013). In territorial oscines, AH is known as an integrator of neural processes related to aggression and defense (Goodson et al. 2012). How the AH might trigger the production of territorial songs or aggressive calls is currently unknown. More generally, the specific pathways between the social behavior network and the circuit for vocal production require further investigation in both males and females to understand how nonauditory social stimuli can elicit appropriate vocal responses.

7.5.5 Summary of Sensory-Motor Integration

With the exception of the song-copying behavior, little is known about the sensory-motor integration of vocalizations in the larger context of oscine communication. Based on the studies to date, it appears that arcopallial regions like RA, RA-cup/AIV, and other regions bordering RA play a key role in linking the outcome of auditory processing (recognition of vocalizations, of signaler identity, and of behavioral state), which is putatively occurring in the auditory nidopallium and mesopallium, with appropriate social behaviors. Thus, just as for call production and perception, it appears that the circuits involved in the song-copying behavior, in particular in the arcopallium, are also involved in the sensory-motor integration required for eliciting appropriate social behaviors in response to all vocalizations, including conspecific song. Going forward, experiments that assess how manipulations of activity in the arcopallium affect not only song production but also behavioral responses to a range of vocalizations will be important.

7.6 Chapter Summary

This chapter highlights the power and complexity of a neuroethological approach for studying a complete vocal communication system in oscines. The oscine vocal communication system is a flexible language whose complexity correlates with the social complexity observed in an animal's behavior. The complexity of oscine vocal communication rivals that of any other social animal, including nonhuman primates. Vocal communication is the glue of the oscine social networks. Studying the production and perception of all vocalizations provides a unique window into the poorly explored field of social neuroscience. The vocal complexity of oscine vocal communication goes well beyond what can be very explicitly heard by all humans when listening to the oscine song, which inspired many musicians and naturalists well before inspiring modern researchers.

There has been extensive scientific work revealing the neural basis of the song-copying behavior in oscines, and this trove of research is very relevant to the perception and processing of all vocalizations. The neural plasticity that has been implicated for the formation of the tutor song template, and its use for guiding song imitation, might be relevant for other forms of vocal plasticity in the production, use, and perception of other vocalizations. In this respect, the neuroethology of vocal communication in oscines clearly benefits from investigations into the neurobiology of song copying. Similarly, the neurobiology of song copying will benefit from experiments that take advantage of the complexity and diversity of the complete repertoire of vocal communication behaviors. A short list of some of these experiments is provided here. First, a comparison of the neural mechanisms for copying the song versus the distance call in male zebra finches could provide key information on circuits that are solely implicated in learning sequences (found only in song) versus those that are involved in spectrotemporal features (found in song and distance call). Second, circuits that mediate perceptual memories for song can be compared to those involved in memories for nonsong vocalizations and, in oscines (whose song behavior is sexually dimorphic), these circuits can be compared between females and males. Third, circuits that mediate vocal plasticity for timing of song syllables could be compared to circuits that mediate vocal plasticity in call-back responses. Fourth, developmental studies could compare the maturation of different vocalization types and their neural underpinning. More generally, another useful strategy would be to leverage an evolutionary comparative approach: comparisons can be made not only across vocalization types and sexes within the same species but also across species that vary in song repertoire complexity, in their flexibility in the use of vocalization types, and in their abilities for imitation and for individual recognition.

Acknowledgements The authors thank Profs. Martin Wild, Marc Schmidt, Sarah Woolley, and Jon Sakata for their insightful comments and feedback on draft versions of this chapter.

Compliance with Ethics Requirements Julie Elie declares that she has no conflict of interest.

Frédéric Theunissen declares that he has no conflict of interest.

References

Akutagawa E, Konishi M (2010) New brain pathways found in the vocal control system of a songbird. J Comp Neurol 518(15):3086–3100

Alward BA, Balthazart J, Ball GF (2013) Differential effects of global versus local testosterone on singing behavior and its underlying neural substrate. Proc Natl Acad Sci US A 110(48):19573–19578

Amin N, Gill P, Theunissen FE (2010) Role of the zebra finch auditory thalamus in generating complex representations for natural sounds. J Neurophysiol 104(2):784–798

Appeltants D, Ball G, Balthazart J (2002) The origin of catecholaminergic inputs to the song control nucleus RA in canaries. Neuroreport 13:649–653

Aronov D, Andalman AS, Fee MS (2008) A specialized forebrain circuit for vocal babbling in the juvenile songbird. Science 320(5876):630–634

Ashmore RC, Renk JA, Schmidt MF (2008) Bottom-up activation of the vocal motor forebrain by the respiratory brainstem. J Neurosci 28(10):2613–2623

Baker MC, Bjerke TK, Lampe HU et al (1987) Sexual-response of female yellowhammers to differences in regional song dialects and repertoire sizes. Anim Behav 35:395–401

Bauer EE, Coleman MJ, Roberts TF et al (2008) A synaptic basis for auditory-vocal integration in the songbird. J Neurosci 28(6):1509–1522

Beecher MD, Brenowitz EA (2005) Functional aspects of song learning in songbirds. Trends Ecol Evol 20(3):143–149

Belyk M, Brown S (2017) The origins of the vocal brain in humans. Neuroscience and Biobehav Rev 77:177–193

Benichov JI, Benezra SE, Vallentin D et al (2016) The forebrain song system mediates predictive call timing in female and male zebra finches. Curr Biol 26(3):309–318

Bottjer SW, Miesner EA, Arnold AP (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. Science 224(4651):901–903

Brenowitz EA, Beecher MD (2005) Song learning in birds: diversity and plasticity opportunities and challenges. Trends Neurosci 28(3):127–132

Coleman MJ, Vu ET (2005) Recovery of impaired songs following unilateral but not bilateral lesions of nucleus *uvaeformis* of adult zebra finches. J Neurobiol 63(1):70–89

Coleman MJ, Roy A, Wild JM et al (2007) Thalamic gating of auditory responses in telencephalic song control nuclei. J Neurosci 27(37):10024–10036

D'Amelio PB, Klumb M, Adreani MN et al (2017) Individual recognition of opposite sex vocalizations in the zebra finch. Sci Rep 7:5579

Daley M, Goller F (2004) Tracheal length changes during zebra finch song and their possible role in upper vocal tract filtering. J Neurobiol 59(3):319–330

Dalziell AH, Magrath RD (2012) Fooling the experts: accurate vocal mimicry in the song of the superb lyrebird *Menura novaehollandiae*. Anim Behav 83(6):1401–1410

Dunning JL, Maze SE, Atwood EJ et al (2018) Caudal mesopallial neurons in female songbirds bridge sensory and motor brain regions. J Comp Neurol 526(10):1703–1711

Durand SE, Tepper JM, Cheng MF (1992) The shell region of the nucleus *ovoidalis*: a subdivision of the avian auditory thalamus. J Comp Neurol 323(4):495–518

Düring DN, Elemans CPH (2016) Embodied motor control of avian vocal production In: Suthers R Fitch W Fay R Popper A (eds) Vertebrate Sound Production and Acoustic Communication Springer Handbook of Auditory Research vol 53

Elie JE, Theunissen FE (2015) Meaning in the avian auditory cortex: neural representation of communication calls. Eur J Neurosci 41(5):546–567

Elie JE, Theunissen FE (2016) The vocal repertoire of the domesticated zebra finch: a data-driven approach to decipher the information-bearing acoustic features of communication signals. Anim Cogn 19(2):285–315

Elie JE, Theunissen FE (2018) Zebra finches identify individuals using vocal signatures unique to each call type. Nat Commun 9(1):4026

Elie JE, Soula HA, Mathevon N et al (2011) Dynamics of communal vocalizations in a social songbird the zebra finch (*Taeniopygia guttata*). J Acoust Sco Am 129(6):4037–4046

Engesser S, Crane JM, Savage JL et al (2015) Experimental evidence for phonemic contrasts in a nonhuman vocal system. PLoS Biol 13(6):e1002171

Engesser S, Ridley AR, Townsend SW (2016) Meaningful call combinations and compositional processing in the southern pied babbler. Proc Natl Acad Sci U S A 113(21):5976–5981

Fitch WT, Huber L, Bugnyar T (2010) Social cognition and the evolution of language: constructing cognitive phylogenies. Neuron 65(6):795–814

Flower TP, Gribble M, Ridley AR (2014) Deception by flexible alarm mimicry in an african bird. Science 344(6183):513–516

Freeberg TM (2008) Complexity in the chick-a-Dee call of carolina chikcadees. Auk 125(4):896–907

Fukushima Y, Aoki K (2000) The role of the dorsomedial nucleus (DM) of *intercollicular* complex with regard to sexual difference of distance calls in Bengalese finches. Zool Sci 17(9):1231–1238

Gammon DE, Altizer CE (2011) Northern mockingbirds produce syntactical patterns of vocal mimicry that reflect taxonomy of imitated species. J Field Ornithol 82(2):158–164

Gentner TQ (2004) Neural systems for individual song recognition in adult birds. Ann N Y Acad Sci 1016:282–302

Gentner TQ, Hulse SH (1998) Perceptual mechanisms for individual vocal recognition in European starlings *Sturnus vulgaris*. Anim Behav 56(3):579–594

Gentner TQ, Margoliash D (2003) Neuronal populations and single cells representing learned auditory objects. Nature 424(6949):669–674

Gill SA, Bierema AM, Hauber M (2013) On the meaning of alarm calls: a review of functional reference in avian alarm calling. Ethology 119(6):449–461

Gobes SM, Bolhuis JJ (2007) Birdsong memory: a neural dissociation between song recognition and production. Curr Biol 17(9):789–793

Gobes SM, Zandbergen MA, Bolhuis JJ (2010) Memory in the making: localized brain activation related to song learning in young songbirds. Proc Roy Soc B-Biol Sci 277(1698):3343–3351

Goller M, Shizuka D (2018) Evolutionary origins of vocal mimicry in songbirds. Evol Lett 2(4):417–426

Goller F, Mallinckrodt MJ, Torti SD (2004) Beak gape dynamics during song in the zebra finch. J Neurobiol 59(3):289–303

Goodson JL (2005) The vertebrate social behavior network: evolutionary themes and variations. Horm Behav 48(1):11–22

Goodson JL, Kelly AM, Kingsbury MA et al (2012) An aggression-specific cell type in the anterior hypothalamus of finches. Proc Natl Acad Sci 109(34):13847–13852

Griesser M, Wheatcroft D, Suzuki TN (2018) From bird calls to human language: exploring the evolutionary drivers of compositional syntax. Curr Opin Behav Sci 21:6–12

Hahnloser RH, Kozhevnikov AA, Fee MS (2002) An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature 419:65–70

Hamaguchi K, Mooney R (2012) Recurrent interactions between the input and output of a songbird cortico-basal ganglia pathway are implicated in vocal sequence variability. J Neurosci 32:11671–11687

Hara E, Kubikova L, Hessler NA et al (2007) Role of the midbrain dopaminergic system in modulation of vocal brain activation by social context. Eur J Neurosci 25(11):3406–3416

Hessler NA, Doupe AJ (1999) Social context modulates singing-related neural activity in the songbird forebrain. Nat Neurosci 2(3):209–211

Hisey E, Kearney MG, Mooney R (2018) A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. Nat Neurosci 21(4):589–597

Hoffmann LA, Saravanan V, Wood AN et al (2016) Dopaminergic contributions to vocal learning. J Neurosci 36(7):2176–2189

Honarmand M, Riebel K, Naguib M (2015) Nutrition and peer group composition in early adolescence: impacts on male song and female preference in zebra finches. Anim Behav 107:147–158

Hsu A, Woolley SM, Fremouw TE et al (2004) Modulation power and phase spectrum of natural sounds enhance neural encoding performed by single auditory neurons. J Neurosci 24(41):9201–9211

Janik VM (2014) Cetacean vocal learning and communication. Curr Opin Neurobiol 28:60–65

Janik VM, Slater PJB (1997) Vocal learning in mammals. Adv Stud Behav 26:59–99

Jarvis ED (2013) Evolution of brain pathways for vocal learning in birds and humans. In: Bolhuis J, Everaert M (eds) Birdsong speech and language: exploring the evolution of mind and brain. MIT Press, Boston, pp 63–107

Jarvis ED, Yu J, Rivas MV et al (2013) Global view of the functional molecular organization of the avian cerebrum: mirror images and functional columns. J Comp Neurol 521(16):3614–3665

Jeanne JM, Thompson JV, Sharpee TO, et al (2011) Emergence of learned categorical representations within an auditory forebrain circuit. J Neurosci 31(7):2595–2606

Kingsbury MA, Kelly AM, Schrock SE et al (2011) Mammal-like organization of the avian midbrain central gray and a reappraisal of the intercollicular nucleus. PLoS One 6(6):e20720

Knörnschild M, Nagy M, Metz M et al (2010) Complex vocal imitation during ontogeny in a bat. Biol Lett 6(2):156–159

Konishi M (2003) Coding of auditory space. Ann Rev Neurosci 26:31–55

Kroodsma DE, Konishi M (1991) A suboscine bird (eastern phoebe *Sayornis phoebe*) develops normal song without auditory feedback. Anim Behav 42:477–487

Krützfeldt NO, Logerot P, Kubke MF et al (2010) Connections of the auditory brainstem in a songbird *Taeniopygia guttata*. II projections of *nucleus angularis* and *nucleus laminaris* to the superior olive and lateral lemniscal nuclei. J Comp Neurol 518(11):2135–2148

Langmore NE, Maurer G, Adcock GJ et al (2008) Socially acquired host-specific mimicry and the evolution of host races in Horsfield's bronze-cuckoo *Chalcites basalis*. Evolution 62(7):1689–1699

Larsen ON, Goller F (2002) Direct observation of syringeal muscle function in songbirds and a parrot. J Exp Biol 205(1):25–35

Lehongre K, Aubin T, Robin S et al (2008) Individual signature in canary songs: contribution of multiple levels of song structure. Ethology 114(5):425–435

Lewandowski B, Vyssotski A, Hahnloser RH et al (2013) At the interface of the auditory and vocal motor systems: NIf and its role in vocal processing production and learning. Journal of physiology Paris 107(3):178–192

Ligout S, Dentressangle F, Mathevon N et al (2016) Not for parents only: begging calls allow nest-mate discrimination in juvenile zebra finches. Ethology 122:193–206

Long MA, Fee MS (2008) Using temperature to analyse temporal dynamics in the songbird motor pathway. Nature 456(7219):189–194

Maddox RK, Billimoria CP, Perrone BP et al (2012) Competing sound sources reveal spatial effects in cortical processing. PLoS Biol 10(5):e1001319

Mandelblat-Cerf Y, Las L, Denissenko N et al (2014) A role for descending auditory cortical projections in songbird vocal learning. elife 3:e02152

Marler P (1956) The voice of the chaffinch and its function as a language. Ibis 98:231–261

Marler P (1982) Avian and primate communication: the problem of natural categories. Neurosci Biobehav Rev 6(1):87–94

Marler P (1967) Animal communication signals. Science 157(3790):769–774

Marler P (1970) A comparative approach to vocal learning: song development in white-crowned sparrows. J Comp Physiol Psychol 71(22):1–25

Marler P (1997) Three models of song learning: evidence from behavior. J Neurobiol 33(5):501–516

Marler P (2004) Bird calls: their potential for neurobiology. In: Zeigler HP, Marler P (eds) Behavioral neurobiology of birdsong. The New York Academy of Science, New York, pp 31–44

Meliza CD, Margoliash D (2012) Emergence of selectivity and tolerance in the avian auditory cortex. J Neurosci 32(43):15158–15168

Mello C, Vates G, Okuhata S et al (1998) Descending auditory pathways in the adult male zebra finch. J Comp Neurol 395:137–160

Menardy F, Touiki K, Dutrieux G et al (2012) Social experience affects neuronal responses to male calls in adult female zebra finches. Eur J Neurosci 35(8):1322–1336

Miller DB (1979a) The acoustic basis of mate recognition by female zebra finches (*Taeniopygia guttata*). Anim Behav 27:376–380

Miller DB (1979b) Long-term recognition of fathers song by female zebra finches. Nature 280(5721):389–391

Moore RC, Lee T, Theunissen FE (2013) Noise-invariant neurons in the avian auditory cortex: hearing the song in noise. Plos Comp Biol 9(3):e1002942

Mouterde SC, Elie JE, Mathevon N et al (2017) Single neurons in the avian auditory cortex encode individual identity and propagation distance in naturally degraded communication calls. J Neurosci 37(13):3491–3510

Nottebohm F, Stokes TM, Leonard CM (1976) Central control of song in canary *Serinus canarius*. J Comp Neurol 165(4):457–486

Nottebohm F, Kelley D, Paton J (1982) Connections of vocal control nuclei in the canary telencephalon. J Comp Neurol 207:344–357

Perez EC, Elie JE, Boucaud IC et al (2015) Physiological resonance between mates through calls as possible evidence of empathic processes in songbirds. Horm Behav 75:130–141

Phan ML, Pytte CL, Vicario DS (2006) Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds. Proc Natl Acad Sci 103(4):1088–1093

Potvin DA, Ratnayake CP, Radford AN et al (2018) Birds learn socially to recognize heterospecific alarm calls by acoustic association. Curr Biol 28(16):2632

Riede T, Goller F (2010) Peripheral mechanisms for vocal production in birds - differences and similarities to human speech and singing. Brain Lang 115(1):69–80

Riede T, Schilling N, Goller F (2013) The acoustic effect of vocal tract adjustments in zebra finches. J Comp Physiol A 199(1):57–69

Riters LV, Alger SJ (2004) Neuroanatomical evidence for indirect connections between the medial preoptic nucleus and the song control system: possible neural substrates for sexually motivated song. Cell Tissue Res 316(1):35–44

Roberts TF, Hisey E, Tanaka M et al (2017) Identification of a motor-to-auditory pathway important for vocal learning. Nat Neurosci 20(7):978–986

Schmidt MF, Wild JM (2014) The respiratory-vocal system of songbirds. anatomy physiology and neural control Prog Brain Res 212:297–335

Schneider DM, Woolley SM (2013) Sparse and background-invariant coding of vocalizations in auditory scenes. Neuron 79(1):141–152

Searcy WA, Beecher MD (2009) Song as an aggressive signal in songbirds. Anim Behav 78(6):1281–1292

Seyfarth RM, Cheney DL (2010) Production usage and comprehension in animal vocalizations. Brain Lang 115(1):92–100

Sewards TV, Sewards MA (2003) Representations of motivational drives in mesial cortex medial thalamus hypothalamus and midbrain. Brain Res Bull 61(1):25–49

Shaevitz SS, Theunissen FE (2007) Functional connectivity between auditory areas field L and CLM and song system nucleus HVC in anesthetized zebra finches. J Neurophysiol 98(5):2747–2764

Simpson HB, Vicario DS (1990) Brain pathways for learned and unlearned vocalizations differ in zebra finches. J Neurosci 10(5):1541–1556

Sitt JD, Amador A, Goller F et al (2008) Dynamical origin of spectrally rich vocalizations in birdsong. Physic Rev E 78(1):e011905

Slabbekoorn H, Smith TB (2002) Bird song ecology and speciation. Phylos T Roy Soc B 357(1420):493–503

Slater PJ, Jones AE (1995) The timing of song and distance call learning in zebra finches. Anim Behav 49(2):548–550

Sober JS, Wohlgemuth MJ, Brainard MS (2008) Central contributions to acoustic variation in birdsong. J Neurosci 28(41):10370–10379

Stoeger AS, Manger P (2014) Vocal learning in elephants: neural bases and adaptive context. Curr Opin Neurobiol 28:101–107

Suthers RA, Goller F, Hartley RS (1994) Motor dynamics of song production by mimic thrushes. J Neurobiol 25(8):917–936

Suzuki TN (2018) Alarm calls evoke a visual search image of a predator in birds. Proc Natl Acad Sci 115(7):1541–1545

Suzuki TN, Wheatcroft D, Griesser M (2018) Call combinations in birds and the evolution of compositional syntax. PLoS Biol 16(8):e2006532

Tanaka M, Sun FM, Li YL et al (2018) A mesocortical dopamine circuit enables the cultural transmission of vocal behavior. Nature 563(7729):117–119

Ter Maat A, Trost L, Sagunsky H et al (2014) Zebra finch mates use their forebrain song system in unlearned call communication. PLoS One 9(10):e109334

Vates GE, Broome BM, Mello CV et al (1996) Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches (*Taenopygia guttata*). J Comp Neurol 366:613–642

Vicario DS, Simpson HB (1995) Electrical stimulation in forebrain nuclei elicits learned vocal patterns in songbirds. J Neurophysiol 73(6):2602–2607

Vicario DS, Naqvi NH, Raksin JN (2001) Behavioral discrimination of sexually dimorphic calls by male zebra finches requires an intact vocal motor pathway. J Neurobiol 47(2):109–120

Vignal C, Mathevon N, Mottin S (2004) Audience drives male songbird response to partner's voice. Nature 430(6998):448–451

Villain AS, Boucaud IC, Bouchut C et al (2015) Parental influence on begging call structure in zebra finches (*Taeniopygia guttata*): evidence of early vocal plasticity. Roy Soc Open Sci 2(11):e150497

Vu ET, Mazurek ME, Kuo YC (1994) Identification of a forebrain motor programming network for the learned song of zebra finches. J Neurosci 14(11):6924–6934

Wang L, Narayan R, Grana G et al (2007) Cortical discrimination of complex natural stimuli: can single neurons match behavior? J Neurosci 27(3):582–589

Wang Y, Brzozowska-Prechtl A, Karten HJ (2010) Laminar and columnar auditory cortex in avian brain. Proc Natl Acad Sci 107(28):12676–12681

Wild JM (1994a) The auditory-vocal-respiratory axis in birds. Brain Behav Evol 44(4):192–209

Wild JM (1994b) Visual and somatosensory inputs to the avian song system via nucleus *uvaeformis* (Uva) and a comparison with the projections of a similar thalamic nucleus in a nonsongbird *Columba livia*. J Comp Neurol 349(4):512–535

Wild JM (2017) The ventromedial hypothalamic nucleus in the zebra finch (*Taeniopygia guttata*): afferent and efferent projections in relation to the control of reproductive behavior. J Comp Neurol 525(12):2657–2676

Wild JM, Botelho JF (2015) Involvement of the avian song system in reproductive behavior. Biol Lett 11(12):e20150773

Wild JM, Krützfeldt NE (2010) Neocortical-like organization of avian auditory 'cortex'. Brain Behav Evol 76(2):89–92

Wild JM, Krützfeldt NE (2012) Trigeminal and telencephalic projections to jaw and other upper vocal tract premotor neurons in songbirds: sensorimotor circuitry for beak movements during singing. J Comp Neurol 520(3):590–605

Wild JM, Li DF, Eagleton C (1997) Projections of the dorsomedial nucleus of the intercollicular complex (DM) in relation to respiratory-vocal nuclei in the brainstem of pigeon (*Columba livia*) and zebra finch (*Taeniopygia guttata*). J Comp Neurol 377(3):392–413

Woolley SC, Doupe AJ (2008) Social context - induced song variation affects female behavior and gene expression. PLoS Biol 6(3):525–537

Woolley SM, Portfors CV (2013) Conserved mechanisms of vocalization coding in mammalian and songbird auditory midbrain. Hear Res 305:45–56

Woolley SM, Gill PR, Fremouw T et al (2009) Functional groups in the avian auditory system. J Neurosci 29(9):2780–2793

Woolley SC, Rajan R, Joshua M et al (2014) Emergence of context-dependent variability across a basal ganglia network. Neuron 82(1):208–223

Xiao L, Chattree G, Oscos FG et al (2018) A basal ganglia circuit sufficient to guide birdsong learning. Neuron 98(1):208–221

Yanagihara S, Yazaki-Sugiyama Y (2016) Auditory experience-dependent cortical circuit shaping for memory formation in bird song learning. Nat Commun 7:e11946

Zann R (1985) Ontogeny of the zebra finch distance call. 1 effects of cross-fostering to Bengalese finches. Zeitschrift Fur Tierpsychologie (Journal of Comparative Ethology) 68(1):1–23

Zann R (1990) Song and call learning in wild zebra finches in south-East Australia. Anim Behav 40:811–828

Zann R (1996) The Zebra finch: a synthesis of field and laboratory studies. Oxford University Press, Oxford

Chapter 8
Linking Features of Genomic Function to Fundamental Features of Learned Vocal Communication

Sarah E. London

Abstract Learned vocal communication emerges from the coordination of sensory and motor learning, reflects the function of a distributed but integrated neural circuit, and unfolds across several timescales, often occurring in maturing animals. Because nearly all brain organization and function originates from patterns of genomic activation, it is crucial to understand principles of how the genome works in order to understand how learned vocal communication arises. In this chapter, the fact that genome functions have high evolutionary conservation will be leveraged to provide a conceptual guide for how research using a species of songbird, the zebra finch, can deepen and expand clinical findings from humans. Additionally, this chapter provides examples for how studies in the zebra finch can uncover fundamental processes of learned vocal communication that are of value for understanding human speech and language. Examples include the organization of specialized neural circuits, responses to social communication experiences, activation of motor plans, and consideration of how the age and sex of the individual intersect with vocal communication skills, all of which have potential to inform on vocal learning mechanisms in humans. Together, our current state of knowledge advances the idea that humans and songbirds do not simply share superficial parallels; rather, they share deep biological properties to accomplish the complex, multi-level processes required for learning and producing meaningful communication patterns.

Keywords Behavior · Critical period · Epigenetics · Genome · Histone · Learning and memory · mTOR · Neural development · Sensitive period · Sex differences · Song · Songbird · Zebra finch

---

S. E. London (✉)
Department of Psychology, Institute for Mind and Biology, Grossman Institute for Neuroscience, Quantitative Biology and Human Behavior, University of Chicago, Chicago, IL, USA
e-mail: london@uchicago.edu

J. T. Sakata et al. (eds.), *The Neuroethology of Birdsong*, Springer Handbook of Auditory Research 71, https://doi.org/10.1007/978-3-030-34683-6_8

8.1 Introduction

This chapter relates the dynamics occurring at the genomic and whole-animal levels to deepen the appreciation for how complex and important natural behaviors, such as learned vocal communication, emerge from biological substrates (also see Sakata and Woolley, Chap. 1). Here, several dimensions of genomic function, which include the sequence, regulation, and function of RNA, protein products, and epigenetic modifications, are considered in the context of learned vocal communication.

The genome is not static, thus the relationships between the genome, the brain, and behavior are interdependent. This chapter does not comprehensively describe any specific example of the genome-brain-behavior interrelationship but does consider vocal learning mechanisms in light of biological dimensions that influence vocal communication. Those major dimensions include sex, age, prior experience, social context, individual brain areas, and neural circuits. The stories presented herein demonstrate how to initiate discoveries to deepen mechanistic understanding of what biological and experiential factors influence vocal learning across species. Examples come almost exclusively from one species of songbird, the zebra finch (*Taeniopygia guttata*), as it shares multiple key features of vocal learning with humans and has the most comprehensive data across biological dimensions.

8.2 Parallels in Human and Zebra Finch Vocal Learning

8.2.1 Behavioral Similarities

The behavioral similarities between human speech and language acquisition and song learning in songbirds have been described before (Doupe and Kuhl 1999) and in this volume (Sakata and Woolley, Chap. 1; Sakata and Yazaki-Sugiyama, Chap. 2). A few broad strokes describing the process of zebra finch developmental song learning here will serve to ground later discussions of the genome within this context (Fig. 8.1).

Zebra finches live in a rich social environment throughout their lives; multi-family colonies can have over one-hundred members (Zann 1996). Song is a tool to communicate in this complex environment. Only male zebra finches can sing. Each male sings one stereotyped song his entire adult life, which can be 80% similar to another bird's song but is unique. The combination of song uniqueness and stability facilitates individual recognition within the colony across time. Males sing as part of their courtship display (called *directed song*) and after a female chooses a male, the pair forms a tight, exclusive, and long-lasting mate bond. Interestingly, male zebra finches also sing in nonreproductive contexts. This *undirected song* is thought to function as rehearsal to maintain song stereotypy (see Podos and Sung, Chap. 9).

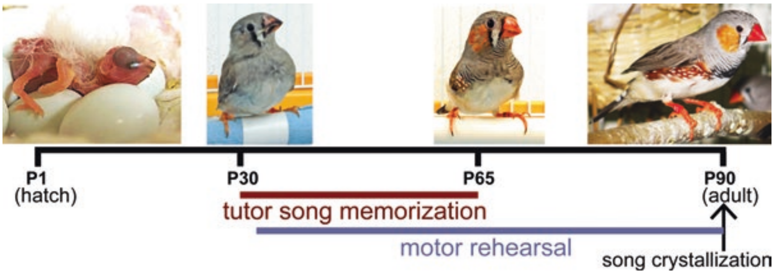


Fig. 8.1 Juvenile zebra finches learn to produce song much like humans acquire speech. Shown is a timeline of post-hatch (P) development in the zebra finch. Birds hatch on P1 and are considered adults at P90. Male zebra finches memorize the song of an adult tutor during social interactions; the ability for tutor song memorization is normally restricted to a critical period that spans P30–65. There is a period of ~30 days when the young males can memorize *tutor song*. Through a process of sensorimotor error correction, the young males use feedback to shape their own song to eventually largely resemble the tutor's song structure. As adults, each male sings one *crystalized* song that is based on his experiences with the tutor but is unique to him. In this way, song is culturally transmitted. (also see Sakata and Woolley, Chap. 1; Sakata and Yazaki-Sugiyama, Chap. 2)

Males acquire their song during posthatch development. They have one *critical period*, defined as a restricted phase when a specific experience has profound and lasting effects on a particular brain system and behavior (more on critical periods in Sect. 8.6). During the critical period for song, they can form an auditory representation of an adult *tutor song* in a process termed *tutor song memorization*. Using their memory of the tutor's song as a kind of template, young males undergo a process of sensory-motor error correction during which they alter their initial, immature vocalizations such that they come to resemble the syllable structure and order of the tutor's song. From the multi-modal integration of sensory, sensorimotor, and motor learning, each male enters adulthood with a single, unique, and highly stereotyped song (Gobes et al. 2017; London 2017).

8.2.2 *Functional Similarities in Neural Circuits*

The neural circuitry for learned vocal communication does not superficially appear equivalent in songbirds and humans. The human cortex has a typical mammalian laminar structure whereas songbird brains are organized into *nuclei*. However, the differences in macroscopic organization belie the remarkable conservation in both form and function of neural circuits across the species.

Like humans, songbirds have brain areas specialized for the learning and production of vocal communication (Fig. 8.2) (Petkov and Jarvis 2012; Pfenning et al. 2014). This includes brain areas for processing complex auditory stimuli that are integrated with the social context (the auditory forebrain or auditory lobule, AL; see Table 8.1 for all abbreviations) (details in Woolley and Woolley, Chap. 5),

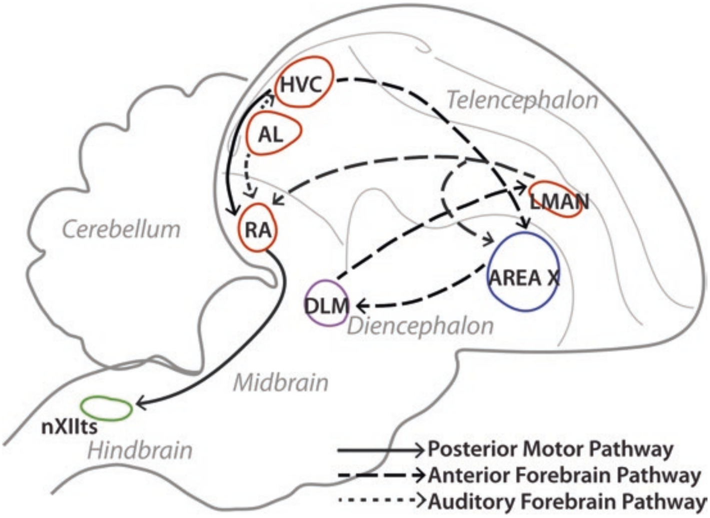


Fig. 8.2 Songbirds have a specialized neural circuit for learned vocal communication that has equivalencies in humans. Shown is a schematic of a pseudosagittal section through an adult male zebra finch brain. *Gray* outlines large regions of the brain and the regional names are labeled in *gray*. Colored *circles* and *arrows* depict the location and connections among major nodes of the song circuit. These avian brain areas function as in human circuits for speech and language. *Dark red nodes* are telencephalic nuclei similar to human cortical regions, the *blue node* designates the basal ganglia (*Area X*), *purple* outlines a thalamic relay nucleus, and *green* denotes hindbrain nuclei (*nXIIts* for syringeal control). The circuit is interconnected, and commonly divided into the posterior motor pathway (*solid arrows*), the anterior forebrain pathway (*dashed arrows*), and the auditory forebrain pathway (*dotted arrows*)(see Sakata and Yazaki-Sugiyama, Chap. 2; Woolley and Woolley, Chap. 5 for further discussions of these circuits) *AL*, the auditory forebrain, also called the auditory lobule; *DLM*, medial portion of the dorsolateral thalamic nucleus; *HVC* (used as a proper name); *LMAN*, lateral magnocellular nucleus of the anterior nidopallium; *nXIIts*, tracheosyringeal nucleus of the twelfth cranial nerve; *RA*, robust nucleus of the arcopallium

a cortico-basal ganglia-thalamic-cortical loop for fine-grained sensorimotor practice and performance (HVC, Area X, DLM, LMAN)(see Murphy, Lawley, Smith, and Prather, Chap. 3; Leblois and Perkel, Chap. 4), and areas that drive precise syringeal motor outputs coordinated with tongue and respiratory patterns (HVC, RA, nXIIts, nRA)(see Elie and Theunissen, Chap. 7). Indeed, direct functional analogies between song and language areas have been proposed (Bolhuis et al. 2010; Phenning et al. 2014), and patterns of gene expression have revealed that, although laminar structure is not a characteristic of avian brains, the genes that characterize cortical layers in mammals are expressed in songbird brains (Dugas-Ford et al. 2012; Karten 2013). Thus, songbirds and humans may share deeper features for vocal learning than the word “parallel” suggests. Future discussions may find a more suitable word that moves beyond the implication that vocal learning in humans and songbirds is outwardly similar but occurs without any shared mechanistic underpinnings.

Table 8.1 Abbreviations

AL	Auditory lobule, auditory forebrain
ASW	(gene) avian sex-specific w-linked
Cas9	CRISPR-associated protein 9
CM	Caudal mesopallium
CNO	Clozapine N-oxide
CNTNAP2	Contactin associated protein like 2
CRISPR	Clustered regularly interspaced short palindromic repeats
DLM	Medial portion of the dorsolateral thalamic nucleus
DNA	Deoxyribonucleic acid
DREADDS	Designer receptors exclusively activated by designer drugs
ERK	Extracellular signal regulated kinase
FoxP2	(protein) forkhead box P2
FoxP2	(gene) forkhead box P2
GFP	Green fluorescent protein
HDAC	Histone deacetylase
HTT	(gene) Huntington gene
HVC	Used as proper noun for vocal motor nucleus in the nidopallium
IEG	Immediate early gene
LMAN	Lateral magnocellular nucleus of the anterior nidopallium
TOR	Mechanistic target of rapamycin
mRNA	Messenger RNA
NCM	Caudomedial nidopallium
ncRNA	Noncoding RNA
nRA	Nucleus retroambigualis
nXIIts	Tracheosyringeal nucleus of the twelfth cranial nerve
P	Posthatch day
PKCi	(gene) protein kinase C iota
PTM	Post-translational modification
RA	Robust nucleus of the arcopallium
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
tRNA	Transfer RNA
ZEBrA	Zebra finch expression brain atlas
ZENK	Acronym for *zif268, egr-1, ngfi-a, krox-24*

8.3 Why Study Genomes?

What are the elements of the genome that provide both the program to reliably organize neural systems receptive to experience and the dynamic, experience-triggered responses required for processes like learned vocal communication? One of the

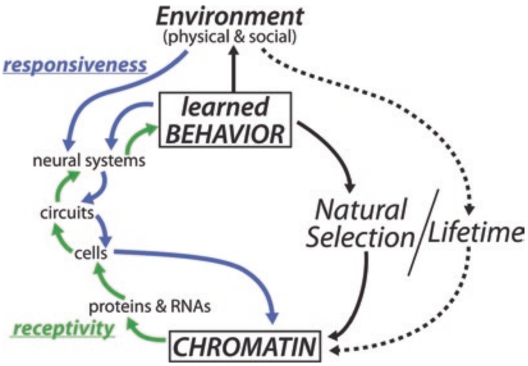


Fig. 8.3 Chromatin (the combination of genomic DNA and histone proteins) is a significant bridge between patterns of behavior and their neurobiological substrates. The role of *natural selection* in shaping the genetics of an individual (the nucleotides that comprise the sequence of genomic DNA) is well-appreciated. Genomic DNA codes for *proteins* and *RNA*s that create *neural systems* that support behavior (*receptivity, green arrows*). Genomic function can also be altered via neural responses to the environment (*responsiveness, blue arrows*). These same interdependencies exist on the timescale of a *lifetime* (*dashed arrows*)

fascinating features of the genome is that there are multiple timescales that influence how it functions. On the longest scale, there are evolutionary pressures. Evolutionary influences are reflected in the genomic DNA as sequence changes in specific regions of stability that can give clues to function when compared across generations or species. On the shortest scale, transcription can be regulated within minutes of an experience, and experiences accumulate in epigenetic modifications to the DNA and histones (proteins around which the genomic DNA wraps). Collectively, these features of the genomic DNA and histone proteins serve as a kind of biological archive of an individual, representing selection pressures placed on prior generations and the accumulation of lifetime experiences to date. The following section provides a brief overview of chromatin, the combination of genomic DNA and histone proteins, spanning the time frames relevant to the emergence of complex learned behavior (Fig. 8.3).

8.3.1 Genomic Sequence as the Central Dogma of Molecular Biology

The sequence of genomic DNA describes the genetics of an individual. Evolutionary-scale selection pressures influence DNA sequence, which manifests in signatures specific to particular species and even individuals of a particular familial lineage. DNA sequence is important because it encodes the RNAs and proteins that construct the cells that comprise brain areas and networks; the set of RNAs and proteins required for cellular structure and function is one way to define the output of the genome (Fig. 8.3).

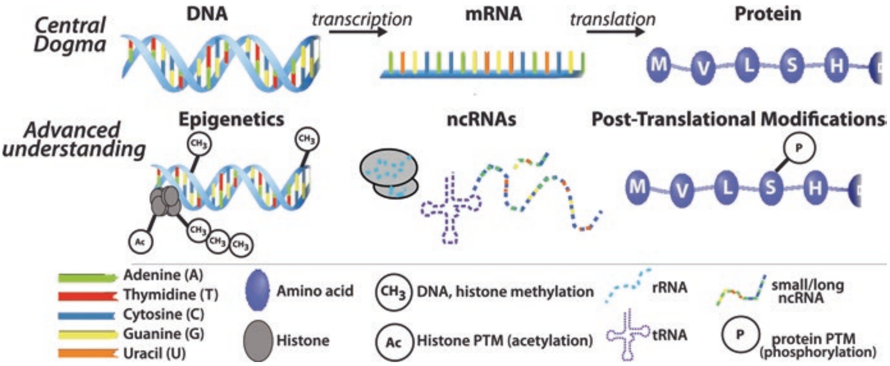


Fig. 8.4 The multilevel and interdependent regulation of genomic function provides the complement of proteins for the organization of neural circuits and for their ability to respond after experience. A major output of the genome is various proteins. The *top row* shows the linear relationship between genomic DNA, mRNA, and protein as described in the *central dogma of* molecular biology. In part because of whole genome sequencing, the number of interacting features of the process of regulating the genome has grown in complexity. The *bottom row* shows examples of our more *advanced understanding* of genomic activation at the level of DNA, RNA, and protein. *Amino acids*: D, aspartic acid; H, histidine; L, leucine; M, methionine; S, serine; V, valine (portions modified from Genome Research Limited; https://www.yourgenome.org/facts/what-is-the-central-dogma)

The *Central Dogma of Molecular Biology* states that genomic DNA is transcribed into messenger RNA (mRNA), which is transported from the nucleus to the cytoplasm for translation into protein (Fig. 8.4). The Central Dogma explains production of proteins that include the building blocks of cell morphology, the enzymes for cellular metabolism, and creation of signaling molecules, including neurotransmitters, the receptors for cell-cell signaling, the transcription factor proteins that regulate gene expression, and the hormones that can signal whole-body states to the brain. The focus on protein-coding portions of the genome has led to major breakthroughs in how brains are organized during maturation, how experience can be rapidly signaled through neural circuits, and how cells and synapses are remodeled to encode experience as memory. Researchers are collecting massive datasets of mRNAs with the aim of profiling sets of processes that occur in the brain, and experiments guided by the Central Dogma continue to elucidate neural processes. Much of the sequence of the genome, however, does not code for proteins, suggesting that research must look beyond coding regions to understand genome function.

8.3.2 Moving Beyond the Central Dogma to a More Complex View of the Genome

The sequencing and assembly of whole animal genomes, including those of the human and zebra finch (Venter et al. 2001; Warren et al. 2010), forced the revelation that the sequence of an individual's protein-coding DNA alone would not elucidate the causal relationships between an individual's genetics and how his/her brain functions to support behavior. DNA sequence that does not code for a protein, once considered "junk" DNA, is also important because much of this DNA is essential to regulate transcription (Fig. 8.4) (The ENCODE Project Consortium, 2012). For example, regulatory regions of the genomic DNA that do not get synthesized into proteins are essential for understanding how transcription is directed and can have species-specific functional consequences without significant alteration in the protein-coding sequence (Hammock and Young 2005; Gilad et al. 2006).

There are also many types of RNAs that are categorized as *noncoding* because, unlike mRNAs, they are not translated into proteins. The most abundant noncoding RNAs (ncRNAs) are ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs) that directly contribute to protein translation. There are also small and large ncRNAs that specifically and combinatorially regulate the availability of mRNAs and, therefore, the population of proteins (Wang et al. 2012; Hollins and Cairns 2016). Understanding the diversity and functional roles of ncRNAs in the brain continues to grow rapidly. New discoveries will likely be essential for understanding how the genome organizes the neural circuits for vocal learning and regulates the dynamic genomic response to sensory and motor experiences that shape vocal communication patterns (Nguyen et al. 2018; Marty and Cavaillé 2019)

Further, we now know that an individual's environment works through the brain to alter the structure–not the sequence–of the genome. Structural changes in the genome are mediated by epigenetic mechanisms (see Sect. 8.5.3). Epigenetic modifications include methylation of the genomic DNA itself, methylation of RNA, and post-translational modifications (PTMs) of histone proteins (Strahl and Allis 2000; Fu et al. 2014a; Allis and Jenuwein 2016). Both DNA and histone modifications locally alter the probability that the associated protein-coding gene will be transcribed and, therefore, meaningfully shift the output of the genome. RNA methylation alters the stability and structure of RNAs that influence how available they are for function (Zhao et al. 2016). Epigenetic modifications are almost exclusively accumulated within an individual's lifetime and thus represent a more immediate process than natural selection to affect the relationship between genomic function, brain, and behavior. Additionally, the effects of chromatin modifications are known to alter long-distance three-dimensional conformations of genomic DNA. This three-dimensional folding provides another way that chromatin structure influences transcription by bringing distal portions of genomic DNA, for example a regulatory region, in close proximity to a protein-coding gene where it can alter transcription (Lin et al. 2018).

8.3.3 Chromatin Is the Biological Hinge Point for Neural Receptivity and Responsivity

Learned behaviors such as vocal communication require an organized neural circuit that can be remodeled; the brain must be receptive to, and properly responsive to, experience. Chromatin can be regarded as a *master regulator* of cellular function and, therefore, is a hub for mechanisms of receptivity and responsiveness (Fig. 8.3). Chromatin regulates the synthesis of sets of RNAs and proteins, which build cells that assemble into circuits that form nodes in neural systems that can drive learned behavior. Receptivity results from the establishment of those neural systems. Responses to environmental stimuli or signals generated by the individual's own behavior trigger receptive neural systems, and the neural plasticity required for learning and memory results from chromatin function. Sections 8.3.1 and 8.3.2 provided a brief overview of some of the major chromatin components that are currently understand as regulators of the output of the genome. An integrated view of the response from cell membrane signaling to the nucleus and the mechanisms by which the genome's response remodels cells for plasticity has been well-conceptualized as a "genomic action potential" (Clayton 2000).

8.3.4 Summary of Chromatin in Brain and Behavior

Understanding of the complexity of genomic regulation continues to grow. As described in Sect. 8.1, the genome is the basis for nearly all of the structural and functional components of a brain, and genomic features are highly conserved evolutionarily. Each new discovery provides additional power to meaningfully investigate what makes songbirds and humans capable of the complex behavior of vocal learning as bound by shared chromatin-related processes. No one chromatin feature is likely to explain how an individual acquires learned vocalizations. Rather, vocal learning is almost certainly a combination of genes, molecular signaling activation patterns, ncRNAs, and epigenetic mechanisms that determine neural organization and function. There is an obvious need to keep investigating each of these features independently and how they may interact.

8.4 The Case of *FoxP2*

Given the features common to both human speech acquisition and zebra finch song learning, discoveries linking genomic features with speech abilities in humans can be mechanistically dissected in zebra finches. The ability to perform invasive measures and manipulations in the zebra finch is invaluable for uncovering fundamental features of vocal learning in humans, too. Investigations into the function of the gene for the Forkhead box protein P2 (*FoxP2*) demonstrate how fruitful such interspecies investigations can be.

8.4.1 Identification and Analysis of Human *FoxP2*

Mutations in the *FoxP2* gene were first discovered through classic pedigree analysis of a family known as KE. Members of the KE family displayed a speech disorder characterized by language disruptions and deficits in controlling the fine movements of the tongue and lips needed to produce speech clearly. Pedigree inspection suggested a single-gene etiology with autosomal dominance, and subsequent genetic analysis identified a narrow genomic region that contained the *FoxP2* gene (Fisher et al. 1998; Lai et al. 2001). An independent case of speech disruption convincingly connected mutations in *FoxP2* to disorders in speech (MacDermot et al. 2005).

The type of gene mutation that affected KE family members was a single *point mutation*, which alters one amino acid in the entire protein structure. The human FoxP2 protein is over 700 amino acids long, so one change seems to have a disproportionate impact. The reason such a small genetic change can have such a dramatic effect on complex behavior is because the FoxP2 protein is a transcription factor. Transcription factors bind to gene regulatory regions of the DNA to influence the expression level of the associated genes. Each transcription factor can regulate many protein-coding genes. The KE family's *FoxP2* mutation was in the segment of the gene that codes for its DNA-binding domain; therefore, the transcriptional regulatory function of FoxP2 amplifies the impact of its mutation in the KE family (Nudel and Newbury 2013).

Curiously, the FoxP2 protein sequence is almost 100% conserved across species. That fact made the discovery of two amino acids that were seemingly unique to humans compared to nonhuman primates even more compelling (Enard et al. 2002). The species difference raised the possibility that these two changes conferred language ability to humans. Notably, these amino acids are not the same as those associated with speech disorders. Instead, the human-specific amino acids are found within coding exon 7. This region of exon 7 does not encode for a known protein functional domain, suggesting that the effect of the amino acid changes may broadly influence the three-dimensional structure of FoxP2 in ways that affect its function other than its direct ability to bind DNA. Despite the insights gained from the myriad of clinical and comparative *FoxP2* studies, a number of questions about the importance and mechanisms of FoxP2 function remained unanswered from these endeavors.

8.4.2 Mechanistic Questions About *FoxP2* Function Addressed in Songbirds

Findings in humans set up three big research questions about FoxP2 functions in vocal communication that are best answered with a nonhuman animal model:

1. Do species-typical *FoxP2* sequences dictate the ability for learned vocal communication?

2. What does the pattern of *FoxP2*expression tell us about its function in learned vocal communication?
3. What are the downstream genes regulated by FoxP2 that may explain its influence on learned vocal communication?

The first question was tackled by doing a comparative analysis of *FoxP2* transcript sequences across multiple species, including vocal learners (humans and songbirds) and vocal nonlearners (mice and birds that are not songbirds) (Haesler et al. 2004; Scharff and Haesler 2005). These analyses yielded two important insights. First, sequence comparison showed a remarkably high degree of predicted protein sequence conservation across species: >98% (Teramitsu et al. 2004; Haesler et al. 2004). In comparative analyses, greater sequence similarity provides evidence that a specific stretch of sequence has functional significance. Data indicate that mammals and birds last shared a common ancestor more than 300 million years ago; thus, the nearly identical protein sequence across the four groups of vertebrates tested suggests that the function of FoxP2 is also unchanged (O'Leary et al. 2013; Prum et al. 2016). Second, the amino acids that distinguished human *FoxP2* from other primate gene sequences—the individual changes that were postulated to confer vocal learning capability—were not observed in songbirds. In other words, the sequence of the *FoxP2* gene does not systematically sort with vocal learning ability; genetics did not indicate that specific variants of the *FoxP2* gene predict vocal learning in a species.

The second question focuses on ways to infer the function of FoxP2 in vocal learning by examining its gene expression patterns in brain areas required for song. For example, one of the structural brain differences in affected KE members compared to other members was found in the basal ganglia (Vargha-Khadem et al. 1998). The basal ganglia have several functions, including the learning and production of finely tuned motor skills like those required for speech (see Leblois and Perkel, Chap. 4). In humans, it is not possible to definitively determine if neural differences in affected KE family members were the cause of the speech impairments or if they arose after years of impaired speech production. However, it is possible to construct an argument for a causal relationship in zebra finches. Indeed, *FoxP2* is expressed in the songbird basal ganglia (Area X; Fig. 8.2), starting at the earliest stages of neural development (Teramitsu et al. 2004). It is expressed in the major cell type of the striatal portion of the basal ganglia in humans and songbirds: the medium spiny neurons (Haesler et al. 2004; Kreitzer 2009). Further, its transcription is rapidly reduced in Area X when males sing (Scharff and Adam 2013), indicating that FoxP2 may be involved in the process of developmental song learning.

These studies laid essential groundwork to implicate FoxP2 in the same component of vocal communication in zebra finches as was affected in the KE family. One causal test would be to reduce FoxP2 production in Area X and ask if the resulting song phenotype was disrupted; reducing the abundance FoxP2 in the zebra finch would be functionally akin to a human gene mutation that minimizes the function of FoxP2. FoxP2 knockdown was first reported in 2007 and, indeed, reducing FoxP2 levels in Area X of juvenile males during the phase of song learning when they depend on it for sensorimotor learning led to deficits in song structure (Haesler et al. 2007).

In addition, because of the patterns of regulated transcription when birds sing, overexpression could disrupt vocal learning (Murugan et al. 2013; Heston et al. 2018) as there are multiple systems that follow a “Goldilocks scenario” whereby too much or too little of a cellular process prevents the signal-to-noise ratio required to convey information. With expression of designer receptors exclusively activated by designer drugs (DREADDs), it is possible to inhibit or potentiate cell firing using a nonbiological ligand (Urban and Roth 2015; Roth 2016). For existing DREADDs, the intended ligand was an antibiotic called clozapine N-oxide (CNO), although it can be reverse metabolized into clozapine, resulting in unintended effects that were not always accounted for in early studies. However, DREADDs expressed in Area X revealed a complicated relationship with LMAN in the execution of moment-to-moment variability in adult song production (Heston et al. 2018). Generally, these types of manipulations, which cannot be done in humans, were essential to support the hypothesis that FoxP2 mutations can have causal effects on basal ganglia function that lead to deficits in vocal production patterns.

The baseline and singing-regulated pattern of *FoxP2* transcription, as well as results from its manipulated expression patterns and data on Area X function, were consistent with the notion that FoxP2 contributes to vocal production. Thus, it was necessary to address the third question that requires identification of the genes that FoxP2 transcriptionally regulates. This is important because perturbations in the availability of these factors would be most directly related to deficits in basal ganglia function and vocal production.

Several experimental strategies can be employed to discover individual genes or sets of genes regulated by a transcription factor. One strategy is to survey the genome for the short DNA regulatory sequences that FoxP2 proteins recognize as locations for binding and then determine which protein-coding genes are associated with those regulatory regions. One of the individual genes identified in this way was Contactin Associated Protein Like 2 (CNTNAP2)(Fisher and Scharff 2009). CNTNAP2 is an intriguing protein because it is an adhesion molecule that affects cellular properties that direct cell-to-cell communication (Fisher and Scharff 2009). In the zebra finch, the FoxP2-CNTNAP2 interaction was confirmed in brain areas relevant to vocal communication and was regulated in ways consistent with a role in song production (Panaitof et al. 2010; Condro and White 2014; Adam et al. 2017). A second strategy is to manipulate FoxP2 in zebra finches and identify genes differentially expressed as a result. Using a combination of experimental conditions, testable hypotheses regarding the shifting gene networks can be formulated. This strategy has revealed transcriptional networks in Area X that are perturbed by alterations in FoxP2 DNA binding and that may be specifically involved in developmental song-motor learning (Burkett et al. 2018). Other strategies that revealed miRNAs that regulate *FoxP2* mRNAs (see Sect. 8.7 for more information) added an epigenetic layer of modulation onto the genetic and genomic processes by which FoxP2 influences vocal learning (Shi et al. 2013; Fu et al. 2014b).

8.4.3 FoxP2 Summary

As the case study of *FoxP2* demonstrates, studies in songbirds provide integrative mechanistic data on a variety of questions that would be difficult to acquire in humans. For example, work in songbirds can reveal when a gene is expressed during development, if its expression is restricted to specific cell types, and how its expression is localized across the neural network for vocal learning. Songbird research also allows causal manipulations that link gene expression to behavior. Each of these general advantages are part of the *FoxP2* story. In short, studies in zebra finches can confirm but also expand our knowledge base of how genes can affect the acquisition and regulation of complex behaviors such as vocal communication.

8.5 Genome-Brain-Behavior Interdependencies in Songbirds Inform On Human Communication

Songbirds have demonstrated value for identifying neural and experiential mechanisms that influence learned vocal communication. For example, the zebra finch model permits meaningful investigation into both developmental and adult processes and the separation of auditory and motor components of learned vocal communication. Importantly, these studies can be combined with different readouts of genomic function in the context of cells, circuits, and the whole animal. Further, the nuclear structure of the songbird brain confers advantages for investigation because functional areas can be identified with the naked eye. Visible structures allow for specific and reliable anatomical localization of genomic features to test how they segregate among behavioral components of vocal learning. The following subsections describe the difference between *genetics*, the static DNA sequence, and *genomics*, the dynamic regulation of transcriptional output of the genome. In the following subsections, chromatin is positioned as the center of both upstream and downstream molecular processes that regulate genome function. In addition, sex differences, epigenetics, and immediate early gene expression are used as a backdrop to understand the interplay between genes, brain, and behavior in vocal learning.

8.5.1 Sex Differences as a Path to Mechanism (Receptivity)

Behaviors typically acquired during development, such as vocal communication, depend on the construction of a neural network that is sufficiently organized to encode the experiences that guide vocal output patterns. Biological sex is one organizing process that affects the brain and begins very early in maturation; therefore, biological sex may provide some key insights into mechanisms that create the neural network for vocal learning.

In humans, sex differences in language disorders are documented, although they are more difficult to parse mechanistically than in songbirds. Often speech and language deficits are associated with broader syndromes, such as *autism spectrum disorders* (Halladay et al. 2015), *attention deficit hyperactivity disorder* (Cohen et al. 2000), and *schizophrenia* (Walder et al. 2006), which can complicate investigation. Further, sex differences can be blurred by the broad range of individual variation and can be influenced by environmental factors such as differential application of special education intervention measures (Barbu et al. 2015; Kvande et al. 2018). These factors mean that differences in speech and language abilities between boys and girls are often difficult to parse by gender alone. Additionally, there are some distinctions in how language function is represented in the brains of adult men and women, but determining if the structural distinctions are a cause or effect of potential sex differences in speech and language is nearly impossible (Wallentin 2009; Etchell et al. 2018). Ultimately, the effect of sex on vocal communication phenotypes is not clear.

In zebra finches, however, there is one of the starkest sex differences described in brains. The sex difference is apparent with the naked eye, and it encompasses the motor and sensorimotor nuclei of the song circuit (remember, females cannot produce song, but males can) (Nottebohm and Arnold 1976; Wade and Arnold 2004). These differences make zebra finches valuable for discovering mechanistic underpinnings of sex differences with the potential to uncover organizational principles of vocal learning circuits.

The dogma of sexual differentiation, as defined from mammalian studies, is essentially that gonadal steroids do it all: a gene on the sex chromosomes determines whether or not testes develop, and testicular secretions form a masculine brain and body (Arnold and Schlinger 1993; Arnold et al. 2004). Gonadally derived steroids are undoubtedly powerful mediators of maturation, but they do not explain all of sexual differentiation. In fact, understanding how the genome is regulated in the brain has revealed how the dogma fails to fully explain how brains are organized. Notably, steroids can be locally synthesized in the brain to influence specific functions (see Remage-Healey, Chap. 6), and the complement of sex chromosomes themselves alter the phenotype of brain cells, in part because sex chromosome-linked genes are transcribed in the brain (Arnold et al. 2004; London et al. 2009b). Because extreme differences can be useful for initial discovery steps, the zebra finch affords a unique opportunity to consider mechanisms of vocal circuit organization that can then be applied to more subtle systems like those in humans.

8.5.1.1 Direct Genetic Effects: Sex Chromosome Gene Expression

Each sex chromosome can code for unique protein variants of the same gene. Because males and females differ in their complement of sex chromosomes—mammalian males are XY and females are XX; avian males are ZZ and females are ZW—genetic differences can directly influence the brain via neural expression of genes localized to sex chromosomes. Importantly, the sex chromosome complement that is expressed in rodent brains affects essential cellular features, including the abundance of neurons with specific neurochemical phenotypes (Carruth et al. 2002; Arnold and Chen 2009). Quite possibly, similar processes occur in humans.

Perhaps the most striking demonstration of sex chromosome gene expression in the brain comes from a naturally occurring gynandromorphy: an individual that is nearly perfectly split hemispherically as male and female (Agate et al. 2003). In a zebra finch gynandromorph, transcription of a gene variant found on the W chromosome, *ASW*, was restricted to the side of the bird that matched the female plumage. Further, the version of the protein kinase gene *PKCi* localized to the Z chromosome was more highly expressed in the right hemisphere, matching the side where male-typical plumage was present (there is minimal Z-inactivation and therefore male:female dosage of Z-linked genes is typically close to 2:1) (Agate et al. 2003; Itoh et al. 2007).

Because the entirety of the gynandromorph's brain would have been exposed to the same environment of circulating gonadal hormones, this individual provided a rare opportunity to test for direct genetic effects on song circuit sexual dimorphism. Indeed, when the volumes of major singing nuclei were measured, they were larger in the right hemisphere (greater Z gene-expressing and male typical plumage) than in the left, as compared to the more symmetrical volumes found in normal males. Interestingly, however, nuclei in the left (W gene-expressing female hemisphere) were also partially masculinized, indicating that while direct genetic effects likely determined the majority of the song nuclei volumes, there may be additional, local signaling molecules such as neurally synthesized steroids (see Sect. 8.5.1.2) that could act on both hemispheres.

8.5.1.2 Effects of Regulated Gene Expression: Autosomal Chromosomes

The singing circuitry is masculinized by the steroid estradiol. All steroids are synthesized from cholesterol molecules through a series of enzymatic conversions. The spatiotemporal distribution of steroidogenic enzymes, therefore, determines the steroid-producing capacity of particular regions. In developing and adult zebra finches, circulating levels of estradiol are indistinguishable in males and females, and estradiol is produced within the brain, indicating that the enzyme required for the last step of estrogen synthesis is present in the brain (Adkins-Regan et al. 1990; Schlinger and Arnold 1992). However, the song circuitry is still masculinized in genetic males gonadectomized early in development, leading to the hypothesis that all five major enzymes needed to convert cholesterol to estradiol are in the brain (Arnold 1975; Arnold and Schlinger 1993). Steroids synthesized from enzymes expressed within the brain itself are termed *neurosteroids* to distinguish them from steroids originating in the periphery (London et al. 2009b).

Sex differences in the zebra finch song circuit are detectable as early as nine days after hatching (posthatch day 9, P9)(Gahr and Metzdorf 1999; Kim et al. 2004), and estradiol is most masculinizing during the first week of posthatch life (Adkins-Regan et al. 1994). Consistent with neurosteroid contributions to these early processes of sexual differentiation, the genes for steroidogenic enzymes are expressed at P1 and P5 (London and Schlinger 2007). Interestingly, transcription occurs within the cells along the lateral ventricle, especially where neurogenesis is particularly prolific (Dewulf and Bottjer 2005), indicating that neurosteroids may be affecting brain organization at its earliest stages. Genes for steroidogenic enzymes continue to be expressed in song nuclei at later developmental ages and are transcribed in different combinations across brain areas (London et al. 2006). Each steroid can have multiple effects and drive fundamental elements of brain organization in other systems (London 2016); thus, control of neurosteroid production via genomic regulation may influence sex differences in vocal communication abilities and may hold some potential for understanding sex differences in communication disorders in humans.

8.5.2 Dynamic Experience-Dependent Processes for Vocal Learning

Learning is by definition a dynamic process and, therefore, cannot be completely explained by the static genetic sequence of an individual. Long term memory formation requires new transcription and translation as a result of an experience. There are several ways to examine mechanisms that modulate patterns of transcription and translation and to detect patterns of genomic activation that correspond to experience-dependent neural processing. The following subsections provide an overview for how two of these strategies, immediate early gene (IEG) expression and patterns of epigenetic modifications, can promote our understanding of gene, brain, and behavioral interdependencies.

8.5.2.1 Immediate Early Genes as a Tool for Anatomical and Functional Discovery

Transcription of IEGs is regulated by pre-existing transcription factor proteins that can be activated within milliseconds after cell firing. Immediate early genes, therefore, are among the first new mRNAs and proteins generated after a cell has fired, and they can be used to identify cells and brain areas that were active during an experience. Their transcription depends on cellular activation, but the absence of IEG expression does not mean that a cell has not fired. Instead, IEGs represent the activation of selective molecular processes triggered by cell firing. Their expression provides two levels of information: the cell has fired and a specific molecular process was initiated (Tischmeyer and Grimm 1999; Minatohara et al. 2016). ZENK, an IEG with multiple names (*zif268*, *egr-1*, *ngfi-a*, *krox-24*) has been most comprehensively studied in zebra finches (Mello et al. 2004). This gene was described as a necessary component of behavioral learning in other animal systems and has been leveraged to probe features of vocal learning and production (Bozon et al. 2002; Alberini 2009). Sections 8.5.2.1.1 and 8.5.2.1.2 provide an overview of adult song-recognition learning and developmental song learning in the zebra finch and how IEG expression studies lend insight into the molecular cascades initiated in learning and plasticity.

8.5.2.1.1 Sensory Learning in Adults: Song Recognition Learning

Male and female adult zebra finches learn to recognize the songs of others, which helps them to distinguish individuals within their colony. In adults, IEGs such as ZENK are rapidly transcribed after a bird hears songs of other zebra finches that are unfamiliar to them (Mello et al. 2004). Interestingly, the numbers of cells that induce ZENK transcription are reduced as exposure to the same song is repeated (Dong and Clayton 2008, 2009). The characteristics of this process are consistent with canonical definitions of habituation, which is a form of nonassociative learning (Thompson and Spencer 1966; Rankin et al. 2009).

Two regions in the brain show ZENK transcription upon hearing novel conspecific songs: the caudomedial nidopallium (NCM) and caudal mesopallium (CM). Both NCM and CM are major components of the *auditory forebrain*, which is a composite brain area outside of the traditional song circuit for motor control (Vates et al. 1996). Interestingly, NCM and CM are tightly interconnected with an adjacent primary auditory cortical area, Field L. However, unlike NCM and CM, neurons in Field L do not express ZENK in response to hearing novel conspecific songs (Mello et al. 2004). This underscores the molecular specificity of the IEG response. Because NCM and CM receive their information from Field L, neurons in Field L must be firing for NCM and CM to be activated, yet the molecular cascade to transcribe ZENK is not triggered in Field L cells. With one exception, *dusp-1*, data to date indicate that hearing complex and biologically meaningful sounds initiates genomic regulation in NCM and CM, but not Field L (Horita et al. 2010). Of course, future studies may provide additional examples of IEGs transcribed in primary auditory areas and greater complexity of IEG-mediated auditory processing will be revealed.

The initial discovery of song-induced IEG expression led to further insights about the mechanistic complexity of sensory song learning in adults. For example, the numbers of cells in NCM and CM expressing IEGs positively correlate with the biological relevance of the particular song being heard, including directed versus undirected song (Mello et al. 2004), mate versus unfamiliar male (Woolley and Doupe 2008), and higher-order structural song complexity (Lin et al. 2014). All of these findings were consistent with the idea that IEG activation was a feature of NCM and CM neurons responding to higher-order features of the song, not simply responding to the basic acoustic features. This was confirmed by studies that demonstrated that changing the physical and social contexts in which song was experienced could alter the magnitude of the genomic response in NCM and CM (Kruse et al. 2004; Vignal et al. 2005) and by experiments that demonstrated how prior social contacts influenced the response to hearing song (Woolley and Doupe 2008;

Lin et al. 2014). These findings also indicated that NCM and CM processing is more complicated than pure auditory processing. Instead, contextual features modulate the processing of song stimuli (see Woolley and Woolley, Chap. 5). How and why this can occur is unknown.

8.5.2.1.2 Sensory Learning in Juveniles: Tutor Song Memorization

As juveniles, males memorize their tutor's song, and this sensory learning serves as the foundation of their own song structure (see Sakata and Yazaki-Sugiyama, Chap. 2). There is also evidence that young females learn the song of their dad, though this is more difficult to assess because females cannot sing (Braaten et al. 2006; Braaten et al. 2008).

The baseline density of ZENK expression in NCM and CM was as high at the age of onset of the critical period for tutor song memorization as it was in adults who had heard biologically relevant song (Jin and Clayton 1997; Roper and Zann 2006). Therefore, it was possible that ZENK expression was necessary for tutor song memorization to occur. In the early 2000s, it was not technically possible to directly "knock-down" a single gene in the songbird brain, but it was possible to disrupt an upstream protein signal that was necessary for ZENK transcription, ERK (extracelluar signal regulated kinase) (Cheng and Clayton 2004). Combining a transient disruption of ZENK induction in NCM and CM during a juvenile male's tutor experiences prevented him from producing high fidelity copies of the tutor's song (London and Clayton 2008). This experiment thus made two contributions: (1) molecular regulation of experience-dependent IEG expression was causally linked with sensory learning, and (2) the NCM/CM regions were identified as essential anatomical loci for tutor song memorization.

8.5.2.1.3 Motor Learning and Production

A large portion of the song circuit integrates sensory information with motor output during developmental song learning and drives song production across the lifespan of the individual (see Sakata and Yazaki-Sugiyama, Chap. 2; Murphy, Lawley, Smith, and Prather, Chap. 3). A lot of attention is given to the highly stereotyped nature of adult song, but IEG studies have provided some interesting mechanistic clues about how subtle changes in song structure, especially across maturation and in different social contexts, may occur.

When adult males sing, IEGs are expressed within the components of the circuit that control song production (Jarvis and Nottebohm 1997; Kimpo and Doupe 1997). Singing adults do not need to hear for ZENK to be expressed in the motor circuitry: deafened birds show the same pattern of induction across motor circuitry as hearing birds when they sing undirected song (Jarvis and Nottebohm 1997). Thus, at least by maturity, the motor circuit appears to generate song independent of the signaling required for NCM and CM sensory processing in juveniles.

Singing does not induce the same distribution of ZENK across the motor circuitry in juvenile males compared to adult males (Jin and Clayton 1997). In particular, ZENK mRNA is restricted to the posterior portion of RA (robust nucleus of the arcopallium) in adults but is expressed throughout RA in P35 males who have just begun the process of vocal learning (Fig. 8.1). There is an intriguing shift in synaptic inputs from LMAN and HVC onto neurons in RA as the birds develop their song; perhaps the synaptic anatomy and genomic function are functionally connected, and the developmental change in connectivity alters the distribution of molecular signaling in RA (Mooney and Konishi 1991; Aronov et al. 2008). If so, this would be an interesting example of how IEG induction reveals specific molecular processes underlying neural function.

Additionally, within adults, the social context in which the male sings changes the neural pattern of ZENK expression. After an adult male sings directed song to a female, ZENK expression in LMAN, Area X, and RA is lower than after the bird sings undirected song, even if he is surrounded by other birds but not singing directly to any of them (Jarvis et al. 1998). Not everyone is comfortable integrating information across different experiments within a study, but it is possible that, unlike what occurs in adults who sing undirected song, directed song induces ZENK in NCM and CM, and this induction during directed song is prevented by deafening (Jarvis et al. 1998). There may be additional modulatory signals that combine with auditory input during directed song to initiate cascades that promote ZENK transcription in these higher-order sensory processing areas. Lastly, it is important to note that there may be additional molecular regulation on IEG function in song control nuclei after transcription has occurred as ZENK protein distributions do not always recapitulate ZENK mRNA patterns (Whitney et al. 2000). More work tracking mRNA and protein dynamics will be needed to understand what, if any, the functional ramifications of this disconnect are.

8.5.2.2 Epigenetics in Adult Song

As introduced in Sect. 8.3.2, epigenetic mechanisms are defined as those that change the function of the genome without altering the sequence of the genomic DNA. There are several types of epigenetic mechanisms: ncRNAs, modifications to DNA (DNA methylation), modifications to RNA (RNA methylation), and a diversity of modifications to histone proteins. All of these processes have the effect of modulating the abundance and mixture of mRNAs available for translation in a cell. Like other chromatin features, epigenetic modifications appear to be highly conserved evolutionarily in terms of function. The understanding of epigenetic mechanisms active in the song circuit is still nascent but will likely grow in the coming years.

The following sections will review recent research that investigated ncRNAs and histone PTMs in adults to demonstrate that these approaches have value in testing mechanisms of recognition learning and vocal production (epigenetic modifications in juvenile zebra finches are discussed in Sect. 8.6). There is much remaining to be discovered about how epigenetic mechanisms influence vocal learning and production.

8.5.2.2.1 miRNAs

The number of transcripts available in the cell for translation can be lowered by miRNAs via binding to short recognition sequences in mRNAs (O'Brien et al. 2018; Gebert and MacRae 2018). Predicting the effect of one miRNA is difficult because each mRNA often includes multiple miRNA recognition sequences, there can be more than one type of miRNA that recognizes the same mRNA, and there may be more than one site per mRNA for a particular miRNA (Kim et al. 2016). Upon the miRNA binding to the mRNA, the mRNA is cleaved by a protein complex recruited by the bound miRNA, preventing the mRNA from being translated into protein (Shukla et al. 2011; O'Brien et al. 2018).

Many more miRNAs have been predicted in the zebra finch than have been functionally studied (Luo et al. 2012). Initial reports, however, indicate that they are involved in auditory processing and song production in adults and may regulate sex-specific responses as well as the availability of FoxP2 (Gunaratne et al. 2011; Shi et al. 2013). For now, these reports demonstrate that there is great potential for additional miRNA investigations. As data are acquired from future studies, miRNAs will likely become recognized as important regulators of the dynamic processes underlying vocal communication.

8.5.2.2.2 Histone Post-Translational Modifications

Like other proteins, histones can be post-translationally modified by the addition of molecular side chains such as phosphate, acetyl, and methyl groups. The specific type of PTM, the number of PTMs, the amino acid that receives the PTM, and the specific histone protein with PTMs all alter chromatin structure, which in turn shifts the probability that the associated DNA regions will be transcribed (Strahl and Allis 2000).

The addition and removal of histone PTMs is performed by a set of enzymes that are often referred to as *writer*s and *erasers*. The full diversity of histone modifications that act in adult songbirds is not known, but at least one eraser that removes acetyl groups (a histone deacetylase, HDAC) influences adult song recognition. Prior work in rodents demonstrated that accumulation of histone acetylations via HDAC3 deletion enhanced learning and memory such that a subthreshold learning experience was transformed into one that coded a memory that lasted at least 24 hrs (McQuown and Wood 2011; McQuown et al. 2011). In adult zebra finches, pharmacological inhibition of HDAC3 in the auditory forebrain after a subthreshold song playback experience also resulted in neural measures indistinguishable from a more robust experience known to support song recognition learning (Phan et al. 2017). This is consistent with a pattern seen in other systems, including in human auditory processing, in which increased histone acetylation improves learning, although perceptual training alone can have similar effects in humans (Gervain et al. 2013; Van Hedger et al. 2015).

After adult males have sung, epigenetic modifications may also influence the transcriptional probability of genes involved in song stability. One type of histone acetylation (on Lysine 27 of histone H3: H3K27ac) is used to identify regions that can be actively transcribed. Using a procedure called *chromatin immunoprecipitation* followed by DNA-seq (ChIPseq), analysis of H3K27ac-associated DNA confirmed that there is a set of genes that can be transcribed selectively in each of four major song control nuclei: HVC, LMAN, RA, and Area X (Whitney et al. 2014). Collectively, approximately 2000 genes were found to be actively regulated upon singing across these regions. Discoveries like these open the door to many more investigations about how epigenetic mechanisms can stably and dynamically influence the acquisition and production of vocal communication.

8.5.3 Conclusion of Dynamic Experience-Dependent Processes

Investigations at various levels of genomic function in the zebra finch continue to uncover features of brain circuit organization and function. The advantages of being able to dissociate how genomic regulation influences sensory functions compared to motor functions, evaluate how maturation is controlled, and assess how the genome associates with behavioral measures of learned vocal communication make zebra finches a powerful resource for revealing meaningful genomic function.

8.6 Critical Period Informs on Mechanisms that Modulate Learning Plasticity

One of the most striking features of song learning in the zebra finch is that the process of tutor song memorization is defined by a critical period, as described previously (Fig. 8.1). Further, that same experience has no measurable effect on behavior before or after the critical period. While the critical period "open" may be set by genetically determined maturational processes, the "close" depends on experience, and preventing the relevant experience extends the age at which brain and behavior remain open to its effects (Figs. 8.1, 8.5) (e.g., Knudsen 2004; Takesian and Hensch 2013). Because brain areas with critical periods undergo extreme fluctuations in experience-dependent plasticity, they are excellent models in which to disentangle chromatin mechanisms of maturation from those required for experience-dependent learning.

In the zebra finch, there is evidence that the critical period regulates the social sensory process of tutor song memorization rather than the motor rehearsal component of developmental song learning (see Sakata and Yazaki-Sugiyama, Chap. 2). Notably, preventing young males from hearing song between P30–65 extends the

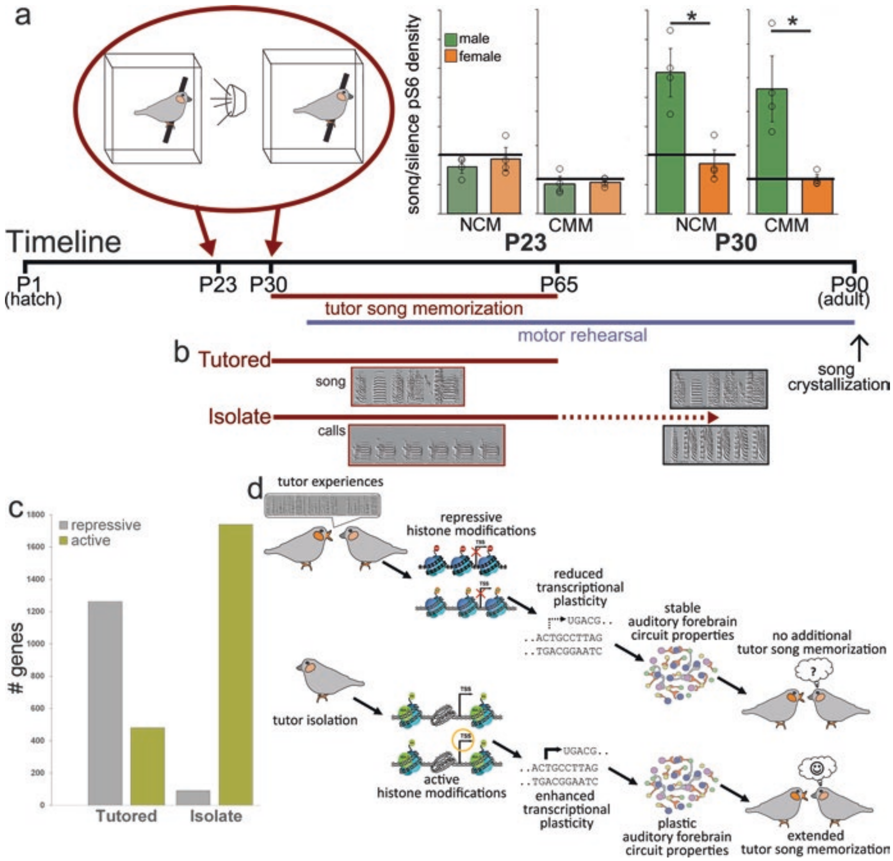


Fig. 8.5 Multiple timescales of song learning can be measured with appropriate levels of neurobiology. (**a**) A schematic depiction of a timeline highlighting the critical period for tutor song memorization and major elements of developmental song learning. A paradigm of acute song playbacks (left, in *red circle*), in which juveniles were either played song or left in silence, revealed sex and age differences in mTOR cascade activation, quantified by the density of S6 phosphorylation (pS6), in the caudomedial nidopallium (*NCM*) and caudal mesopallium (*CM*) (histograms on right; asterisk indicates statistical significance $p < 0.05$). (**b**) Juvenile males reared under a tutor (reared with an adult male; example song depicted) during P30–65 have their critical period closed, while song-isolated juveniles (reared with adult females that do not sing but do produce calls, which have similar acoustic features to some song syllables, example depicted) have an extended critical period (extension is denoted by *dashed red line*). Without additional tutoring, tutored males sing faithful copies of the tutor's song as adults and isolate-reared males sing an abnormally structured song. (**c**) Measures of epigenetic histone post-translational modifications in the auditory forebrain from tutored and isolate-reared males showed distinct proportions of repressive and active chromatin. (**d**) These data lead to a hypothesis linking tutor experience with critical period closing, mediated by epigenetic mechanisms. (portions of this figure were modified from Ahmadiantehrani and London 2017a; London 2017; Ahmadiantehrani et al. 2018; Kelly et al. 2018)

age at which tutoring successfully shapes the juvenile's song structure (London 2017); but birds reared without tutoring still sing, with essentially no alteration in the patterns of gene expression in song control nuclei (Mori and Wada 2015). Interestingly, males can be raised with adult females that produce innate calls with acoustic structures similar to some types of song syllables, but this auditory and social experience is not sufficient to close the critical period (Fig. 8.5) (Eales 1985, 1987). Because tutor song memorization relies on genomic and molecular processes in NCM and CM during tutor experiences (London and Clayton 2008; Ahmadiantehrani and London 2017a), it is possible that an important neural locus for the critical period is in the auditory forebrain.

Evidence for genomic regulation as a determining factor in the onset and offset of the critical period exists in the auditory forebrain across biological scales (Fig. 8.5). Single-gene expression patterns were the first indication that baseline and experience-dependent changes occurred in the auditory forebrain at ages relevant to the critical period (Jin and Clayton 1997). In fact, the profile of auditory forebrain RNAs are different at baseline (in silence) and after hearing song when compared between males prior to and after the critical period (London et al. 2009a). Additionally, hearing song activates a particular molecular signaling cascade, the *mechanistic target of rapamycin* (mTOR), in male NCM and CM at P30, the onset of the critical period (Roper and Zann 2006), but not one week earlier and not in females who cannot sing at either age (Fig. 8.5) (Ahmadiantehrani and London 2017a). Manipulations of mTOR signaling in juvenile males experiencing a tutor prevent high fidelity tutor song copying (Ahmadiantehrani and London 2017a). The same mTOR manipulations do not have the equivalent effect on adult song recognition learning as they do on tutor song memorization (Ahmadiantehrani et al. 2018). Interestingly, mTOR regulates the initiation of protein synthesis (Hoeffer and Klann 2010), suggesting that proteins controlled by mTOR activation may underlie the onset and specificity of the critical period for tutor song memorization (Ahmadiantehrani and London 2017a; Ahmadiantehrani et al. 2018).

Importantly, critical periods are distinguished from age-limited learning by the fact that it is experience, not age, that closes the phase of experience-dependent neural plasticity. Knudsen (2004) reviews the logic of this tie between experience, behavior, and neural plasticity in several systems. Therefore, it is important to dissociate age from prior tutor experience to understand how tutor song memorization prevents subsequent tutor experience from being learned. Epigenetic mechanisms that alter the balance of repressed and active chromatin, especially surrounding genes involved in transcription and translation, are associated with closed and extended critical periods in P67 male auditory forebrain (Fig. 8.5) (Kelly et al. 2018). The data support the hypothesis that tutor song memorization leads to an accumulation of repressive chromatin that limits transcriptional and cellular responses following a tutor encounter that occurs after P67 (Fig. 8.5). Because memory formation requires new transcription and translation, this limited genomic response prevents additional tutor song memorization. On the other hand, tutor song isolation during the critical period leads to relatively more abundant regions of active chromatin. Active chromatin has high potential for transcription, which would then permit the complement of new proteins required for memory formation to be synthesized even when the male experiences a tutor past the normal close of the critical period (Fig. 8.5).

In general, language acquisition is believed to be most effective earlier in development, but whether or not there are one or more critical periods that control vocal learning in humans is still contested (Van Hedger et al. 2015; Werker and Hensch 2015). Zebra finch studies that causally relate neural mechanisms with vocal learning capability may therefore provide some insights into these issues as well as broader criteria for learning.

8.7 The Value of Comparative Studies

A diversity of song learning styles exists across the 5000 extant species of songbirds (Clayton et al. 2009). There is great value in leveraging the experiments nature performed, and discovery can come from comparative studies. Just like comparisons of *FoxP2* sequences between humans and nonhuman primate species suggested functional properties of the protein in vocal learning and production, the comparison of genomic features across birds with unique behavioral vocal learning and production traits can identify possible connections between genomes and behavior. The continued accumulation of genomic and transcriptomic sequences from a variety of species will bolster these investigations.

Comparative studies can address several questions. For example, are there genetic or genomic features that differentiate species of birds that can learn vocalizations from those that cannot? Second, are genomes regulated similarly across species as they transition within a lifetime from being able to learn to being closed to learning? Like zebra finches, parasite birds (birds that lay their eggs in the nests of another host species) might have restricted learning abilities to support learning of their own species' song rather than their host species' song. Alternatively, some species continue to engage in vocal learning through adulthood (see Sakata and Woolley, Chap. 1). For example, starlings can acquire song every day. Do mechanisms of the zebra finch critical period also characterize the open and closed phases of learning in other species? Third, some animals have vocal plasticity, permitting them to adjust acoustic features, such as frequency, based on their environment, but they do not have stable communication patterns as a result of experience with other individuals (i.e., vocal learning). Might the genome uncover features intermediate between those with inflexible vocal patterns and those with vocal learning? Finally, similarities and differences in genomic regulation between the different nuclei of the song circuit could elucidate key neural processes that distinguish learning from nonlearning brain areas and distinguish mechanisms that are needed for sensory, sensorimotor, and motor learning.

8.8 The Future of Genome-Brain-Behavior Investigations

The following sections outline a few avenues of research that can provide important insights into genome function with regard to complex behaviors like vocal communication.

8.8.1 *Genomic Identification of Cellular Specifications*

All cells of an individual have the same genomic sequence: how genomic function is regulated determines the transcription of genes characteristic to distinct cell types. The coarsest categorization of cell types in the brain is neurons and glia, but there are many subtypes still being discovered.

Cell types are informative for function because, for example, they can be excitatory or inhibitory, have different electrical properties, distinct sets of membrane receptors, and a range of metabolic activities. Vocal learning occurs during development, and cell types shift across the lifetime, which may be a partial explanation for restricted vocal learning capabilities despite continued experience. Finally, the different brain nuclei in the song circuit are functionally specialized, and the entire circuit is specialized compared to the rest of the brain, suggesting specific complements of cell types within each region.

Much more discovery-based research needs to be done, perhaps by taking advantage of epigenetic features such as H3K27ac-mediated epigenetic approaches for identifying regulatory regions considered enhancers. However, beautiful molecular work has been done dissecting and describing the cellular populations within Area X, a complicated composite brain area for vocal plasticity (Person et al. 2008). Another approach was a widespread initiative termed the Zebra Finch Expression Brain Atlas (ZEBrA) that catalogued patterns of gene expression across the entire adult male brain (http://www.zebrafinchatlas.org/). The result is a resource of over 600 genes (and growing) and the anatomical location of their expression that can be used to identify genes potentially specialized for particular facets of vocal learning. This type of resource pairs well with community-wide initiatives that describe the levels of gene expression across various brain areas, ages, and conditions and provide gene-based insights into neural functions (Replogle et al. 2008). Finally, methods that measure the entire population of extant RNAs from a brain area at a particular time may be useful for identifying cellular traits that relate song nuclei to human speech and language regions (Pfenning et al. 2014).

8.8.2 *Gene Editing and Genetic Manipulation Technologies*

Transitioning from descriptive to causal experimental design requires the ability to do manipulations. The zebra finch has the advantage that experiential manipulations have known functional consequence on the future pattern of song and the ability to learn song past P65. Genetic manipulation has been immensely powerful in revealing mechanisms in rodents and other animals, and the use of such manipulations is advancing in birds.

There are now three reports of transgenic zebra finches that have cells with altered genomes throughout their bodies. One set of transgenic birds expresses green fluorescent protein (GFP), which is useful for cellular anatomy (Agate et al. 2009). Another line of transgenic birds was created with manipulated activation levels of a transcription factor called CREB: these birds displayed deficits in developmental song learning (Abe et al. 2015). And finally, there are transgenics with a mutated Huntington gene (*HTT*) that display motor song performance issues interpreted as consistent with motor deficits in human Huntington's disease (Liu et al. 2015). Although currently they are highly inefficient to create and can be difficult to breed, transgenic zebra finches will be powerful tools moving forward as their creation becomes easier.

An alternative to germ-line creation of transgenics is local gene manipulations using either in vivo electroporation or viral vectors for construct delivery (Heston and White 2015; Ahmadiantehrani and London 2017b). This method has the advantage of not directly manipulating function in brain regions other than the one of interest. There are currently several ways to deliver gene constructs to manipulate brain cell function. It is possible to overexpress a gene or to mutate or eliminate a gene using CRISPR/Cas9 technology (Sander and Joung 2014). Alternatively, cells can be inhibited or activated on very short timescales by expressing optogenetic channels that use light as a ligand or on longer timescales via DREADD receptors, which were designed to have antibiotics as their ligand (Boyden 2011; Smith et al. 2016). These strategies can be combined with other elements of manipulation constructs that confer temporal and cell-type specificity (Hisey et al. 2018; Xiao et al. 2018), and they are beginning to reveal how areas of the traditional sensorimotor singing circuitry operate (Roberts et al. 2012; Tanaka et al. 2018; Xiao et al. 2018). There is much more to be done to optimize and popularize strategies for genetic manipulation, but as knowledge about genome-brain-behavior connections becomes more sophisticated, it will be essential to have these tools available to researchers asking the variety of questions that are collectively required to compile a comprehensive understanding of how genomic features contribute to vocal learning. As a last note, pharmacological agents can be beneficial alternatives to genetic manipulation especially because their effects are transient, permitting carefully timed disruptions in function during experience and in controlled situations.

8.9 Summary

Previously, researchers thought that knowing an individual's genetic sequence would unlock his/her individual biology. It is now clear that genetics is not equivalent to biological determinism; rather, the complexity of our lives alters how our genomes function, and it is the interplay between chromatin, brain, and experience that ultimately produces measurable behavioral patterns such as speech and language. This chapter summarized the basic features of chromatin structure and function and their relationship to vocal learning in humans and songbirds. Strategies to probe the chromatin-brain-behavior interdependencies were explored and discussed with respect to understanding how a complex behavior such as vocal learning emerges. An underlying goal was to demonstrate how cross-species investigations—human to avian and across avian species—can be fruitful when open-minded researchers take advantage of the unique and shared properties of genomes to create meaningful research comparisons. Investigations that focus on various components of genomic function will advance recognition of the fundamental shared features across species that go deeper than superficial parallels.

Compliance with Ethics Requirements Sarah E. London declares that she has no conflict of interest.

References

Abe K, Matsui S, Watanabe D (2015) Transgenic songbirds with suppressed or enhanced activity of CREB transcription factor. Proc Natl Acad Sci USA 112(24):7599–7604. https://doi.org/10.1073/pnas.1413484112

Adam I, Mendoza E, Kobalz U, Wohlgemuth S, Scharff C (2017) CNTNAP2 is a direct FoxP2 target in vitro and in vivo in zebra finches: complex regulation by age and activity. Genes Brain Behav 16(6):635–642. https://doi.org/10.1111/gbb.12390

Adkins-Regan E, Abdelnabi M, Mobarak M, Ottinger MA (1990) Sex steroid levels in developing and adult male and female zebra finches (*Poephila guttata*). Gen Comp Endo 78(1):93–109. https://doi.org/10.1016/0016-6480(90)90051-M

Adkins-Regan E, Mansukhani V, Seiwert C, Thompson R (1994) Sexual differentiation of brain and behavior in the zebra finch: critical periods for effects of early estrogen treatment. J Neurobiol 25(7):865–877. https://doi.org/10.1002/neu.480250710

Agate RJ, Grisham W, Wade J, Mann S, Wingfield J, Schanen C, Palotie A, Arnold AP (2003) Neural, not gonadal, origin of brain sex differences in a gynandromorphic finch. Proc Natl Acad Sci U S A 100(8):4873–4878. https://doi.org/10.1073/pnas.0636925100

Agate RJ, Scott BB, Haripal B, Lois C, Nottebohm F (2009) Transgenic songbirds offer an opportunity to develop a genetic model for vocal learning. Proc Natl Acad Sci U S A 106(42):17963–17967. https://doi.org/10.1073/pnas.0909139106

Ahmadiantehrani S, London SE (2017a) Bidirectional manipulation of mTOR signaling disrupts socially mediated vocal learning in juvenile songbirds. Proc Natl Acad Sci U S A 114(35):9463–9468. https://doi.org/10.1073/pnas.1701829114

Ahmadiantehrani S, London SE (2017b) A reliable and flexible gene manipulation strategy in posthatch zebra finch brain. Sci Rep 7:43244. https://doi.org/10.1038/srep43244

Ahmadiantehrani S, Gores EO, London SE (2018) A complex mTOR response in habituation paradigms for a social signal in adult songbirds. Learn Mem 25(6):273–282. https://doi.org/10.1101/lm.046417.117

Alberini CM (2009) Transcription factors in synaptic plasticity and learning and memory A2 - squire, Larry R. In: Encyclopedia of neuroscience. Academic, Oxford, pp 1081–1092. https://doi.org/10.1016/B978-008045046-9.00828-7

Allis CD, Jenuwein T (2016) The molecular hallmarks of epigenetic control. Nat Rev Genet 17(8):487–500. https://doi.org/10.1038/nrg.2016.59

Arnold AP (1975) The effects of castration and androgen replacement on song, courtship, and aggression in zebra finches (*Poephila guttata*). J Exp Zool 191(3):309–326. https://doi.org/10.1002/jez.1401910302

Arnold AP, Chen X (2009) What does the “four core genotypes” mouse model tell us about sex differences in the brain and other tissues? Front Neuroendocrinol 30(1):1–9. https://doi.org/10.1016/j.yfrne.2008.11.001

Arnold AP, Schlinger BA (1993) Sexual differentiation of brain and behavior: the zebra finch is not just a flying rat. Brain Behav Evol 42(4–5):231–241. https://doi.org/10.1159/000114157

Arnold AP, Xu J, Grisham W, Chen X, Kim YH, Itoh Y (2004) Minireview: sex chromosomes and brain sexual differentiation. Endocrinology 145(3):1057–1062. https://doi.org/10.1210/en.2003-1491

Aronov D, Andalman AS, Fee MS (2008) A specialized forebrain circuit for vocal babbling in the juvenile songbird. Science (New York, NY) 320(5876):630–634. https://doi.org/10.1126/science.1155140

Barbu S, Nardy A, Chevrot J-P, Guellaï B, Glas L, Juhel J, Lemasson A (2015) Sex differences in language across early childhood: family socioeconomic status does not impact boys and girls equally. Front Psychol 6:1874–1874. https://doi.org/10.3389/fpsyg.2015.01874

Bolhuis JJ, Okanoya K, Scharff C (2010) Twitter evolution: converging mechanisms in birdsong and human speech. Nat Rev Neurosci 11(11):747–759. https://doi.org/10.1038/nrn2931

Boyden ES (2011) A history of optogenetics: the development of tools for controlling brain circuits with light. F1000 Biol Rep 3:11. https://doi.org/10.3410/B3-11

Bozon B, Davis S, Laroche S (2002) Regulated transciption of the immediate-early gene Zif268: mechanisms and gene dosage-dependent function in synaptic plasticity and memory formation. Hippocampus 12(5):570–577. https://doi.org/10.1002/hipo.10100

Braaten RF, Petzoldt M, Colbath A (2006) Song perception during the sensitive period of song learning in zebra finches (*Taeniopygia guttata*). J Comp Psychol 120(2):79–88. https://doi.org/10.1037/0735-7036.120.2.79

Braaten RF, Miner SS, Cybenko AK (2008) Song recognition memory in juvenile zebra finches: effects of varying the number of presentations of heterospecific and conspecific songs. Behav Proc 77(2):177–183. https://doi.org/10.1016/j.beproc.2007.10.010

Burkett ZD, Day NF, Kimball TH, Aamodt CM, Heston JB, Hilliard AT, Xiao X, White SA (2018) FoxP2 isoforms delineate spatiotemporal transcriptional networks for vocal learning in the zebra finch. eLife 7:e30649. https://doi.org/10.7554/eLife.30649

Carruth LL, Reisert I, Arnold AP (2002) Sex chromosome genes directly affect brain sexual differentiation. Nat Neurosci 5(10):933–934. https://doi.org/10.1038/nn922

Cheng HY, Clayton DF (2004) Activation and habituation of extracellular signal-regulated kinase phosphorylation in zebra finch auditory forebrain during song presentation. J Neurosci 24(34):7503–7513. https://doi.org/10.1523/JNEUROSCI.1405-04.2004

Clayton DF (2000) The genomic action potential. Neurobiol Learn Mem 74(3):185–216. https://doi.org/10.1006/nlme.2000.3967

Clayton DF, Balakrishnan CN, London SE (2009) Integrating genomes, brain and behavior in the study of songbirds. Curr Biol 19(18):R865–R873. https://doi.org/10.1016/j.cub.2009.07.006

Cohen NJ, Vallance DD, Barwick M, Im N, Menna R, Horodezky NB, Isaacson L (2000) The Interface between ADHD and Language Impairment: An Examination of Language, Achievement, and Cognitive Processing. J Child Psychol Psychiatry and Allied Disc 41 (3):353–362. doi:undefined

Condro MC, White SA (2014) Distribution of language-related Cntnap2 protein in neural circuits critical for vocal learning. J Comp Neurol 522(1):169–185. https://doi.org/10.1002/cne.23394

Dewulf V, Bottjer SW (2005) Neurogenesis within the juvenile zebra finch telencephalic ventricular zone: a map of proliferative activity. J Comp Neurol 481(1):70–83. https://doi.org/10.1002/cne.20352

Dong S, Clayton DF (2008) Partial dissociation of molecular and behavioral measures of song habituation in adult zebra finches. Genes Brain Behav 7(7):802–809. https://doi.org/10.1111/j.1601-183X.2008.00423.x

Dong S, Clayton DF (2009) Habituation in songbirds. Neurobiol Learn Mem 92(2):183–188. https://doi.org/10.1016/j.nlm.2008.09.009

Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. Annu Rev Neurosci 22:567–631. https://doi.org/10.1146/annurev.neuro.22.1.567

Dugas-Ford J, Rowell JJ, Ragsdale CW (2012) Cell-type homologies and the origins of the neocortex. Proc Natl Acad Sci U S A 109(42):16974–16979. https://doi.org/10.1073/pnas.1204773109

Eales LA (1985) Song learning in zebra finches - some effects of song model availability on what is learnt and when. An Behav 33 (Nov):1293–1300. doi:Doi https://doi.org/10.1016/S0003-3472(85)80189-5

Eales LA (1987) Song learning in female-raised zebra finches - Another look at the sensitive phase. An Behav 35:1356–1365. doi:Doi https://doi.org/10.1016/S0003-3472(87)80008-8

Enard W, Przeworski M, Fisher SE, Lai CSL, Wiebe V, Kitano T, Monaco AP, Pääbo S (2002) Molecular evolution of FOXP2, a gene involved in speech and language. Nature 418:869–872. https://doi.org/10.1038/nature01025

Etchell A, Adhikari A, Weinberg LS, Choo AL, Garnett EO, Chow HM, Chang SE (2018) A systematic literature review of sex differences in childhood language and brain development. Neuropsychologia 114:19–31. https://doi.org/10.1016/j.neuropsychologia.2018.04.011

Fisher SE, Scharff C (2009) FOXP2 as a molecular window into speech and language. Trends Genetics 25(4):166–177. https://doi.org/10.1016/j.tig.2009.03.002

Fisher SE, Vargha-Khadem F, Watkins KE, Monaco AP, Pembrey ME (1998) Localisation of a gene implicated in a severe speech and language disorder. Nat Genet 18(2):168–170. https://doi.org/10.1038/ng0298-168

Fu Y, Dominissini D, Rechavi G, He C (2014a) Gene expression regulation mediated through reversible m6A RNA methylation. Nat Rev Genetics 15:293–306. https://doi.org/10.1038/nrg3724

Fu L, Shi Z, Luo G, Tu W, Wang X, Fang Z, Li X (2014b) Multiple microRNAs regulate human FOXP2 gene expression by targeting sequences in its 3′ untranslated region. Mol Brain 7:71. https://doi.org/10.1186/s13041-014-0071-0

Gahr M, Metzdorf R (1999) The sexually dimorphic expression of androgen receptors in the song nucleus Hyperstriatalis Ventrale pars Caudale of the zebra finch develops independently of gonadal steroids. J Neurosci 19(7):2628–2636. https://doi.org/10.1523/jneurosci.19-07-02628.1999

Gebert LFR, MacRae IJ (2018) Regulation of microRNA function in animals. Nat Rev Mol Cell Biol 20:21–37. https://doi.org/10.1038/s41580-018-0045-7

Gervain J, Vines BW, Chen LM, Seo RJ, Hensch TK, Werker JF, Young AH (2013) Valproate reopens critical-period learning of absolute pitch. Front Syst Neurosci 7:102. https://doi.org/10.3389/fnsys.2013.00102

Gilad Y, Oshlack A, Smyth GP, Speed TP, White K (2006) Expression profiling in primates reveals a rapid evolution of human transcription factors. Nature 440:242–245. https://doi.org/10.1038/nature04559

Gobes SMH, Jennings RB, Maeda RK (2017) The sensitive period for auditory-vocal learning in the zebra finch: consequences of limited-model availability and multiple-tutor paradigms on song imitation. Behav Process 163:5–12. https://doi.org/10.1016/j.beproc.2017.07.007

Gunaratne PH, Lin YC, Benham AL, Drnevich J, et al. (2011) Song exposure regulates known and novel microRNAs in the zebra finch auditory forebrain. BMC genomics 12, doi.org/10.1186/1471-2164-12-277

Haesler S, Wada K, Nshdejan A, Morrisey EE, Lints T, Jarvis ED, Scharff C (2004) FoxP2 expression in avian vocal learners and non-learners. J Neurosci 24(13):3164–3175. https://doi.org/10.1523/JNEUROSCI.4369-03.2004

Haesler S, Rochefort C, Georgi B, Licznerski P, Osten P, Scharff C (2007) Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus area X. PLoS Biol 5(12):e321. https://doi.org/10.1371/journal.pbio.0050321

Halladay AK, Bishop S, Constantino JN, Daniels AM, Koenig K, Palmer K, Messinger D, Pelphrey K, Sanders SJ, Singer AT, Taylor JL, Szatmari P (2015) Sex and gender differences in autism spectrum disorder: summarizing evidence gaps and identifying emerging areas of priority. Mol Autism 6(1):36. https://doi.org/10.1186/s13229-015-0019-y

Hammock EA, Young LJ (2005) Microsatellite instability generates diversity in brain and sociobehavioral traits. Science 308:1630–1634. https://doi.org/10.1126/science.1111427

Heston JB, White SA (2015) Behavior-linked foxP2 regulation enables zebra finch vocal learning. J Neurosci 35(7):2885–2894. https://doi.org/10.1523/JNEUROSCI.3715-14.2015

Heston JB, Jt S, Day NF, Coleman MJ, White SA (2018) Bidirectional scaling of vocal variability by an avian cortico-basal ganglia circuit. Physiol Rep 6(8):e13638–e13638. https://doi.org/10.14814/phy2.13638

Hisey E, Kearney MG, Mooney R (2018) A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. Nat Neurosci 21(4):589–597

Hoeffer CA, Klann E (2010) mTOR signaling: at the crossroads of plasticity, memory and disease. Trends Neurosci 33(2):67–75. https://doi.org/10.1016/j.tins.2009.11.003

Hollins SL, Cairns MJ (2016) MicroRNA: small RNA mediators of the brains genomic response to environmental stress. Prog Neurobiol 143:61–81. https://doi.org/10.1016/j.pneurobio.2016.06.005

Horita H, Wada K, Rivas MV, Hara E, Jarvis ED (2010) The dusp1 immediate early gene is regulated by natural stimuli predominantly insensory input neurons. J Comp Neurol 518(14):2873–2901. https://doi.org/10.1002/cne.22370

Itoh Y, Melamed E, Yang X, Kampf K, Wang S, Yehya N, Van Nas A, Replogle K, Band MR, Clayton DF, Schadt EE, Lusis AJ, Arnold AP (2007) Dosage compensation is less effective in birds than in mammals. J Biol 6(1):2. https://doi.org/10.1186/jbiol53

Jarvis ED, Nottebohm F (1997) Motor-driven gene expression. Proc Natl Acad Sci U S 94(8):4097–4102. https://doi.org/10.1073/pnas.94.8.4097

Jarvis ED, Scharff C, Grossman MR, Ramos JA, Nottebohm F (1998) For whom the bird sings: context-dependent gene expression. Neuron 21(4):775–788. https://doi.org/10.1016/S0896-6273(00)80594-2

Jin H, Clayton DF (1997) Localized changes in immediate-early gene regulation during sensory and motor learning in zebra finches. Neuron 19(5):1049–1059. https://doi.org/10.1016/S0896-6273(00)80396-7

Karten HJ (2013) Neocortical evolution: neuronal circuits arise independently of lamination. Curr Biol 23(1):R12–R15. https://doi.org/10.1016/j.cub.2012.11.013

Kelly TK, Ahmadiantehrani S, Blattler A, London SE (2018) Epigenetic regulation of transcriptional plasticity associated with developmental song learning. Proc Royal Soc B: Biol Sci 285(1878):20180160. https://doi.org/10.1098/rspb.2018.0160

Kim D, Sung YM, Park J, Kim S, Kim J, Park J, Ha H, Bae JY, Kim S, Baek D (2016) General rules for functional microRNA targeting. Nat Genetics 48:1517–1526. https://doi.org/10.1038/ng.3694

Kim YH, Perlman WR, Arnold AP (2004) Expression of androgen receptor mRNA in zebra finch song system: developmental regulation by estrogen. J Comp Neurol 469(4):535–547. https://doi.org/10.1002/cne.11033

Kimpo RR, Doupe AJ (1997) FOS is induced by singing in distinct neuronal populations in a motor network. Neuron 18(2):315–325. https://doi.org/10.1016/S0896-6273(00)80271-8

Knudsen EI (2004) Sensitive periods in the development of the brain and behavior. J Cog Neurosci 16(8):1412–1425. https://doi.org/10.1162/0898929042304796

Kreitzer AC (2009) Physiology and pharmacology of striatal neurons. Ann Rev Neurosci 32(1):127–147. https://doi.org/10.1146/annurev.neuro.051508.135422

Kruse AA, Stripling R, Clayton DF (2004) Context-specific habituation of the zenk gene response to song in adult zebra finches. Neurobiol Learn Mem 82(2):99–108. https://doi.org/10.1016/j.nlm.2004.05.001

Kvande MN, Belsky J, Wichstrøm L (2018) Selection for special education services: the role of gender and socio-economic status. European J Sp Needs Edu 33(4):510–524. https://doi.org/10.1080/08856257.2017.1373493

Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. Nature 413(6855):519–523. https://doi.org/10.1038/35097076

Lin D, Hong P, Zhang S, Xu W, Jamal M, Yan K, Lei Y, Li L, Ruan Y, Fu ZF, Li G, Cao G (2018) Digestion-ligation-only hi-C is an efficient and cost-effective method for chromosome conformation capture. Nat Genetics 50(5):754–763. https://doi.org/10.1038/s41588-018-0111-2

Lin LC, Vanier DR, London SE (2014) Social information embedded in vocalizations induces neurogenomic and behavioral responses. PLoS One 9(11):e112905. https://doi.org/10.1371/journal.pone.0112905

Liu W, Kohn J, Szwed SK, Pariser E, Sepe S, Haripal B, Oshimori N, Marsala M, Miyanohara A, Lee R (2015) Human mutant huntingtin disrupts vocal learning in transgenic songbirds. Nat Neurosci 18(11):1617–1622. https://doi.org/10.1038/nn.4133

London SE (2016) Influences of non-canonical neurosteroid signaling on developing neural circuits. Curr Opin Neurobiol 40:103–110. https://doi.org/10.1016/j.conb.2016.06.018

London SE (2017) Developmental song learning as a model to understand neural mechanisms that limit and promote the ability to learn. Behav Process 163:13–23. https://doi.org/10.1016/j.beproc.2017.11.008

London SE, Clayton DF (2008) Functional identification of sensory mechanisms required for developmental song learning. Nat Neurosci 11(5):579–586. https://doi.org/10.1038/nn.2103

London SE, Schlinger BA (2007) Steroidogenic enzymes along the ventricular proliferative zone in the developing songbird brain. J Comp Neurol 502(4):507–521. https://doi.org/10.1002/cne.21335

London SE, Monks DA, Wade J, Schlinger BA (2006) Widespread capacity for steroid synthesis in the avian brain and song system. Endocrinology 147(12):5975–5987. https://doi.org/10.1210/en.2006-0154

London SE, Dong S, Replogle K, Clayton DF (2009a) Developmental shifts in gene expression in the auditory forebrain during the sensitive period for song learning. Dev Neurobiol 69(7):437–450. https://doi.org/10.1002/dneu.20719

London SE, Remage-Healey L, Schlinger BA (2009b) Neurosteroid production in the songbird brain: a re-evaluation of core principles. Front Neuroendocrinol 30(3):302–314. https://doi.org/10.1016/j.yfrne.2009.05.001

Luo GZ, Hafner M, Shi Z, Brown M, Feng GH, Tuschl T, Wang XJ, Li X (2012) Genome-wide annotation and analysis of zebra finch microRNA repertoire reveal sex-biased expression. BMC Genomics 13(1):727. https://doi.org/10.1186/1471-2164-13-727

MacDermot KD, Bonora E, Sykes N, Coupe AM, Lai CS, Vernes SC, Vargha-Khadem F, McKenzie F, Smith RL, Monaco AP, Fisher SE (2005) Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. Am J Hum Genet 76(6):1074–1080. https://doi.org/10.1086/430841

Marty V, Cavaillé J (2019) Imprinted small noncoding RNA genes in brain function and behaviour. Curr Opin Behav Sci 25:8–14. https://doi.org/10.1016/j.cobeha.2018.05.009

McQuown SC, Wood MA (2011) HDAC3 and the molecular brake pad hypothesis. Neurobiol Learn Mem 96(1):27–34. https://doi.org/10.1016/j.nlm.2011.04.002

McQuown SC, Barrett RM, Matheos DP, Post RJ, Rogge GA, Alenghat T, Mullican SE, Jones S, Rusche JR, Lazar MA, Wood MA (2011) HDAC3 is a critical negative regulator of long-term memory formation. J Neurosci 31(2):764–774. https://doi.org/10.1523/JNEUROSCI.5052-10.2011

Mello CV, Velho TAF, Pinaud R (2004) Song-induced gene expression: a window on wong auditory processing and perception. Ann N Y Acad Sci 1016(1):263–281. https://doi.org/10.1196/annals.1298.021

Minatohara K, Akiyoshi M, Okuno H (2016) Role of immediate-early genes in synaptic plasticity and neuronal ensembles underlying the memory trace. Front Mol Neurosci 8:78–78. https://doi.org/10.3389/fnmol.2015.00078

Mooney R, Konishi M (1991) Two distinct inputs to an avian song nucleus activate different glutamate receptor subtypes on individual neurons. Proc Natl Acad Sci U S A 88(10):4075–4079

Mori C, Wada K (2015) Audition-independent vocal crystallization associated with intrinsic developmental gene expression dynamics. J Neurosci 35(3):878–889. https://doi.org/10.1523/JNEUROSCI.1804-14.2015

Murugan M, Harward S, Scharff C, Mooney R (2013) Diminished FoxP2 levels affect dopaminergic modulation of corticostriatal signaling important to song variability. Neuron 80(6):1464–1476. https://doi.org/10.1016/j.neuron.2013.09.021

Nguyen TC, Zaleta-Rivera K, Huang X, Dai X, Zhong S (2018) RNA, action through interactions. Trends Genetics 34(11):867–882. https://doi.org/10.1016/j.tig.2018.08.001

Nottebohm F, Arnold AP (1976) Sexual dimorphism in vocal control areas of the songbird brain. Science 194(4261):211–213. https://doi.org/10.1126/science.959852

Nudel R, Newbury DF (2013) FOXP2. Wiley Interdiscip Rev Cogn Sci 4(5):547–560. https://doi.org/10.1002/wcs.1247

O'Brien J, Hayder H, Zayed Y, Peng C (2018) Overview of microRNA biogenesis, mechanisms of actions, and circulation. Front Endocrinol 9:402–402. https://doi.org/10.3389/fendo.2018.00402

O'Leary MA, Bloch JI, Flynn JJ, Gaudin TJ, Giallombardo A, Giannini NP, Goldberg SL, Kraatz BP, Luo Z-X, Meng J, Ni X, Novacek MJ, Perini FA, Randall ZS, Rougier GW, Sargis EJ, Silcox MT, Simmons NB, Spaulding M, Velazco PM, Weksler M, Wible JR, Cirranello AL (2013) The placental mammal ancestor and the post–K-Pg radiation of placentals. Science 339(6120):662–667. https://doi.org/10.1126/science.1229237

Panaitof SC, Abrahams BS, Dong H, Geschwind DH, White SA (2010) Language-related Cntnap2 gene is differentially expressed in sexually dimorphic song nuclei essential for vocal learning in songbirds. J Comp Neurol 518(11):1995–2018. https://doi.org/10.1002/cne.22318

Person AL, Gale SD, Farries MA, Perkel DJ (2008) Organization of the songbird basal ganglia, including area X. J Comp Neurol 508(5):840–866. https://doi.org/10.1002/cne.21699

Petkov CI, Jarvis ED (2012) Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. Front Evol Neurosci 4:12. https://doi.org/10.3389/fnevo.2012.00012

Pfenning AR, Hara E, Whitney O, Rivas MV, Wang R, Roulhac PL, Howard JT, Wirthlin M, Lovell PV, Ganapathy G, Mountcastle J, Moseley MA, Thompson JW, Soderblom EJ, Iriki A, Kato M, Gilbert MTP, Zhang G, Bakken T, Bongaarts A, Bernard A, Lein E, Mello CV, Hartemink AJ, Jarvis ED (2014) Convergent transcriptional specializations in the brains of humans and song-learning birds. Science 346(6215):1256846–1256846. https://doi.org/10.1126/science.1256846

Phan ML, Gergues MM, Mahidadia S, Jimenez-Castillo J, Vicario DS, Bieszczad KM (2017) HDAC3 inhibitor RGFP966 modulates neuronal memory for vocal communication signals in a songbird model. Front Sys Neurosci 11:65–65. https://doi.org/10.3389/fnsys.2017.00065

Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, Lemmon AR (2016) A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature 534:S7–S8. https://doi.org/10.1038/nature19417

Rankin CH, Abrams T, Barry RJ, Bhatnagar S, Clayton DF, Colombo J, Coppola G, Geyer MA, Glanzman DL, Marsland S, McSweeney FK, Wilson DA, Wu CF, Thompson RF (2009) Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. Neurobiol Learn Mem 92(2):135–138. https://doi.org/10.1016/j.nlm.2008.09.012

Replogle K, Arnold AP, Ball GF, Band M, Bensch S, Brenowitz EA, Dong S, Drnevich J, Ferris M, George JM, Gong G, Hasselquist D, Hernandez AG, Kim R, Lewin HA, Liu L, Lovell PV, Mello CV, Naurin S, Rodriguez-Zas S, Thimmapuram J, Wade J, Clayton DF (2008) The songbird Neurogenomics (SoNG) initiative: community-based tools and strategies for study of brain gene function and evolution. BMC Genomics 9:131. https://doi.org/10.1186/1471-2164-9-131

Roberts TF, Gobes SMH, Murugan M, Ölveczky BP, Mooney R (2012) Motor circuits are required to encode a sensory model for imitative learning. Nat Neurosci 15(10):1454–1459. https://doi.org/10.1038/nn.3206

Roper A, Zann R (2006) The onset of song learning and song tutor selection in fledgling zebra finches. Ethology 112(5):458–470. https://doi.org/10.1111/J.1439-0310.2005.01169.X

Roth BL (2016) DREADDs for neuroscientists. Neuron 89(4):683–694. https://doi.org/10.1016/j.neuron.2016.01.040

Sander JD, Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotech 32:347–355. https://doi.org/10.1038/nbt.2842

Scharff C, Adam I (2013) Neurogenetics of birdsong. Curr Opin Neurobiol 23(1):29–36. https://doi.org/10.1016/j.conb.2012.10.001

Scharff C, Haesler S (2005) An evolutionary perspective on FoxP2: strictly for the birds? Curr Opin Neurobiol 15(6):694–703. https://doi.org/10.1016/j.conb.2005.10.004

Schlinger BA, Arnold AP (1992) Circulating estrogens in a male songbird originate in the brain. Proc Natl Acad Sci U S A 89(16):7650–7653. https://doi.org/10.1073/pnas.89.16.7650

Shi Z, Luo G, Fu L, Fang Z, Wang X, Li X (2013) miR-9 and miR-140-5p target FoxP2 and are regulated as a function of the social context of singing behavior in zebra finches. J Neurosci 33(42):16510–16521. https://doi.org/10.1523/jneurosci.0838-13.2013

Shukla GC, Singh J, Barik S (2011) MicroRNAs: processing, maturation. target recognition and regulatory functions Mol Cell Pharmacol 3(3):83–92

Smith KS, Bucci DJ, Luikart BW, Mahler SV (2016) DREADDS: use and application in behavioral neuroscience. Behav Neurosci 130(2):137–155. https://doi.org/10.1037/bne0000135

Strahl BD, Allis CD (2000) The language of covalent histone modifications. Nature 403(6765):41–45

Takesian AE, Hensch TK (2013) Balancing plasticity/stability across brain development. Prog Brain Res 207:3–34. https://doi.org/10.1016/B978-0-444-63327-9.00001-1

Tanaka M, Sun F, Li Y, Mooney R (2018) A mesocortical dopamine circuit enables the cultural transmission of vocal behaviour. Nature 563(7729):117–120. https://doi.org/10.1038/s41586-018-0636-7

Teramitsu I, Kudo LC, London SE, Geschwind DH, White SA (2004) Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction. J Neurosci 24(13):3152–3163. https://doi.org/10.1523/JNEUROSCI.5589-03.2004

The ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489:57–74. https://doi.org/10.1038/nature11247

Thompson RF, Spencer WA (1966) Habituation: a model phenomenon for the study of neuronal substrates of behavior. Psych Rev 73(1):16–43. https://doi.org/10.1037/h0022681

Tischmeyer W, Grimm R (1999) Activation of immediate early genes and memory formation. Cell Mol Life Sci CMLS 55(4):564–574. https://doi.org/10.1007/s000180050315

Urban DJ, Roth BL (2015) DREADDs (designer receptors exclusively activated by designer drugs): Chemogenetic tools with therapeutic utility. Ann Rev Pharm Toxicol 55(1):399–417. https://doi.org/10.1146/annurev-pharmtox-010814-124803

Van Hedger SC, Heald SLM, Koch R, Nusbaum HC (2015) Auditory working memory predicts individual differences in absolute pitch learning. Cognition 140:95–110. https://doi.org/10.1016/j.cognition.2015.03.012

Vargha-Khadem F, Watkins KE, Price CJ, Ashburner J et al (1998) Neural basis of an inherited speech and language disorder. Proc Natl Acad Sci U S A 95(21):12695–12700. https://doi.org/10.1073/pnas.95.21.12695

Vates GE, Broome BM, Mello CV, Nottebohm F (1996) Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches. J Comp Neurol 366(4):613–642. https://doi.org/10.1002/(SICI)1096-9861(19960318)366:4. <613::AID-CNE5>3.0.CO;2-7

Venter JC, Adams MD, Myers EW, Li PW et al (2001) The sequence of the human genome. Science 291(5507):1304–1351. https://doi.org/10.1126/science.1058040

Vignal C, Andru J, Mathevon N (2005) Social context modulates behavioural and brain immediate early gene responses to sound in male songbird. Eur J Neurosci 22(4):949–955. https://doi.org/10.1111/j.1460-9568.2005.04254.x

Wade J, Arnold AP (2004) Sexual differentiation of the zebra finch song system. Ann N Y Acad Sci 1016:540–559. https://doi.org/10.1196/annals.1298.015

Walder DJ, Seidman LJ, Cullen N, Su J, Tsuang MT, Goldstein JM (2006) Sex differences in language dysfunction in schizophrenia. AmJ psychiatry 163(3):470–477. https://doi.org/10.1176/appi.ajp.163.3.470

Wallentin M (2009) Putative sex differences in verbal abilities and language cortex: a critical review. Brain Lang 108:175–183. https://doi.org/10.1016/j.bandl.2008.07.001

Wang W, Kwon EJ, Tsai L-H (2012) MicroRNAs in learning, memory, and neurological diseases. Learn Mem 19(9):359–368. https://doi.org/10.1101/lm.026492.112

Warren WC, Clayton DF, Ellegren H, Arnold AP et al (2010) The genome of a songbird. Nature 464(7289):757–762. https://doi.org/10.1038/nature08819

Werker JF, Hensch TK (2015) Critical periods in speech perception: new directions. Annu Rev Psychol 66:173–196. https://doi.org/10.1146/annurev-psych-010814-015104

Whitney O, Soderstrom K, Johnson F (2000) Post-transcriptional regulation of zenk expression associated with zebra finch vocal development. Brain Res Mol Brain Res 80(2):279–290. https://doi.org/10.1016/S0169-328X(00)00178-9

Whitney O, Pfenning AR, Howard JT, Blatti CA et al (2014) Core and region-enriched networks of behaviorally regulated genes and the singing genome. Science 346(6215):1256780. https://doi.org/10.1126/science.1256780

Woolley SC, Doupe AJ (2008) Social context–induced song variation affects female behavior and gene expression. PLoS Biol 6(3):e62. https://doi.org/10.1371/journal.pbio.0060062

Xiao L, Chattree G, Oscos FG, Cao M, Wanat MJ, Roberts TF (2018) A Basal Ganglia circuit sufficient to guide birdsong learning. Neuron 98(1):208–221. e205. https://doi.org/10.1016/j.neuron.2018.02.020

Zann RA (1996) The Zebra finch: a synthesis of Field and laboratory studies. Oxford University Press, New York

Zhao BS, Roundtree IA, He C (2016) Post-transcriptional gene regulation by mRNA modifications. Nat Rev Mol Cell Biol18:31. doi:https://doi.org/10.1038/nrm.2016.132

Chapter 9
Vocal Performance in Songbirds: From Mechanisms to Evolution

Jeffrey Podos and Ha-Cheol Sung

Abstract Song production across many songbird species is shaped by largely conserved sets of constraints associated with song control, production, and perception. This chapter addresses two main hypotheses: (1) that mechanisms of birdsong production and performance set constraints on vocal structure in ways that run parallel across taxa and that help to explain the evolution of vocal phenotypes; and (2) that resulting variations in birdsong structure influence communicative functions for both intrasexual and intersexual interactions in ways that favor the evolution of vocal phenotypes that challenge vocal performance capacities. Substantial evidence for both hypotheses has accumulated, showing that performance constraints can influence the structure of diverse vocal features and that birds indeed attend and respond differentially to performance-based vocal variations. Topics highlighted for future work include: (1) exploration of neural and endocrine mechanisms that underlie vocal performance, (2) elucidation of cognitive and sampling processes that underlie song performance assessment, and (3) evaluation of the overall impact of song performance on the evolution of songbird vocal phenotypes.

Keywords Bioacoustics · Birdsong · Female preferences · Sexual selection · Song perception · Vocal communication · Vocal evolution

9.1 Introduction

A main goal of this volume is to explore the mechanistic bases of bird vocal communication with special attention to neural control, auditory processing, genetic architecture, computational challenges, and cognition. This chapter considers

---

J. Podos (✉)
Department of Biology, University of Massachusetts, Amherst, MA, USA
e-mail: jpodos@bio.umass.edu

H.-C. Sung
Department of Biological Sciences, Chonnam National University, Gwangju, South Korea
e-mail: shcol2002@chonnam.ac.kr

J. T. Sakata et al. (eds.), *The Neuroethology of Birdsong*, Springer Handbook of Auditory Research 71, https://doi.org/10.1007/978-3-030-34683-6_9

birdsong from an evolutionary perspective. Bird songs are phenotypic traits that have evolved over eons and whose parameters reflect a broad palette of selective pressures, mechanistic constraints, and phylogenetic histories. Accordingly, mechanisms that underlie song learning, production, and perception ultimately evolve in service of the vocal phenotype, which is the principal locus of song evolution. The main theme developed in this chapter is that evolutionary and mechanistic studies of bird vocal communication can be mutually informative when singing behavior is assessed across multiple levels of analysis (see Sakata and Woolley, Chap. 1). More precisely, this chapter develops the case that recent studies of vocal performance and function offer new and profitable insights into the processes that shape vocal evolution.

As is often the case in animal behavior, a good place to begin an inquiry into birdsong is to consider the natural field context. Birds produce many kinds of sounds. Among these, *songs* typically function in mating contexts, especially in territorial defense and mate attraction (Marler and Slabbekoorn 2004; Webster and Podos 2018). By singing, birds are able to broadcast their occupancy of a territory and their availability as a prospective mate to fellow conspecifics over long distances (Catchpole and Slater 2008). In this light, bird songs could, in theory, have evolved to be structurally simple. Yet bird songs are, on the whole, highly diverse and complex, presumably due to a history of sexual selection favoring these attributes. In particular, diverse and complex songs can help singers to: (1) stand out from the flock, insofar as song transmits honest information about singer quality (e.g., Hasselquist et al. 1996; Byers et al. 2016); (2) negotiate territorial disputes and convey aggressive intent with enhanced precision (Searcy and Beecher 2009; Vehrencamp et al. 2014); and (3) more effectively tap into aesthetic preferences of prospective mates (Prum 2012; Ryan 2018).

Against this backdrop of sexual selection, song evolution is shaped by a host of additional factors. Song diversity and complexity can be enhanced by selection against costly hybrid matings, which is thought to drive sympatric species toward distinct song structures (Nelson and Marler 1990; Luther 2009). Learned song components can diverge in myriad ways through cultural evolution as song models are copied with incomplete precision across generations. Cultural evolution presumably favors songs with acoustic structures that transmit well in a given environment; for example, songs with low frequencies are favored in forested habitats (Brumm and Naguib 2009; Peters et al. 2012). Other factors that drive song evolution are independent of the receiver, including genetic and cultural drift as well as natural selection on vocal morphology for functions other than communication (e.g., selection on beaks for feeding performance: Podos 2001; Derryberry et al. 2012). As a result of all of these factors, songs have diverged to the point where nearly every songbird species produces songs that are acoustically distinct, and many species show substantial vocal diversity both within and across populations. In addition, many individuals show marked acoustic variation across multiple song types in their repertoires.

Given this background, attempts to explain song evolution would seem to demand focus on vocal plasticity and its role in generating novel phenotypes. What mechanistic factors, for example, would enable some species to evolve expanded vocal repertoires? Why should some species evolve to be open-ended learners, enabling song learning throughout adulthood, whereas other species evolve to be close-ended learners, with song imitation restricted to critical periods in development? What factors favor the evolution of vocal dialects? While these and related questions are central to the modern study of birdsong, a focus on explaining diversity itself might not provide the most productive route for identifying common rules that govern song evolution. This is because each instance of song divergence, perhaps best visualized in a phylogenetic framework (e.g., ten Cate 2004), might be regarded as a unique evolutionary event triggered by highly variable sets of factors specific to the lineage, time, and place in question. That is, pressures within bird lineages that drive the evolution of complex repertoires (or learning programs) likely differ on a case-by-case basis across and within species. For example, song evolution should vary in accordance with wide-ranging natural variation in local habitats, social structures, territory sizes and compositions, biotic communities, and related factors that shape how song functions in natural contexts. In other words, explaining a thousand cases of song divergence will probably require a thousand distinct explanations.

A complementary approach for explaining song evolution, which is the focus of this chapter, recognizes the highly conserved nature of physiological systems that govern vocal control, production, and perception (e.g., Brainard and Doupe 2002; Fitch and Suthers 2016). The mechanistic commonalities across songbirds are striking: all birds sing using their syrinx; the syrinx in all birds is activated by airflow from the lungs; across the oscine songbirds, the activity of the syrinx and supporting vocal motor systems is mediated by homologous forebrain circuits; the tonal structure of song in all species is presumably modified in parallel ways by vocal tract resonances; and auditory perceptual systems are highly conserved within and across bird groups. The conserved nature of mechanical elements in song production and perception invokes common constraints on how song can be shaped across distinct evolutionary events. These constraints, in turn, impose a common barrier on the evolution of song complexity and diversity. Thus, a thousand cases of mechanistic constraints on song divergence might require only a handful of explanations.

The primary aim of this chapter is to review recent studies, with an emphasis on work in the past decade, that address two hypotheses central to the study of mechanistic constraints on song evolution:

1. Mechanisms of birdsong production and performance set constraints on vocal structure in ways that run parallel across taxa and that help to explain the evolution of vocal phenotypes.
2. Resulting variations in birdsong structure influence communicative functions for both intrasexual and intersexual interactions in ways that favor the evolution of vocal phenotypes that challenge vocal performance capacities.

Note that these are generalized statements of hypotheses H1 and H2 that are outlined in Podos (2017). As pointed out in that review, testing hypotheses about vocal mechanics and song function involve distinct sets of approaches and methods (Fig. 9.1). Recent studies in both areas are offering new and thought-provoking insights into causal relationships between vocal mechanisms and song evolution.

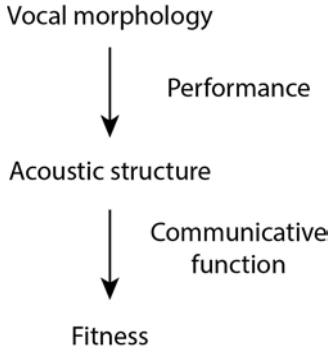


Fig. 9.1 Two-stage process by which vocal morphology influences fitness for animals that employ acoustic signals as mating displays. In stage 1, animals generate sound by activating *vocal morphology* under the control and action of neural, endocrinological, and motor systems. The term *performance* encompasses these processes and provides a framework for analyzing constraints or trade-offs in sound production that limit the range of vocal phenotypes that animals can express. Presumably, individuals in a population will vary in their ability to generate high-performance acoustic phenotypes. Such variations might, in some species, influence stage 2, which is how signal structure impacts communicative function. Sexual selection theory predicts that signal receivers will evolve to respond differentially to varying signal phenotypes and also that higher-performance displayers will enjoy higher fitness outcomes. This is because high-performance displayers are of relatively high quality, or because they can tap more effectively into a receiver's perceptual or aesthetic biases. Fitness benefits for generating high-performance acoustic displays can influence vocal morphology and performance so as to favor displays of increasingly high performance in a feedback loop

9.2 Constraints on Song Production and Performance

9.2.1 Conceptual Background

Across many animal taxa, sexual selection drives mating ornaments and displays to evolve to be increasingly complex or elaborate, and there is growing recognition that this tendency can be counteracted by proximate organismal constraints. In the case of mating ornaments, selection for elaborate phenotypes can be tempered by developmental or metabolic trade-offs that diminish the expression of other traits. Illustrations of such have been identified in certain beetle species in which sexual selection for large horns is accompanied by reductions in the size of other critical morphological systems, including eyes and testes (Emlen 2001; Simmons and Emlen 2006). In the case of mating displays, selection for intricate, dynamic displays can be tempered by limits imposed by morphology, musculoskeletal architecture, and neural capacities (Barske et al. 2011; Manica et al. 2017). Corresponding variations in display traits, in both stamina and skill, potentially influence how signalers fare in social and sexual interactions (e.g., Byers et al. 2010). Evidence for the communicative salience of display features limited by display mechanics has been accumulating rapidly in diverse animal taxa (e.g., Manica et al. 2016; Briffa and Lane 2017). For vocal signals, the idea that performance limits could shape the evolution of signal structure was first articulated over twenty years ago (Lambrechts 1996; Podos 1997), and experimental advances for the first decade have been reviewed in detail elsewhere (Podos et al. 2009; Sakata and Vehrencamp 2012).

Hypotheses about how performance limits influence display evolution are probably best framed in language developed by functional and ecological morphologists who have asked how, in general, animal morphology and mechanics enable and limit the abilities of animals to achieve ecologically relevant behavioral outcomes (Arnold 1983) such as those associated with feeding (Wainwright 1988) or locomotion (Garland and Losos 1994; Irschick et al. 2008). Across taxa and across behavioral contexts, the concept of performance helps to explain how morphology and mechanics are translated into behavioral outcomes and how morphological and mechanical variation both enable and constrain the range of outcomes animals can achieve (Wainwright and Reilly 1994; Irschick et al. 2015).

The case for mechanistic and phylogenetic constraints on display evolution has been well-developed for acoustic signals, not just for birdsong (Clark et al. 2011; Miles et al. 2018). Acoustic signals are powered by sets of muscles that, often in synchrony with other motor systems (e.g., respiration and locomotion), generate sound by activating and sustaining time-varying pressure waves that transmit through the relevant media (air, water, or objects such as twigs or leaves). Accordingly, limits on the speed, power, timing, and coordination of acoustic musculature can constrain signal outcomes (Podos and Patek 2015). Approaches that have been useful in documenting mechanical constraints on sound production include analyses of scaling and allometry (Ophir et al. 2010), power output (Clark 2012), kinematics (Westneat et al. 1993), development (Podos 1996; Zollinger and Suthers 2004), coordination and control (Suthers et al. 2012; Prather et al. 2012), morphology (Riede and Goller 2014; Elemans et al. 2015), and phylogenetics (Cardoso and Hu 2011). Building on these types of studies and on theory, it is also possible to infer performance limitations via analysis of acoustic structure and to test expectations about upper boundaries and trade-offs among acoustic traits (e.g., Podos 1997; Cardoso 2017).

9.2.2 Source-Based Constraints

In the case of songbird vocalizations, it is particularly useful to distinguish constraints operating at the sound source (syrinx) versus at the periphery (respiratory and vocal tract motor systems) (Nowicki et al. 1992). With respect to the former, the syrinx is the primary source of song and operates pneumatically as airflow from the lungs activates vibrations in syringeal tissues (Riede et al. 2019). A number of features make the syrinx highly versatile and, accordingly, enhance the diversity of sounds birds can generate. For example, the two sides of the syrinx can operate in synchrony or independently with each side potentially emphasizing different frequency ranges (e.g. Suthers 1990; Secora et al. 2012). In addition, the syrinx is relatively small, which allows for rapid frequency and amplitude modulations, and syringeal tissue configurations in most avian taxa are mediated by multiple sets of muscles, which enable fine-scale motor control.

A recent study of eight songbird species has identified another attribute of the syrinx that appears to enhance vocal versatility: the structural complexity of syringeal labia (vibrating tissues). Riede and Goller (2014) applied histological approaches to characterize the structure of syringeal labia and showed that species with more complex labia (defined as having greater asymmetry and more layers of extracellular matrix) tend to produce songs with wider ranges of fundamental frequencies and, thus, with greater acoustic complexity. It would be interesting to learn, applying a phylogenetic approach, whether labial structure tends to increase in complexity over evolutionary time under selection to enhance the versatility of syringeal function. In any case, the labial structure of the syrinx seems very well-adapted for enabling birds to produce a wide range of sounds.

9.2.3 Peripheral Constraints

By contrast, the vocal periphery, namely the respiratory and vocal tract motor systems, seems more likely to set constraints on the evolution of song. Unlike the syrinx, these motor systems are relatively large in size and, therefore, harder to modulate at higher speeds (a scaling effect). These systems are also subject to competing selection pressures associated with breathing, feeding, and locomotion (Nowicki et al. 1992; Podos and Hendry 2006). Elegant demonstrations of peripheral constraints on song structure have been provided by a series of laboratory studies by Rod Suthers and colleagues, including a recent study on respiration and singing in canaries, *Serinus canaria* (Suthers et al. 2012).

Prior research showed that female canaries tend to solicit copulations more readily from males who sing songs with higher vocal performance, in particular, males with songs featuring syllables produced at rapid rates with wide frequency bandwidths (Vallet and Kreutzer 1995; Vallet et al. 1998) and extended durations (Pasteau et al. 2009). Suthers et al. (2012) expanded on this finding by documenting patterns of airflow during the production of these and other types of syllables. Low-performance syllables, on the one hand, appeared to be generated predominantly by only one side of the syrinx and with the opportunity for birds to breathe between notes. Such patterns presumably do not tax the motor system to a significant extent. High-performance syllables, on the other hand, involved precise coordination between both sides of the syrinx and also required the use of *pulsatile respiration* in which air sac pressure remains positive throughout the song. These adaptations enable birds to sing higher-performance songs; yet at the same time, they must be limited to some degree by musculoskeletal capacities. Motor constraints should thus limit the expression of timing traits and could also drive trade-offs between timing traits and other acoustic features such as amplitude and phrase duration. One can also extrapolate that evolutionary changes in respiratory motor patterns, by selection for flight adaptations or scaling, will impose secondary impacts on the structure of high performance vocal phrases.

Additional studies have focused on variation in the morphology and function of the vocal tract, including the beak. Because beaks play prominent roles in song production, the core hypothesis applied so far is that variation in beak form and function, which tends to be prominent in many avian groups, can impose differential constraints on song production. One way this can occur is via selection for increased bite force, which should in turn lead to reduced speeds of gape modulation because of trade-offs between force and velocity in cyclical motor systems (Herrel et al. 2009). Accordingly, birds with large beaks should be especially constrained in the production of rapid or wide-bandwidth songs, which require rapid or broad beak gape modulations.

Support for this hypothesis was first reported for Darwin's finches of the Galápagos Islands, both across (Podos 2001; Herrel et al. 2009) and within species (Herrel et al. 2005; Huber and Podos 2006). Those studies revealed that birds with larger beaks, both absolutely and scaled for body size, tend to produce songs with slower trill rates and narrower frequency bandwidths in the direction predicted by the hypothesis of vocal motor constraints. Two subsequent studies (Derryberry et al. 2012, 2018) explored the relationship between beaks and songs within another Neotropical passerine group, the suboscine passerine family Furnariidae (woodcreepers and ovenbirds). The first study analyzed relationships between songs and beak morphology, correcting for phylogenetic non-independence, across 51 species of woodcreepers within the subfamily Dendrocolaptinae (Derryberry et al. 2012). Across the subfamily, species with larger beaks produced songs with slower rates of note repetition and narrower frequency bandwidths. That finding is consistent with both an hypothesis of motor constraints and with the aforementioned results from Darwin's finches (Podos et al. 2009). The second study expanded the scale of analysis to 276 species across the family and, as with the prior analysis, supported the hypotheses that beaks impose constraints on song diversification (Derryberry et al., 2018). It is intriguing to learn that beaks can constrain song in a conserved manner across a species-rich and diverse avian radiation. More generally, selection on beaks is now regarded as an incidental driver of changes in the vocal phenotype, especially in how it modifies the range of vocalizations birds could conceivably produce within the constraints of their motor systems (Podos et al. 2004; Garcia et al. 2018). An important caveat here is that a correlation between beak morphology and song structure would not be expected to be universal. It would be unrealistic to expect that the morphology of one element of the vocal periphery would influence all of the diverse vocal phenotypes birds express (Podos et al. 2004, 2009). Indeed, beak morphology and song structure in some groups appear to map onto each other only loosely or not at all (Slabbekoorn and Smith 2000; Porzio et al. 2018).

9.2.4 Acoustic Parameters Constrained by Performance

Whether grounded in the vocal source or periphery, performance constraints are considered to potentially shape the structure of wide ranges of acoustic parameters across multiple levels of organization (Gil and Gahr 2002; Podos and Nowicki 2004). Some studies on this topic have focused on the consistency or stereotypy with which birds repeat vocal features across renditions. Greater levels of consistency are thought to reflect the quality or ability of the singer (Sakata and Vehrencamp 2012) and in at least one species are detectable by the audience targeted by the singer (Woolley and Doupe 2008). Song features for which consistency has been quantified include spectral structure in tropical mockingbirds (*Mimus gilvus*) (Botero et al. 2009) and banded wrens (*Thryophilus pleurostictus*) (de Kort et al. 2009b; Vehrencamp et al. 2013), song frequency ratios in black-capped chickadees (*Poecile atricapillus*) (Grava et al. 2012), and broad sets of frequency and timing features in prairie warblers (*Setophaga discolor*) (Byers et al. 2015, 2016), zebra finches (*Taeniopygia guttata*) (Kao and Brainard 2006; Kojima et al. 2018), and Bengalese finches (*Lonchura striata* var. *domestica*) (James and Sakata 2014, 2015). In these and related studies, variation within singers is typically characterized using coefficients of variation or pairwise spectrogram cross-correlation calculations with the typical outcome being that acoustic consistency varies greatly across singers.

Other recent analyses of vocal performance have focused on trilled song phrases in which notes or syllables are repeated in rapid succession (Fig. 9.2). In trilled song phrases, faster rates of repetition (*trill rates*) should be harder to achieve because of increased demands on the vocal motor system. Trill rate is often evaluated in tandem with frequency bandwidth (the range of frequencies expressed in a song phrase) following the hypothesis that maximizing one of these parameters should diminish the expression of the other because of time-by-frequency trade-offs in production mechanics (Podos 1997; Podos and Nowicki 2004). Several dozen studies have characterized trill rate by frequency bandwidth trade-offs (Wilson et al. 2014), which have been identified not just in songbirds but also other animal taxa, including two species of seabirds (brown skuas [*Stercoranius antarcticus*], Janicke et al. 2008; great frigatebirds [*Fregata minor*], Juola and Searcy 2011), and a mammal (the Neotropical singing mouse [*Scotinomys teguina*], Pasch et al. 2011). Relatedly, trill rates in introductory segments of calls in Bornean gibbons (*Hylobates muelleri*) trade off with segment durations (Clink et al. 2018). Trilled acoustic signals in grey tree frogs (*Hyla versicolor*) and field crickets (*Gryllus pennsylvanicus*) exhibit trade-offs in two other features: call duration and calling rate (Reichert and Gerhardt 2012; Wagner et al. 2012). Broadly speaking, trade-offs among acoustic parameters in trills might originate in numerous ways, such as through correlational selection, limits on energy allocation, or biomechanical limits (e.g., Wagner et al. 2012).

Performance of vocal sequences, be they trilled or not, can also be quantified as magnitudes or rates of frequency and timing modulation (Stowell and Plumbley 2014; Podos et al. 2016). One such metric, *frequency excursion*, was used to measure

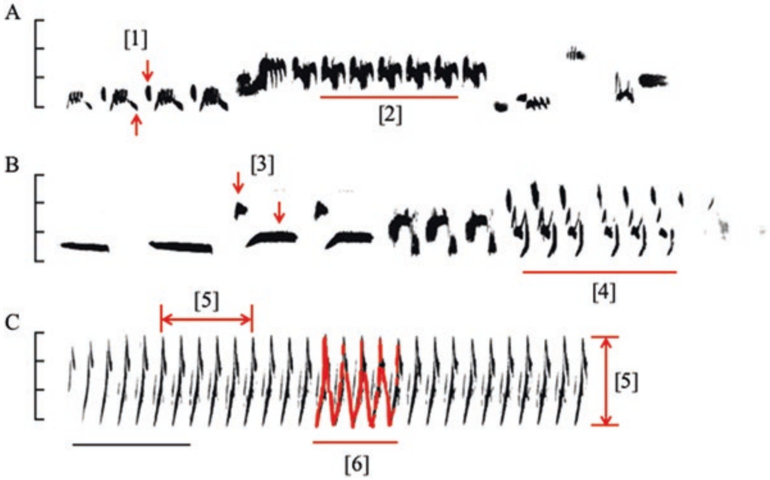


Fig. 9.2 Examples of vocal performance measurements. Spectrograms (*black*) of representative trilled songs from three North American sparrow species: (A) Lincoln's sparrow (*Melospiza lincolnii*), (B) Vesper sparrow (*Pooectes gramineus*), and (C) Chipping sparrow (*Spizella passerina*). Red ink overlays depict six different methods that have been used to infer vocal performance from trilled songs. [1] Singers may face constraints on shifting frequencies from the offset of one note to the onset of the subsequent note (*red arrows*), especially over short durations (Geberzahn and Aubin 2014). [2] It may be difficult for singers to maintain structural consistency (*red bar*) across syllables within trills (Sakata and Vehrencamp 2012). [3] Singers may face production challenges in modulating amplitude quickly between low and high values (*red arrows*) within syllables (Forstmeier et al. 2002, amplitude represented by the darkness of pixels). [4] Illustration of broken syntax (*red bar*) in which the normal trilled rhythm is perturbed; may indicate that the singer has encountered a performance limitation (Podos 1996; Zollinger and Suthers 2004). [5] Birds may encounter mechanical limits in the production of rapid trill rates (*horizontal red arrows*) and wide frequency bandwidths (*vertical red arrows*), the latter best measured not from spectrograms but from amplitude spectra, with the two trading off at upper levels (Podos 1997; Wilson et al. 2014). [6] One way to measure vocal performance of trills is by tracing maximum amplitude contours (*red tracing* above *red bar*) and summing the distance traversed standardized to per unit time (frequency excursion) (Podos et al. 2016). For all rows, the y-axis represents frequency (2–8 kHz) and the x-axis represents time; scale bar shown in (C) is 0.5 sec

vocal performance of songs both within and across individuals in a population of Adelaide's warblers (*Setophaga adelaidae*) (Schraft et al. 2017). One finding of this study was that individual birds ramped up their vocal performance as the morning progressed. This suggests that a bird's vocal performance on a given day benefits from practice, and thus the songs are performance limited. Moreover, this finding implies that birds who can sing at high performance levels at dawn may be high-quality singers, and males who sing more at dawn are better able to showcase their capacities as singers (Schraft et al. 2017).

Similarly, Geberzahn and Aubin (2014) performed a study of nontrilled vocal performance in skylarks (*Alauda arvensis*), a species that sings songs while airborne. Here the authors focused on two acoustic variables: the duration of silent gaps between successive notes, and the magnitude of separation between ending frequencies of an individual note and starting frequencies of the next note in succession (*inter-note frequency shift*). The authors hypothesized that birds who minimized gap duration would be restricted in their magnitudes of inter-note frequency shifts because of limits on the speed with which the vocal apparatus could be reconfigured, as required to accommodate such shifts (also see Podos et al. 2009). Indeed, there was a trade-off of these two parameters at maximal degrees of expression, which was parallel to trade-offs between trill rate and frequency bandwidth (Podos 1997; Wilson et al. 2014). Additionally, different singers achieved different levels of performance in this trade-off, which was quantified using a composite index termed *vocal gap deviation* (Geberzahn and Aubin 2014; also see Cardoso 2014).

9.3 The Relevance of Performance Constraints for Communicative Function

A major goal of the research reviewed in Sect. 9.2 was to characterize the challenges birds face, mechanically speaking, in singing different kinds of songs. The outcome of these challenges is manifested over evolutionary time as variation in vocal acoustic structure. It is critical to note that research on vocal performance by itself cannot reveal if and how performance constraints impact communicative function and, accordingly, fitness. In other words, the term "performance" merely characterizes how morphological and motor systems are translated into action—analogous to how the term "expression" describes, in a neutral way, how genotypes are converted into phenotypes (Fig. 9.1).

Determining whether performance variations matter with regard to fitness requires direct tests of signal function. This has been another area of significant research activity. From the standpoint of sexual selection theory, the prediction might be that animals will evolve to discern and respond differentially to vocal performance variations if and when these variations offer reliable windows into a singer's quality, motivation, or other related attributes (Searcy and Nowicki 2005). Vocal performance variations might generate differential responses through mechanistic biases in listener's perceptual systems, for example, through pre-existing perceptual biases (Ryan 2018). Alternatively, listeners might detect and perceive vocal performance variations equivalently, yet make active choices in how they respond, for example, through weighted assessment decisions or biased aesthetic preferences (Prum 2017; Ryan 2018).

This discussion starts by offering a reminder that there is no reason to predict that all species will attend to performance-based vocal variations (Cramer 2013b). Rather, attention to performance-based variation would be predicted for species that show wide natural ranges of vocal performance, for which some songs push the performance capacities of the singer, and for which performance variations map onto singer attributes. For other species, sexual selection and mate choice will certainly favor the evolution of other types of song traits, such as repertoire complexity or the accuracy of song model copying, which could also provide listeners with indicators of signaler attributes. In any given species, determining which song traits are most salient in mating contexts and what proximate and evolutionary pressures shape those traits is probably best regarded as an empirical rather than a conceptual line of inquiry. Recent empirical tests of the possible functions of variations in vocal performance fall into four main categories: vocal performance modulations, territorial responses to vocal performance variations, sexual responses to vocal performance variations, and fitness correlates of vocal performance variation.

9.3.1 Category 1: Vocal Performance Modulations

The first category concerns male-male territorial interactions and, in that context, asks if and how males might adjust their own vocal activity when interacting with territorial rivals. Increases in vocal performance under such circumstances might be read as an indicator of elevated aggressive intent. In swamp sparrows (*Melospiza georgiana*) territorial males challenged with playbacks of conspecific songs (simulating a territorial intrusion) responded with faster trill rates and broader frequency bandwidths as compared to songs they produced during playback of heterospecific song stimuli (DuBois et al. 2009). Similarly, chiffchaffs (*Phylloscopus collybita*) and tree pipits (*Anthus trivialis*) presented with conspecific playback increased vocal performance levels, which were measured as syllable repetition rates (Linhart et al. 2013; Petruskova et al. 2014). In a study of serins (*Serinus serinus*), a species for which baseline levels of vocal performance are comparatively high (Cardoso et al. 2007), Funghi et al. (2015) found that some serins challenged with conspecific playback were able to further increase their performance, even if only slightly, for two vocal parameters: syllable rate and song duration. In a study of skylarks mentioned previously (Sect. 9.2), which is another species with high baseline levels of vocal performance, Geberzahn and Aubin (2014) showed that birds confronted with playback tended to adjust their songs toward smaller vocal gap deviations, representing higher performance output. Finally, a study of banded wrens revealed that, in response to playback, the levels of aggression in male wrens correlated with the magnitude of increases in trill rates above baseline (dawn chorus) levels (Vehrencamp et al. 2013).

It should be mentioned that these types of effects are not limited to birds. For example, Reichert and Gerhardt (2012) showed that grey tree frogs that were spaced close to one another, thus experiencing an elevated competitive threat, called with enhanced performance levels in a range of measures, including call duration, call duty cycle, and call deviation, which is a composite measure of performance. The degree to which territorial challenges affect vocal performance in other species should be explored in further detail.

9.3.2 Category 2: Territorial Responses to Vocal Performance Variations

The second category of recent studies on vocal performance also centers on male-male territorial interactions but instead asks how responses of territorial males covary with playbacks of different performance-related song variants. Following up on their previous study (Dubois et al. 2009), Dubois et al. (2011) conducted two experiments in which they challenged territorial male swamp sparrows with song stimulus pairs that were matched to song type (songs with the same note sequences) yet varied in two performance features: trill rate and frequency bandwidth. The magnitude of differences in these features within pairs of test stimuli corresponded either to natural variation in performance between individuals (Experiment 1, relatively high variation) or to intra-individual variation across aggressive contexts (Experiment 2, relatively low variation). The research team observed that receivers readily discriminated performance variations at magnitudes that naturally occur between but not within singers.

Also working with swamp sparrows, Moseley et al. (2013) showed that males can discriminate songs based on artificially introduced variations in trill rate alone. In this study, birds generally gave weak responses to low-performance songs (with slower trill rates), yet they fled in response to high-performance songs (with faster trill rates). These results paralleled the outcomes of playbacks to territorial banded wrens in which playback stimuli were constructed to vary not in trill rates (as in Moseley et al. 2013) but rather in frequency bandwidth (Illes et al. 2006; de Kort et al. 2009a). Chipping sparrows (*Spizella passerina*) also discriminate between low-trill rate and high-trill rate versions of natural songs, but in contrast to swamp sparrows and banded wrens, territory holders responded more aggressively and attacked a taxidermic mount positioned near the broadcast speaker more often in response to high performance stimuli (Goodwin and Podos 2014). Recent playback studies by Phillips and Derryberry (2017a, b) suggest that white-crowned sparrows (*Zonotrichia leucophrys*) attend to variations in both trill rate and frequency bandwidth, responding more strongly to higher performance versions of both. Moreover, birds appear to treat variations in both parameters similarly, thus lending support to the functional relevance of a composite index that reflects both parameters: *vocal deviation*. Finally, one species in which vocal performance (at least as defined using traditional metrics) does not appear to influence male territorial responses is the house wren (*Troglodytes aedon*); males of this species respond equivalently to playbacks of low-performance or high-performance song stimuli (Cramer 2013b).

An unexpected finding in several of the studies reviewed here is that the responses of territorial males to playback covaried with their own vocal attributes. For example, male chiffchaffs that produced songs with short intersyllable gaps (high performance songs) were more likely to attack a simulated territorial intruder (Linhart et al. 2013). Similarly, male swamp sparrows that sang at higher performance levels tended to approach the playback speaker more closely than did males that produced songs at lower performance levels (Moseley et al. 2013). In serins, increases in vocal performance in response to playback were observed in birds who approached the playback speaker but not in birds who did not approach the speaker (Funghi et al. 2015). Finally, in chipping sparrows, males challenged by playback sometimes received temporary assistance from a neighbor, and the formation of territorial alliances could be predicted by the relative vocal performance levels of the territory holder, the ally, and the simulated territorial intruder (Goodwin and Podos 2014). All of these results provide additional support for the hypothesis that a bird's vocal performance serves as a reliable indicator of his quality or aggressive intent (Searcy and Beecher 2009).

9.3.3 Category 3: Sexual Responses to Vocal Performance Variations

A third category of recent studies focuses on how variations in vocal performance are perceived and acted upon by females. Initial work measured preferences using the copulation solicitation assay. Female canaries and swamp sparrows perform more solicitation displays to playbacks of high performance song stimuli (Vallet and Kreutzer 1995; Ballentine et al. 2004). More recent studies have emphasized phonotaxis assays in which captive females are given an opportunity to choose (via approach) between pairs of stimuli presented on either side of a testing arena. Female zebra finches prefer to approach speakers playing songs that males had sung in the presence of females (*directed song*) versus in the absence of conspecifics (*undirected song*) (Woolley and Doupe 2008). Because the directed songs are considered higher performance by virtue of their comparatively high syllable-to-syllable stereotypy, these data support the notion that females prefer higher performance songs. Similarly, Lincoln's sparrows (*Melospiza lincolnii*) associate more closely with speakers playing back songs that have been manipulated to faster trill rates (Caro et al. 2010; Weiss et al. 2012). Also in this species, recent experiences of hearing songs can influence females' preferences: prior exposure to low performance songs amplified later preferences for high performance songs (Caro et al. 2010; also see Lyons et al. 2014). Finally, female Lincoln's sparrows and nightingales (*Luscinia megarhynchos*) were more active (another related metric of female preference) when hearing high-performance songs (Caro et al. 2010, Weiss et al. 2012).

Once again, methods and findings from songbirds are being extrapolated to other taxa. For example, female neotropical singing mice, captured and housed temporarily at a field site in Costa Rica, spent more time near a speaker playing back songs with faster trill rates than near a speaker playing songs at their natural trill rates (Pasch et al. 2011).

9.3.4 Category 4: Fitness Correlates of Vocal Performance Variation

The fourth category of studies explores possible correlations between vocal performance and measures or proxies of reproductive success, mainly under field conditions. The basic prediction here, as discussed in Sect. 9.1, is that singers that achieve higher levels of vocal performance will enjoy enhanced fitness benefits (Podos et al. 2009; Sakata and Vehrencamp 2012). In support of this idea, Janicke et al. (2008) found that male brown skuas that sang with higher performance (measured as vocal deviation and peak performance) enjoyed higher reproductive success, which was measured as the number of chicks raised annually over five consecutive breeding seasons; however, the birds did not vary with regard to a body condition index. Juola and Searcy (2011) found the inverse trend in great frigatebirds in which vocal performance mapped positively onto body condition but not pairing success.

Other trends have been reported in other species. For example, male tropical mockingbirds that sang with higher syllable consistency enjoyed higher dominance status and reproductive success (Botero et al. 2009); male black-capped chickadees that produced songs with greater consistency occupied preferred forest types (Grava et al. 2012); and male prairie warblers that produced songs with higher vocal performance measures sired eggs laid at earlier dates (Byers et al. 2015, 2016). In addition, vocal performance levels predicted age in nightingales (Sprau et al. 2013), banded wrens (de Kort et al. 2009b; Vehrencamp et al. 2013), and Bengalese finches (James and Sakata 2014). Vocal performance measurements predicted body condition in Java sparrows (*Lonchura oryzivora*) (Kagawa and Soma 2013) and singing mice (Pasch et al. 2011) and also predicted which tree pipit males were able to defend and maintain territories (Petruskova et al. 2014). By contrast, no fitness correlates with vocal performance were identified for house wrens, a result that corresponds to the finding for this species that territorial males do not attend to variations in vocal performance (Cramer 2013a, b).

While these studies provide general support for the functional importance of vocal performance variations, it is important to note that correlations by themselves cannot demonstrate causation. Ideally, descriptive studies along these lines would be accompanied by functional tests of the type reviewed earlier in this section.

9.4 Future Directions

Three topics for future research on vocal performance could prove especially promising. All of these topics are the subjects of current research, but much remains to be learned.

9.4.1 Neural and Endocrine Mechanisms

First, it will be helpful to better document how neural and endocrinological factors help to shape song performance and perception. An example of possible neural constraints on song performance was presented by Prather et al. (2012), who quantified the neural bases of rapid trill processing in swamp sparrows. These authors recorded the activity of neurons in the sensorimotor nucleus HVC in response to playbacks of songs with rapid trills and found that neural activity became more temporally variable as trill rates increased. Patterns of neural firing failed to keep time with syllables when repeated at the fastest trill rates. Moreover, the observed neural firing pattern closely matched the timing of broken syntax in songs of birds reared with rapid trill models (Podos 1996). Prather et al. (2012) thus proposed that observed temporal disruptions in auditory processing might also manifest as disruptions in the neural control of rapid song output, given their shared mechanistic and anatomical bases, and that neural limits could interact with motor limits to constrain vocal phenotypes. In cases where neural capacity limits the form of mating displays, it becomes reasonable to consider the evolutionary effects of variations in cellular processes that might limit neural function. A review by Koch and Hill (2018) aims to draw such connections by providing evidence from the human medical literature that both neural and motor capacities are impacted directly by mitochondrial function. As such, displays could provide reliable and direct indicators of signaler condition and quality at the cellular level (Koch and Hill 2018).

Lyons and Sockman (2017), in a study of Lincoln's sparrows, took another approach by monitoring birds' neural responses to playback of songs with low-performance or high-performance levels. Playback of high-performance songs but not low-performance songs led to relatively diminished monoaminergic activity across multiple brain regions in female Lincoln's sparrows. At minimum, this result indicates that females can respond differentially to low-performance versus high-performance songs, as predicted given behavioral data for this species (Caro et al. 2010). Ultimately, analyses of changes in neurochemical activity might point the way to mechanisms that mediate differential preferences (Lyons and Sockman 2017; Barr and Woolley 2018).

On the endocrine side, new approaches are starting to test causal relationships between vocal performance and hormonal status. A first study to draw such connections was conducted with neotropical singing mice, a species that shows trade-offs in vocal features and thus appears to push its vocal performance limits (Pasch et al. 2011). A major question this research team posed was whether variations in androgen activity would be reflected as variation in vocal performance. Castration led to reductions in vocal performance; whereas, castration followed by androgen administration (testosterone or dihydrotestosterone) recovered baseline levels of vocal performance (Pasch et al. 2011). Given these data, it was proposed that androgens help mice to maintain elevated vocal performance through impacts on both neural and morphological components of vocal output.

Alward et al. (2016), studying canaries, used another approach to manipulate androgen function. Canaries were implanted with bicalutamide, which is an androgen receptor antagonist that does not cross the blood-brain barrier and whose effects are thus limited to peripheral motor systems. These birds showed reduced performance in a number of song features, and also showed reduced syringeal mass. Both of these results support the hypothesis that androgens play a causal role in vocal performance through their effects on peripheral motor systems. Alward et al. (2016) recommended that future studies broaden the suite of peripheral systems examined to include not just the syrinx but also respiratory and vocal tract motor systems. The two papers just reviewed (on singing mice and canaries) are part of a growing trend of studies documenting relationships between androgen activity and acoustic signal structure in general (Apfelbeck et al. 2012; Fuxjager et al. 2014).

9.4.2 Cognitive and Sampling Processes

A second major objective for future research on vocal performance should be to better document the cognitive and sampling processes that guide how birds assess and evaluate performance-based song features. A bird's assessment of many vocal traits, such as repertoire size, vocal output rates, song-to-song consistency, and copying accuracy, require integration of immediate perceptual events and memories of songs heard previously. If a bird were tasked with evaluating a singer's repertoire size, for instance, it would need to maintain some sort of running tally (or memory) of the diversity of song types the singer had produced before and be able to recognize whether a song being heard was new or similar to a type heard and memorized previously. The mechanisms that mediate this kind of matching process are not easy to envision, especially when memory spans required are long term.

By contrast, a bird's assessment of vocal performance could presumably be made using simpler and quicker cognitive or sampling mechanisms. Logue and Forstmeier (2008) proposed that birds might readily assess vocal performance in direct side-by-side comparisons of multiple singers, particularly when those singers produce the same song type (song type matching). The assessment would thus be conducted on a relative rather than an absolute basis. Moreover, the time delay and corresponding need for encoding and memorizing prior songs would thus be diminished, analogous to an optometrist offering rapid-fire A-B-A-B options for lenses. An interesting corollary of this hypothesis is that if song type matching does indeed facilitate side-by-side comparisons of vocal performance, then sexual selection might favor higher performance singers that match their neighbors (and thus advertise their skill) yet favor lower performance singers who avoid matching (and thus obscure their poor output). This corollary adds an interesting twist to the literature regarding the role of song type matching as a reliable indicator of a singer's aggressive intent (Searcy and Beecher 2009).

Cardoso (2013) offered a related perspective. He posited that display features that evolve toward maximal performance levels, including many song features reviewed in this chapter, are favored by sexual selection because they are more prone to structural mistakes. Moreover, as Cardoso (2013, p. 1036) states, "[m]istakes should be easily perceptible when they are large deviations to signal structure, such as atypical motor patterns or interrupted displays." It is possible that a perceptual mechanism for recognizing display mistakes would not require comparisons to specific prior displays but rather comparison to a generic construct of what successful displays entail. A follow-up study by Ferreira et al. (2016) offers a particularly thorough survey of potential mistakes in the songs of dark-eyed juncos. These authors detected mistakes in about 6% of songs from a large song sample. Mistakes were distributed disproportionately and largely as predicted across males of varying habitat, age, and reproductive success (Ferreira et al. 2016).

The cognitive and perceptual bases of vocal performance assessment also have been addressed by Lyons et al. (2014) in a laboratory study of song assessment by female Lincoln's sparrows. During a week-long pre-test time block, birds were played either high-performance or low-performance songs exclusively, and approaches to the playback speaker were quantified. At first, birds presented with high-performance songs made closer approaches to the speaker than did birds presented with low-performance songs. However, this difference dissipated after a week of daily trials, presumably as a result of habituation. On the eighth day, birds from both treatments were presented with identical sets of intermediate-performance songs. Birds that had previously experienced only low-performance songs approached the speaker more closely than did birds that had previously heard only high-performance songs. The authors concluded that birds perceive and act not just on absolute vocal performance levels, but also upon contrasts to the vocal performance levels of songs they had heard previously.

9.4.3 Evolution of Song Performance

A third major potential direction for future research will be to consider data on vocal performance within the broader context of other factors that shape birdsong evolution. One direction of growing interest in this regard concerns the evolution of functional and morphological systems that contribute to song production and performance. The evolution of the syrinx (as the sound source) is probably shaped mainly by selection for effective song production. This is because the major functional benefit of the syrinx, as far as we know, is for song production. By contrast, other motor components that shape song production, notably the respiratory and vocal tract motor systems, are subject to other substantial selection pressures (e.g., for breathing or feeding capacity). As such, selection for these functions could potentially impose indirect effects on song evolution. It follows that the evolutionary divergence of morphology for noncommunicative functions might facilitate the divergence of vocal structure and thus potentially foster reproductive isolation and speciation (Podos et al. 2004; Podos and Hendry 2006).

Songs may also experience multiple selective pressures that are directly related to communicative function (Wilkins et al. 2013). Such factors include selection for optimal signal transmission through the habitat, cultural evolution, and selection for vocal complexity. A study of song structure and responses in a Galápagos finch species, the medium ground finch (*Geospiza fortis*), attempted to compare and rank the relative impacts of culturally evolved divergence versus beak-related divergence on song function (Podos et al. 2013). Across two sites on one Galápagos island (Santa Cruz), song divergence appeared to be driven to a greater extent by beak-related evolution than by cultural divergence. Correspondingly, responses of birds to song variations were more pronounced for stimuli sets that diverged in beak-related factors compared to culturally driven factors (Podos et al. 2013). In their analysis of song diversity across species of furnarids, Derryberry et al. (2018) made the case that songs in these birds have been impacted both indirectly by the evolution of morphology (beak and body size) and directly by selection for optimal transmission. As found by Podos et al. (2013), the indirect influence of morphological factors seems to play a disproportionately large role in song divergence. It would be interesting to determine whether patterns of perception and responses in furnarid species are ranked as they are for the Galápagos medium ground finch, with greater salience for morphologically linked song features.

Cardoso and Hu (2011) also adopt a comparative approach and integrate information about multiple factors in birdsong evolution. These authors analyzed the songs of forty-eight species of wood warblers, focusing on variations in vocal performance and structural complexity. Their main result was that the two types of parameters appear to trade off with each other. In particular, species that achieved high performance levels tended to sing relatively simple songs, and species singing complex songs scored low on performance. This finding leads to several inferences: (1) that sexual selection can impact song evolution in different ways in different species groups; (2) that songs favoring either complexity or performance evolve readily, given their multiple origins even within a single group; and (3) that selection for either performance or complexity might preclude the other through its effects on motor mechanisms, perceptual mechanisms, or some combination thereof.

9.5 Conclusion

Research on vocal performance has been advancing quickly and is providing many new insights into the mechanisms and evolution of bird vocalizations. Multiple aspects of vocal mechanisms, including musculoskeletal, neural, and endocrinological attributes, together set limits on the ranges of acoustic phenotypes birds can generate (i.e., vocal performance). The resulting variation in vocal performance outcomes are being shown to impact song function across diverse species and in diverse ways. As such, population-level variations in reproductive success can favor the evolution of high-performance vocal phenotypes through selection on mechanistic attributes that enable their production. Analyses of vocal performance variations, in both cause and consequence, can help enrich our understanding of the evolution of bird songs and vocal communication.

Acknowledgements We are grateful to Sarah Woolley, Jon Sakata, and Arthur Popper for their helpful editorial suggestions.

Compliance with Ethics Requirements Jeffrey Podos declares that he has no conflict of interest.

Ha-Cheol Sung declares that he has no conflict of interest.

References

Alward BA, Madison FN, Gravley WT, Ball GF (2016) Antagonism of syringeal androgen receptors reduces the quality of female-preferred male song in canaries. Anim Behav 119:201–212. https://doi.org/10.1016/j.anbehav.2016.07.010

Apfelbeck B, Kiefer S, Mortega KG, Goymann W, Kipper S (2012) Testosterone affects song modulation during simulated territorial intrusions in male black redstarts (*Phoenicurus ochruros*). PLoS One 7(12):e52009. https://doi.org/10.1371/journal.pone.0052009

Arnold SJ (1983) Morphology, performance and fitness. Amer Zool 23(2):347–361

Ballentine B, Hyman J, Nowicki S (2004) Vocal performance influences female response to male bird song: an experimental test. Behav Ecol 15(1):163–168. https://doi.org/10.1093/beheco/arg090

Barr HJ, Woolley SC (2018) Developmental auditory exposure shapes responses of catecholaminergic neurons to socially-modulated song. Sci Reports 8:11717. https://doi.org/10.1038/s41598-018-30039-y

Barske J, Schlinger BA, Wikelski M, Fusani L (2011) Female choice for male motor skills. Proc Roy Soc B 278(1724):3523–3528. https://doi.org/10.1098/rspb.2011.0382

Botero CA, Rossman RJ, Caro LM, Stenzler LM, Lovette IJ, de Kort SR, Vehrencamp SL (2009) Syllable type consistency is related to age, social status and reproductive success in the tropical mockingbird. Anim Behav 77(3):701–706. https://doi.org/10.1016/j.anbehav.2008.11.020

Brainard MS, Doupe AJ (2002) What songbirds teach us about learning. Nature 417(6886):351–358

Briffa M, Lane SM (2017) The role of skill in animal contests: a neglected component of fighting ability. Proc Roy Soc B 284(1863):20171596. https://doi.org/10.1098/rspb.2017.1596

Brumm H, Naguib M (2009) Environmental acoustics and the evolution of bird song. Ad Stud Behav 40:1–33. https://doi.org/10.1016/s0065-3454(09)40001-9

Byers J, Hebets E, Podos J (2010) Female mate choice based upon male motor performance. Anim Behav 79(4):771–778. https://doi.org/10.1016/j.anbehav.2010.01.009

Byers BE, Akresh ME, King DI (2015) A proxy of social mate choice in prairie warblers is correlated with consistent, rapid, low-pitched singing. Behav Ecol Sociobiol 69(8):1275–1286. https://doi.org/10.1007/s00265-015-1940-y

Byers BE, Akresh ME, King DI (2016) Song and male quality in prairie warblers. Ethology 122(8):660–670. https://doi.org/10.1111/eth.12513

Cardoso GC (2013) Sexual signals as advertisers of resistance to mistakes. Ethology 119(12):1035–1043. https://doi.org/10.1111/eth.12165

Cardoso GC (2014) Studying the silent side of birdsong. BMC Biol 12(3):62. https://doi.org/10.1186/s12915-014-0062-8

Cardoso GC (2017) Advancing the inference of performance in birdsong. Anim Behav 125:E29–E32. https://doi.org/10.1016/j.anbehav.2016.11.034

Cardoso GC, Hu Y (2011) Birdsong performance and the evolution of simple (rather than elaborate) sexual signals. Am Nat 178(5):679–686. https://doi.org/10.1086/662160

Cardoso GC, Mota PG, Depraz V (2007) Female and male serins (*Serinus serinus)* respond differently to derived song traits. Behav Ecol Sociobiol 61(9):1425–1436. https://doi.org/10.1007/s00265-007-0375-5

Caro SP, Sewall KB, Salvante KG, Sockman KW (2010) Female Lincoln's sparrows modulate their behavior in response to variation in male song quality. Behav Ecol 21(3):562–569. https://doi.org/10.1093/beheco/arq022

Catchpole CK, Slater PJB (2008) Bird song: biological themes and variations, 2nd edn. Cambridge University Press, Cambridge, UK

Clark CJ (2012) The role of power versus energy in courtship: what is the 'energetic cost' of a courtship display? Anim Behav 84(1):269–277. https://doi.org/10.1016/j.anbehav.2012.04.012

Clark CJ, Elias DO, Prum RO (2011) Aeroelastic flutter produces hummingbird feather songs. Science 333(6048):1430–1433. https://doi.org/10.1126/science.1205222

Clink DJ, Grote MN, Crofoot MC, Marshall AJ (2018) Understanding sources of variance and correlation among features of Bornean gibbon (*Hylobates muelleri*) female calls. J Acoustic Soc Amer 144:698–708. https://doi.org/10.1121/1.5049578

Cramer ERA (2013a) Physically challenging song traits, male quality, and reproductive success in house wrens. PLoS One 8(3):e59208. https://doi.org/10.1371/journal.pone.0059208

Cramer ERA (2013b) Vocal deviation and trill consistency do not affect male response to playback in house wrens. Behav Ecol 24(2):412–420. https://doi.org/10.1093/beheco/ars178

de Kort SR, Eldermire ERB, Cramer ERA, Vehrencamp SL (2009a) The deterrent effect of bird song in territory defense. Behav Ecol 20(1):200–206. https://doi.org/10.1093/beheco/arn135

de Kort SR, Eldermire ERB, Valderrama S, Botero CA, Vehrencamp SL (2009b) Trill consistency is an age-related assessment signal in banded wrens. Proc Roy Soc B 276(1665):2315–2321. https://doi.org/10.1098/rspb.2009.0127

Derryberry EP, Seddon N, Claramunt S, Tobias JA, Baker A, Aleixo A, Brumfield RT (2012) Correlated evolution of beak morphology and song in the Neotropical woodcreeper radiation. Evolution 66(9):2784–2797. https://doi.org/10.1111/j.1558-5646.2012.01642.x

Derryberry EP, Seddon N, Derryberry GE, Claramunt S, Seeholzer GF, Brumfield RT, Tobias JA (2018) Ecological drivers of song evolution in birds: disentangling the effects of habitat and morphology. Ecol Evol 8(3):1890–1905. https://doi.org/10.1002/ece3.3760

DuBois AL, Nowicki S, Searcy WA (2009) Swamp sparrows modulate vocal performance in an aggressive context. Biol Lett 5(2):163–165. https://doi.org/10.1098/rsbl.2008.0626

DuBois AL, Nowicki S, Searcy WA (2011) Discrimination of vocal performance by male swamp sparrows. Behav Ecol Sociobiol 65(4):717–726. https://doi.org/10.1007/s00265-010-1073-2

Elemans CPH, Rasmussen JH, Herbst CT, During DN, Zollinger SA, Brumm H, Srivastava K, Svane N, Ding M, Larsen ON, Sober SJ, Svec JG (2015) Universal mechanisms of sound production and control in birds and mammals. Nat Commun 6:8978. https://doi.org/10.1038/ncomms9978

Emlen DJ (2001) Costs and the diversification of exaggerated animal structures. Science 291(5508):1534–1536. https://doi.org/10.1126/science.1056607

Ferreira AC, Atwell JW, Whittaker DJ, Ketterson ED, Cardoso GC (2016) Communication value of mistakes in dark-eyed junco song. Am Nat 188(3):289–305. https://doi.org/10.1086/687520

Fitch WT, Suthers RA, Fey RR, Popper AN (2016) Vertebrate sound production and acoustic communication. Springer handbook of auditory research, vol 53. Springer Press, London

Forstmeier W, Kempenaers B, Meyer A (2002) A novel song parameter correlates with extra-pair paternity and reflects male longevity. Proc Roy Soc Lond B 269:1479–1485. https://doi.org/10.1098/rspb.2002.2039

Funghi C, Cardoso GC, Mota PG (2015) Increased syllable rate during aggressive singing in a bird with complex and fast song. J Avian Biol 46(3):283–288. https://doi.org/10.1111/jav.00480

Fuxjager MJ, Heston JB, Schlinger BA (2014) Peripheral androgen action helps modulate vocal production in a suboscine passerine. Auk 131(3):327–334. https://doi.org/10.1642/auk-13-252.1

Garcia NC, Tubaro PL (2018) Dissecting the roles of body size and beak morphology in song evolution in the "blue" cardinalids (Passeriformes: Cardinalidae). Auk 135(2):262–275. https://doi.org/10.1642/AUK-17-146.1

Garland TJ, Losos JB (1994) Ecological morphology of locomotor performance in squamate reptiles. In: Wainwright PC, Reilly SM (eds) Ecological morphology: integrative organismal biology. University of Chicago Press, Chicago, pp 240–302

Geberzahn N, Aubin T (2014) Assessing vocal performance in complex birdsong: a novel approach. BMC Biol 12:58. https://doi.org/10.1186/s12915-014-0058-4

Gil D, Gahr M (2002) The honesty of bird song: multiple constraints for multiple traits. Trends Ecol Evol 17(3):133–141. https://doi.org/10.1016/s0169-5347(02)02410-2

Goodwin SE, Podos J (2014) Team of rivals: alliance formation in territorial songbirds is predicted by vocal signal structure. Biol Lett 10(2):20131083. https://doi.org/10.1098/rsbl.2013.1083

Grava T, Grava A, Otter KA (2012) Vocal performance varies with habitat quality in black-capped chickadees (*Poecile atricapillus*). Behaviour 149(1):35–50. https://doi.org/10.1163/156853912x625854

Hasselquist D, Bensch S, von Schantz T (1996) Correlation between male song repertoire, extra-pair paternity and offspring survival in the great reed warbler. Nature 381(6579):229–232. https://doi.org/10.1038/381229a0

Herrel A, Podos J, Huber SK, Hendry AP (2005) Bite performance and morphology in a population of Darwin's finches: implications for the evolution of beak shape. Funct Ecol 19(1):43–48. https://doi.org/10.1111/j.0269-8463.2005.00923.x

Herrel A, Podos J, Vanhooydonck B, Hendry AP (2009) Force-velocity trade-off in Darwin's finch jaw function: a biomechanical basis for ecological speciation? Funct Ecol 23(1):119–125. https://doi.org/10.1111/j.1365-2435.2008.01494.x

Huber SK, Podos J (2006) Beak morphology and song features covary in a population of Darwin's finches (*Geospiza fortis*). Biol J Linn Soc 88(3):489–498. https://doi.org/10.1111/j.1095-8312.2006.00638.x

Illes AE, Hall ML, Vehrencamp SL (2006) Vocal performance influences male receiver response in the banded wren. Proc Roy Soc B 273(1596):1907–1912. https://doi.org/10.1098/rspb.2006.3535

Irschick DJ, Meyers JJ, Husak JF, Le Galliard J-F (2008) How does selection operate on whole-organism functional performance capacities? A review and synthesis. Evol Ecol Res 10(2):177–196

Irschick DJ, Briffa M, Podos J (2015) Animal Signaling and function: an integrative approach. John Wiley and Sons Inc., Hoboken

James LS, Sakata JT (2014) Vocal motor changes beyond the sensitive period for song plasticity. J Neurophysiol 112(9):2040–2052. https://doi.org/10.1152/jn.00217.2014

James LS, Sakata JT (2015) Predicting plasticity: acute context-dependent changes to vocal performance predict long-term age-dependent changes. J Neurophysiol 114(4):2328–2339. https://doi.org/10.1152/jn.00688.2015

Janicke T, Hahn S, Ritz MS, Peter HU (2008) Vocal performance reflects individual quality in a nonpasserine. Anim Behav 75:91–98. https://doi.org/10.1016/j.anbehav.2007.04.007

Juola FA, Searcy WA (2011) Vocalizations reveal body condition and are associated with visual display traits in great frigatebirds (*Fregata minor*). Behav Ecol Sociobiol 65(12):2297–2303. https://doi.org/10.1007/s00265-011-1240-0

Kagawa H, Soma M (2013) Song performance and elaboration as potential indicators of male quality in Java sparrows. Behav Proc 99:138–144. https://doi.org/10.1016/j.beproc.2013.07.012

Kao MH, Brainard MS (2006) Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. J Neurophysiol 96(3):1441–1455. https://doi.org/10.1152/jn.01138.2005

Koch RE, Hill GE (2018) Behavioural mating displays depend on mitochondrial function: a potential mechanism for linking behaviour to individual condition. Biol Rev 93:1387–1398. https://doi.org/10.1111/brv.12400

Kojima S, Kao MH, Doupe AJ, Brainard MS (2018) The avian basal ganglia are a source of rapid behavioral variation that enables vocal motor exploration. J Neurosci 38(45):9635–9647. https://doi.org/10.1523/jneurosci.2915-17.2018

Lambrechts MM (1996) Organization of birdsong and constraints on performance. In: Kroodsma D, Miller E (eds) Ecology and evolution of acoustic communication in birds. Comstock Publishing Associates, Ithaca, pp 305–320

Linhart P, Jaska P, Petruskova T, Petrusek A, Fuchs R (2013) Being angry, singing fast? Signalling of aggressive motivation by syllable rate in a songbird with slow song. Behav Proc 100:139–145. https://doi.org/10.1016/j.beproc.2013.06.012

Logue DM, Forstmeier W (2008) Constrained performance in a communication network: implications for the function of song-type matching and for the evolution of multiple ornaments. Am Nat 172(1):34–41. https://doi.org/10.1086/587849

Luther D (2009) The influence of the acoustic community on songs of birds in a neotropical rain forest. Behav Ecol 20(4):864–871. https://doi.org/10.1093/beheco/arp074

Lyons SM, Sockman KW (2017) Sex differences in forebrain monoaminergic response to song performance. Brain Behav Evol 89(3):219–230. https://doi.org/10.1159/000471784

Lyons SM, Beaulieu M, Sockman KW (2014) Contrast influences female attraction to performance-based sexual signals in a songbird. Biol Lett 10(10):20140588. https://doi.org/10.1098/rsbl.2014.0588

Manica LT, Graves JA, Podos J, Macedo RH (2016) Multimodal flight display of a neotropical songbird predicts social pairing but not extrapair mating success. Behav Ecol Sociobiol 70(12):2039–2052. https://doi.org/10.1007/s00265-016-2208-x

Manica LT, Macedo RH, Graves JA, Podos J (2017) Vigor and skill in the acrobatic mating displays of a Neotropical songbird. Behav Ecol 28(1):164–173. https://doi.org/10.1093/beheco/arw143

Marler P, Slabbekoorn H (2004) Nature's music: the science of birdsong. Elsevier Academic Press, San Diego

Miles MC, Schuppe ER, Ligon RM IV, Fuxjager MJ (2018) Macroevolutionary patterning of woodpecker drums reveals how sexual selection elaborates signals under constraint. Proc Roy Soc B doi 285:20172628. https://doi.org/10.1098/rspb.2017.2628

Moseley DL, Lahti DC, Podos J (2013) Responses to song playback vary with the vocal performance of both signal senders and receivers. Proc Roy Soc B 280(1768):20131401. https://doi.org/10.1098/rspb.2013.1401

Nelson DA, Marler P (1990) The perception of birdsong and an ecological concept of signal space. In: Stebbins WC, Berkley MA (eds) Comparative perception, complex signals, vol 2. Wiley, New York, pp 443–478

Nowicki S, Westneat MW, Hoese W (1992) Birdsong: motor function and the evolution of communication. Semin Neurosci 4:385–390

Ophir AG, Schrader SB, Gillooly JF (2010) Energetic cost of calling: general constraints and species-specific differences. J Evol Biol 23(7):1564–1569. https://doi.org/10.1111/j.1420-9101.2010.02005.x

Pasch B, George AS, Campbell P, Phelps SM (2011) Androgen-dependent male vocal performance influences female preference in Neotropical singing mice. Anim Behav 82(2):177–183. https://doi.org/10.1016/j.anbehav.2011.04.018

Pasteau M, Nagle L, Monbureau M, Kreutzer M (2009) Aviary experience has no effect on predisposition of female common canaries (*Serinus canaria*) for longer sexy phrases. Auk 126(2):383–388. https://doi.org/10.1525/auk.2009.08101

Peters S, Derryberry EP, Nowicki S (2012) Songbirds learn songs least degraded by environmental transmission. Biol Lett 8(5):736–739. https://doi.org/10.1098/rsbl.2012.0446

Petruskova T, Kinstova A, Pisvejcova I, Laguna JM, Cortezon A, Brinke T, Petrusek A (2014) Variation in trill characteristics in tree pipit songs: different trills for different use? Ethology 120(6):586–597. https://doi.org/10.1111/eth.12231

Phillips JN, Derryberry EP (2017a) Equivalent effects of bandwidth and trill rate: support for a performance constraint as a competitive signal. Anim Behav 132:209–215. https://doi.org/10.1016/j.anbehav.2017.08.012

Phillips JN, Derryberry EP (2017b) Vocal performance is a salient signal for male-male competition in white-crowned sparrows. Auk 134(3):564–574. https://doi.org/10.1642/auk-17-2.1

Podos J (1996) Motor constraints on vocal development in a songbird. Anim Behav 51:1061–1070. https://doi.org/10.1006/anbe.1996.0107

Podos J (1997) A performance constraint on the evolution of trilled vocalizations in a songbird family (Passeriformes: Emberizidae). Evolution 51:537–551

Podos J (2001) Correlated evolution of morphology and vocal signal structure in Darwin's finches. Nature 409:185–188. https://doi.org/10.1038/35051570

Podos J (2017) Birdsong performance studies: reports of their death have been greatly exaggerated. Anim Behav 125:e17–e24. https://doi.org/10.1111/j.1558-5646.1997.tb02441.x

Podos J, Hendry AP (2006) The biomechanics of ecological speciation. In: Herrel A, Speck T, Rowe NP (eds) Ecology and biomechanics. CRC, Boca Raton, pp 301–321

Podos J, Nowicki S (2004) Performance limits on birdsong. In: Marler P, Slabbekoorn H (eds) Nature's music: the science of birdsong. Academic, New York, pp 318–342

Podos J, Patek SN (2015) Acoustic signal evolution: biomechanics, size and performance. In: Irschick DJ, Briffa M, Podos J (eds) Animal signalling and function: an integrative approach. John Wiley and Sons Inc, Hoboken, pp 175–203

Podos J, Huber SK, Taft B (2004) Bird song: the interface of evolution and mechanism. Ann Rev Ecol Evol Syst 35:55–87. https://doi.org/10.1146/annurev.ecolsys.35.021103.105719

Podos J, Lahti DC, Moseley, DL (2009) Vocal performance and sensorimotor learning in songbirds. In Naguib M, Zuberbuhler K, Clayton NS, Janik VM (eds) Ad Stud Behav 40:159–195

Podos J, Dybboe R, Jensen MO (2013) Ecological speciation in Darwin's finches: parsing the effects of magic traits. Curr Zool 59:8–19. https://academic.oup.com/cz/article/59/1/8/1811479

Podos J, Moseley DL, Goodwin SE, McClure J, Taft BN, Strauss AVH, Rega-Brodsky C, Lahti DC (2016) A fine-scale, broadly applicable index of vocal performance: frequency excursion. Anim Behav 116:203–212. https://doi.org/10.1016/j.anbehav.2016.03.036

Porzio NS, Repenning M, Fontana CS (2018) Do beak volume and bite force influence the song structure of sympatric seedeaters (Thraupidae: *Sporophila*)? Emu: Austral Ornithology 119:71–78. https://doi.org/10.1080/01584197.2018.1535833

Prather JF, Peters S, Mooney R, Nowicki S (2012) Sensory constraints on birdsong syntax: neural responses to swamp sparrow songs with accelerated trill rates. Anim Behav 83(6):1411–1420. https://doi.org/10.1016/j.anbehav.2012.03.012

Prum RO (2012) Aesthetic evolution by mate choice: Darwin's really dangerous idea. Phil Trans Roy Soc B 367(1600):2253–2265. https://doi.org/10.1098/rstb.2011.0285

Prum RO (2017) The evolution of beauty: how Darwin's forgotten theory of mate choice shapes the animal world – and us. Doubleday Press, New York

Reichert MS, Gerhardt HC (2012) Trade-offs and upper limits to signal performance during close-range vocal competition in gray tree frogs *Hyla versicolor*. Am Nat 180(4):425–437. https://doi.org/10.1086/667575

Riede T, Goller F (2014) Morphological basis for the evolution of acoustic diversity in oscine songbirds. Proc Roy Soc B 281(1779):20132306. https://doi.org/10.1098/rspb.2013.2306

Riede T, Thomson SL, Titze IR, Goller F (2019) The evolution of the syrinx: an acoustic theory. PLoS Biol 17(2):e2006507. https://doi.org/10.1371/journal.pbio.2006507

Ryan MJ (2018) A taste for the beautiful: the evolution of attraction. Princeton University Press, Princeton

Sakata JT, Vehrencamp SL (2012) Integrating perspectives on vocal performance and consistency. J Exp Biol 215(2):201–209. https://doi.org/10.1242/jeb.056911

Schraft HA, Medina OJ, McClure J, Pereira DA, Logue DM (2017) Within-day improvement in a behavioural display: wild birds 'warm up. Anim Behav 124:167–174. https://doi.org/10.1016/j.anbehav.2016.12.026

Searcy WA, Beecher MD (2009) Song as an aggressive signal in songbirds. Anim Behav 78(6):1281–1292. https://doi.org/10.1016/j.anbehav.2009.08.011

Searcy WA, Nowicki S (2005) The evolution of animal communication: reliability and deception in signaling systems. Princeton University Press, Princeton

Secora KR, Peterson JR, Urbano CM, Chung B, Okanoya K, Cooper BG (2012) Syringeal specialization of frequency control during song production in the Bengalese Finch (*Lonchura striata domestica*). PLoS One 7(3):e34135. https://doi.org/10.1371/journal.pone.0034135

Simmons LW, Emlen DJ (2006) Evolutionary trade-off between weapons and testes. Proc Natl Acad Sci U S A 103(44):16346–16351. https://doi.org/10.1073/pnas.0603474103

Slabbekoorn H, Smith TB (2000) Does bill size polymorphism affect courtship song characteristics in the African finch *Pyrenestes ostrinus*? Biol J Linn Soc 71:737–753. https://academic.oup.com/biolinnean/article/71/4/737/2638714

Sprau P, Roth T, Amrhein V, Naguib M (2013) The predictive value of trill performance in a large repertoire songbird, the nightingale *Luscinia megarhynchos*. J Av Biol 44(6):567–574. https://doi.org/10.1111/j.1600-048X.2013.00113.x

Stowell D, Plumbley MD (2014) Large-scale analysis of frequency modulation in birdsong data bases. Methods Ecol Evol 5(9):901–912. https://doi.org/10.1111/2041-210x.12223

Suthers RA (1990) Contributions to birdsong from the left and right sides of the intact syrinx. Nature 347(6292):473–477. https://doi.org/10.1038/347473a0

Suthers RA, Vallet E, Kreutzer M (2012) Bilateral coordination and the motor basis of female preference for sexual signals in canary song. J Exp Biol 215(17):2950–2959. https://doi.org/10.1242/jeb.071944

ten Cate C (2004) Birdsong and evolution. In: Marler P, Slabbekoorn H (eds) Nature’s music: the science of birdsong. Elsevier Academic Press, San Diego, pp 296–317

Vallet E, Kreutzer M (1995) Female canaries are sexually responsive to special song phrases. Anim Behav 49(6):1603–1610. https://doi.org/10.1016/0003-3472(95)90082-9

Vallet E, Beme I, Kreutzer M (1998) Two-note syllables in canary songs elicit high levels of sexual display. Anim Behav 55:291–297. https://doi.org/10.1006/anbe.1997.0631

Vehrencamp SL, Yantachka J, Hall ML, de Kort SR (2013) Trill performance components vary with age, season, and motivation in the banded wren. Behav Ecol Sociobiol 67(3):409–419. https://doi.org/10.1007/s00265-012-1461-x

Vehrencamp SL, Ellis JM, Cropp BF, Koltz JM (2014) Negotiation of territorial boundaries in a songbird. Behav Ecol 25(6):1436–1450. https://doi.org/10.1093/beheco/aru135

Wagner WE, Beckers OM, Tolle AE, Basolo AL (2012) Tradeoffs limit the evolution of male traits that are attractive to females. Proc Roy Soc B 279(1739):2899–2906. https://doi.org/10.1098/rspb.2012.0275

Wainwright PC (1988) Morphology and ecology: functional basis of feeding constraints in Caribbean labrid fishes. Ecology 69(3):635–645. https://doi.org/10.2307/1941012

Wainwright PC, Reilly SM (eds) (1994) *Ecological Morphology: Integrative Organismal Biology*. University of Chicago Press, Chicago

Webster MS, Podos J (2018) Acoustic communication. In: Morrison ML, Rodewald AD, Voelker G, Colón MR, Prather JF (eds) *Ornithology: Foundation, Critique, and Application*. Johns Hopkins University Press, pp 409–436

Weiss M, Kiefer S, Kipper S (2012) Buzzwords in Females’ Ears? The use of buzz songs in the communication of nightingales (*Luscinia megarhynchos*). PLoS One 7(9):e45057. https://doi.org/10.1371/journal.pone.0045057

Westneat MW, Long JH, Hoese W, Nowicki S (1993) Kinematics of birdsong: functional correlation of cranial movements and acoustic features in sparrows. J Exp Biol 182:147–171

Wilkins MR, Seddon N, Safran RJ (2013) Evolutionary divergence in acoustic signals: causes and consequences. Trends Ecol Evol 28(3):156–166. https://doi.org/10.1016/j.tree.2012.10.002

Wilson DR, Bitton PP, Podos J, Mennill DJ (2014) Uneven sampling and the analysis of vocal performance constraints. Am Nat 183(2):214–228. https://doi.org/10.1086/674379

Woolley SC, Doupe AJ (2008) Social context-induced song variation affects female behavior and gene expression. PLoS Biology 6(3):e62. https://doi.org/10.1371/journal.pbio.0060062

Zollinger SA, Suthers RA (2004) Motor mechanisms of a vocal mimic: implications for birdsong production. Proc Roy Soc B 271(1538):483–491. https://doi.org/10.1098/rspb.2003.2598