Santosh A. Helekar *Editor*

Animal Models of Speech and Language Disorders

Springer

Animal Models of Speech and Language Disorders

Santosh A. Helekar
Editor

Animal Models of Speech and Language Disorders

Springer

Editor
Santosh A. Helekar
Department of Neurology
and Neuroscience
The Methodist Hospital Research
Institute and Weill Medical
College of Cornell University
Houston, TX, USA

ISBN 978-1-4614-8399-1 ISBN 978-1-4614-8400-4 (eBook)
DOI 10.1007/978-1-4614-8400-4
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013949275

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Dysfunction of motor, sensory, and cognitive aspects of speech and language forms a substantial component of the clinical presentation in neurological practice. It is frequently present in the most common neurological conditions such as stroke and Parkinson's disease, as well as many other neurodegenerative diseases, and neurodevelopmental conditions such as autism spectrum disorders. Primary speech and language disorders also constitute a significant burden to the well-being of our society. For example, developmental stuttering afflicts more than 1 % of the population, causing great emotional and social discomfort to people who suffer from it. There is a great need to find rational approaches to alleviate the suffering caused by all these conditions. Toward this end, scientific research is focused on discovering their underlying causal mechanisms in the brain. However, despite several decades of sustained effort, we are no closer to understanding these mechanisms, as they pertain specifically to speech and language pathology.

Experience from other areas of biomedical research suggests that, arguably, an important reason for this lack of progress is the perceived absence of elementary animal models of speech and language pathologies or a reluctance to recognize them. The latter circumstance results from the fact that speech and language have long been thought to be unique attributes of the human species. Psychological and neurological research in these areas has therefore been confined to human subjects and patients. However, other frontiers of biomedical research have shown us that animal models often pave the way to understanding a disorder at the causal and mechanistic level and thereby enable researchers and physicians to devise rational strategies toward therapy.

Basic research over the last 3 decades or so has uncovered similarities between speech, especially its sensorimotor aspects, and vocal communication in several nonhuman species. The most comprehensive studies so far have been conducted in songbirds. Songbirds offer us a model system to study the interactions between developmental or genetic predispositions and tutor-dependent influences, on the learning of vocal communication. Songbird research has elucidated cellular and molecular mechanisms underlying learning and production of vocal patterns,

auditory processing and perception of vocal sounds, vocal motor control, and vocal neuromotor plasticity. More recently, the entire genome of the songbird zebra finch has been sequenced. These discoveries, along with the identification of several genes implicated in familial human speech and language disorders, have made it possible to look for analogues of speech and language dysfunction in zebra finches, at least at the perceptual and sensorimotor levels. Two approaches in particular have led us closer to the development of animal models of human speech conditions, namely, developmental stuttering and a familial verbal dyspraxia associated with a mutation in the gene for the transcription factor FoxP2.

Work on other animals that show developmental sensorimotor learning of vocal sounds used for communication has also shown significant progress, leading to the possibility of development of models of speech and language dysfunction in them. In nonhuman primates, while vocal learning per se is not very prominent, investigations on their communicative abilities have thrown some light on the rudiments of language. As far as auditory processing is concerned, echolocation in bats has long served as a rich source of fundamental insights.

There is a great need for a synthesis of all observations and ideas that have emerged from basic and clinical research into the neuroscience of vocalization and auditory processing, in order to develop a rational animal model-based framework for understanding and management of speech and language pathologies. The ultimate goal of satisfying this need makes the publication of this book focused on animal models of speech and language disorders, detailing the overall investigative approach of neurobehavioral studies in animals capable of vocal communication and learned vocalizations, a much-needed and worthwhile project. This book is arguably the first of its kind. I believe it serves as a unifying review of research in a new multidisciplinary frontier, spanning the molecular to the behavioral, for clinicians and researchers, as well as a teaching resource for advanced speech pathology and neuroscience students.

The book covers a wide range of disciplines related to speech and language and vocal communication in animals. In Part I, the first chapter deals with the current state of understanding of the neurology of speech and language in terms of brain substrates, representation, and theoretical models. The second chapter is a review of what is known about the genetics of speech and language disorders with special emphasis of the FoxP2 gene mutations, on which there is the most amount of new information. In Part II, Chap. 3 introduces and discusses the behavioral and physiological aspects of the songbird model of vocal learning. It focuses on developmental time scales of changes in vocal sounds and sequences of sound, as well as motor control and the role of sleep in these processes. The auditory pathway for encoding and processing of vocal signals is discussed in Chap. 4. Chapter 5 describes the findings of the zebra finch genome research and its application to molecular biological studies on song learning. The latter task has been extended in Chap. 6 to include current and prospective ways to elucidate detailed molecular mechanisms with translational significance. Chapter 7 wraps up the section on songbird neurobiology by proposing an elementary birdsong-based model of stuttered speech in zebra finches and discusses the possible involvement of perceptual and synaptic plasticity and neuromodulatory influences in the underlying mechanisms.

The last section is concerned with vocal signaling in three different groups of mammals that have contributed substantially to our understanding of neurophysiological and/or cognitive aspects of social communication. In Chap. 8 the authors present a compilation of findings on the rich variety of calls used by bats for communication and echolocation and the manner in which they are processed at the neuronal level. Chapter 9 concentrates on social communication in New World monkeys and the extent to which the study of their complexity and cognitive role contributes to gleaning insights into the rudiments of grammar and meaning of vocal sounds. The final chapter is an examination of the accumulated knowledge on gestures and socially significant sounds produced by our closest evolutionary ancestors, the great apes, in terms of their relevance to evolution of speech and language and their shared brain substrates in man and ape.

I am fortunate to have an illustrious panel of experts graciously agree to write chapters on their respective areas of research and teaching for this book. I am deeply grateful to them for their painstaking efforts, as well as to other experts who have reviewed their chapters and offered valuable suggestions to make them better. This book would not have been possible without the kind assistance, guidance, and hard work of Melissa Higgs and Elektra McDermott of Springer. I hope the readers will find this joint endeavor of ours useful and informative.

Houston, TX, USA Santosh A. Helekar

Contents

Part III Mammalian Models of Vocal Communication

Part I
Introduction to Speech and Language Disorders

Chapter 1
Neurology of Speech and Language

David B. Rosenfield

Abstract Advances in brain imaging, neurophysiology, and computer analysis and modeling provide substantial recent advances in understanding the neurological basis of human language. The property of language, a phenomenon unique to human beings, has an increasingly identifiable underlying cerebral architecture, providing cogent models for understanding clinical and experimental paradigms. This chapter explores the neurological basis for language, reviewing different models and perspectives.

Keywords Speech • Language • Wernicke's area • Broca's area • Speech production

Introduction

One of the most unique faculties of human beings is the property of language. We learn tens of thousands of words/symbols during our respective lifetimes, almost without effort. A 5-year-old child learns 20 words per day, without even attempting to do so. We not only learn how to produce and decode words, whether through writing, listening, or reading, but also acquire and produce meaningful syntax and grammar.

Human beings have a "generative grammar," such that any individual on the planet can produce a sentence heretofore never produced, yet all individuals within that speaker's sphere of language understand this is a normal sentence [1].

Thus, one has probably never heard, "I enjoy attending opera performances and am reading a book about language," yet everyone knows this is a normal sentence. Mix up the words and say, "Opera enjoy read language book me," and all recognize

D.B. Rosenfield, M.D. (✉)
Department of Neurology, The Methodist Hospital Neurological Institute,
6560 Fannin Street, Ste 802, Houston, TX 77030, USA
e-mail: drosenfield@tmhs.org

S.A. Helekar (ed.), *Animal Models of Speech and Language Disorders*,
DOI 10.1007/978-1-4614-8400-4_1,

this is abnormal. Our brains are uniquely capable of producing these productions of sounds and processing their meaning, and recognize normality as well as abnormalities.

Nonhuman primates (NHP), as well as avian species, do not possess capability of language. These animals do possess a system of communication, signaling needs and statements pertaining to sex, aggrandizement of territory, dominance, submission, etc., but they lack formal language. No nonhuman animal has a generative grammar and as large a repertoire of symbols and words as do human beings. Thus, despite some contention decades previously, no one can teach a monkey the following command (a property of language): "Go to the nearest delicatessen, pick up a copy of 'The New York Times,' and if they don't have the 'Times,' bring me a bagel."

Further, human language is a "representational system," in which letters, words, and various symbols have a particular meaning. This system, combined with our generativity of grammar, "drives" the speech motor control system (SMCS), accessing neurological control through the brain, brainstem, and peripheral nervous system. These SMCS mechanisms then access neuromotor production of respiration, phonation (i.e., laryngeal sounds), and articulation, the latter employing movements of the tongue, lips, jaw, and palate in concert with breathing and laryngeal adjustment to produce the sounds we term "speech."

Speech should not be obfuscated with language. Having a hot potato in your mouth or a cleft palate or being a person who stutters does not translate into compromised language. These individuals all know what they want to say, but cannot do so. Their language is normal; their speech is not [1].

How and why do human beings have this capability? In this chapter, we address these issues by analyzing structure and function within the human brain and various models that explain our unique capability of language.

Neurophysiology and Neuroanatomy of Language

Language very much involves cognitively processing what we hear and see (i.e., reading is a part of language) and cognitively orchestrating what we produce (speech). Considerable research has been done on related mechanisms of input, pertaining to processing auditory and visual input in animals. Most investigators agree that humans and NHP have similar physiological mechanisms for modeling visual and auditory processing. There is a "what" and a "where" system that illustrates this concept (see Fig. 1.1).

After the brain receives visual input from the ocular/retina system, this data proceeds to the primary visual cortex (V1) and then anterior to the inferior temporal (IT) lobe for analyzing "what" is the object, and to the adjacent posterior parietal cortex (PPC) for ascertaining "where" is the object.

Similarly, auditory input arrives at the temporal lobe which then provides information to the superior temporal (ST) lobe for analysis of "what" is heard and to the PPC for localization ("where") of the sound. Then, as Fig. 1.1 denotes, the temporal

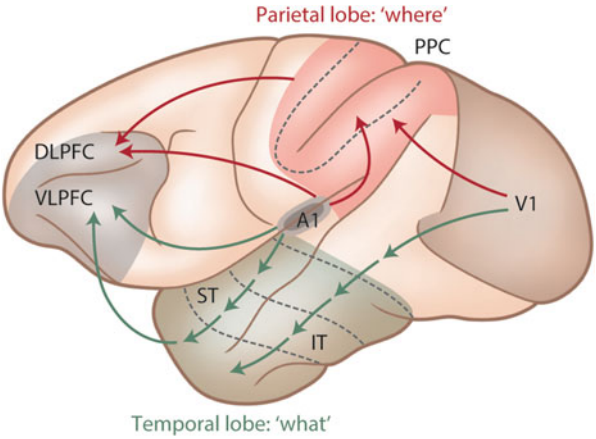


Fig. 1.1 Dual processing for "what" and "where" [4]

lobe system provides the "what" information to the frontal lobe (ventrolateral prefrontal cortex, VLPFC) and the parietal lobe provides the "where" information to the dorsolateral prefrontal cortex (DLPFC) of the frontal lobe.

These connections, reviewed in detail elsewhere [2–4], are not as robust or extensive in NHP as in the human brain. Auditory input, as well as visual input, is very much involved in our language system, and in this context it is helpful to discuss the cerebral connections within the brain that subserve language.

In order to address appropriately these connections, it is important to understand which regions in the brain strongly relate to language: The superior temporal area adjacent to auditory cortex is very much involved in processing the input of what we hear; the inferior frontal cortex adjacent to articulatory motor cortex is important for motoric production of speech. These connections are especially important in the left side of the brain in individuals who are right-handed, and are increasingly shared with the right side of the brain in those who are left-handed [2, 3, 5]. Below, we review important connections between these two areas.

Within the context of production of language, posterior auditory regions connect to anterior motor regions through three main cerebral tracts: extreme capsule, uncinate fasciculus, and the arcuate fasciculus. These connections exist in NHP but are relatively larger in humans, especially the arcuate fasciculus [6].

Models of Speech and Language Production

The classical areas of language function are Broca's area (BA), recognized as orchestrating motor output of speech/language, and Wernicke's area (WA), recognized as processing auditory input. These two areas interrelate, the perspective being that WA subliminally hears what is to be said, transits information to BA,

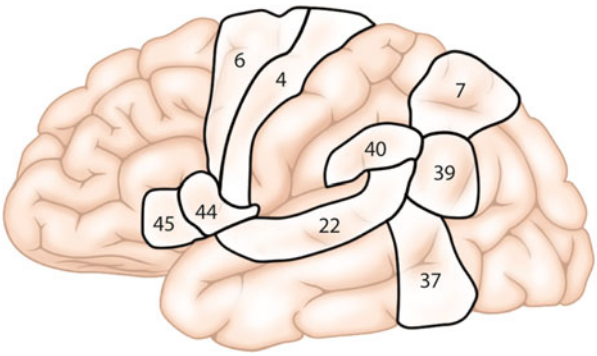


Fig. 1.2 Norman Geschwind's classical approach for cerebral processing of language

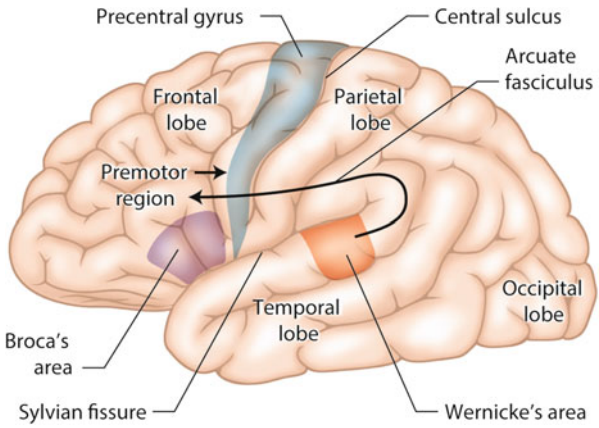


Fig. 1.3 Cerebral areas important in language [14]

and then "hears" whether errors were produced. Again, this paradigm occurs primarily within the left hemisphere of right-handed individuals, moving more to the right hemisphere in those who are left-handed.

In the 1960s and 1970s, Norman Geschwind popularized the above model and recognized the heretofore discarded German literature on Disconnection Syndromes, heralding the field of Behavioral Neurology in the United States. At that time, brain imaging was essentially nonexistent and analysis of speech motor output was embryonic. Nevertheless, Geschwind provided a salient clinical model to explain various types of clinically acquired disruption of language (e.g., aphasias), and his model of relating compromise of comprehension (mild in BA, severe in WA) and fluency/ease of production of words (significantly compromised in BA, minimally compromised in WA) resulted in multiple experimental paradigms that expanded knowledge in the area, contributing to the fields of neuropsychology, neurolinguistics, and psycholinguistics (Figs. 1.2 and 1.3) [2].

One of the difficulties of modeling production and comprehension in language through connections between BA and WA is that neither BA nor WA is cytoarchitectonically distinct: One cannot singularly identify BA or WA under the microscope. Indeed, BA consists of portions of the pars triangularis (PTR) and pars opercularis (PO). PTR is Brodmann area #45, composed of heteromodal cortex and located in the inferior frontal gyrus. The PO is Brodmann area #44, composed of motor association cortex and adjacent to the PTR. Wernicke's area consists of a portion of Brodmann area #22, especially an area termed the "planum temporale," and is composed of auditory association cortex. Thus, BA consists of two areas (#44, #45) and WA consists of a portion of #22 [3, 7].

The BA-WA model, despite clinical salience and popularity for experimental designs, early had certain problems. When external electrical stimulation was applied to the left hemisphere of the brain exposed in patients who were awake, undergoing surgery for epilepsy, stimulation anywhere in the left hemisphere, including BA and WA, caused individuals who were talking to cease talking, as it did when stimulating the motor strip (i.e., anterior central sulcus) in the right hemisphere. And if BA or WA were stimulated in individuals who were not talking, only a grunt type of sound was produced [8].

It became clear the connections between BA and WA were bidirectional and polysynaptic (reviewed in [9]). Further, BA had no connections below the midbrainstem and none to the final neural outflow to the laryngeal muscles (e.g., furthest distal connections are to the periaqueductal gray, with none going to the n. ambiguous) [10].

The above BA-WA system implied an "input system" that filtered sensory input (e.g., auditory or visual input) in a feed-forward manner, consonant with other theories of input systems, in which perceptual processes, after interaction with attention, emotion, and memory modules, influenced motor output systems.

Thus, many investigators posited a sensory receptive system, residing in posterior superior cortex of the temporal lobe that translated data anteriorly to the premotor and motor systems within the frontal cortex. In this paradigm, the motor output of the cortex of the frontal lobe was dependent upon the perceptual and cognitive-related systems, considerably dependent upon the cortex of the temporal lobe.

Recent data has very much reformed and improved upon this modeling. It is now recognized that there are neurons (e.g., "mirror neurons") that are capable of sensorimotor interactions. Within this context, experiments in monkeys reveal that premotor F5 neurons, thought to be similar but not isomorphic to Brodmann area #44, are active during formal execution of a particular skilled movement and similarly active during the monkey witnessing/seeing that same skilled movement when performed by another monkey or even a human being. There are now posited multiple action-perception circuits, with an integrated view of perceptual, cognitive, and motor control systems sharing neuronal mechanisms in which the above-noted sensorimotor neurons are increasingly important [6].

The above modeling is reminiscent of Hebbian models, in which neurons adapt during learning. This model maintains that repeated and persistent stimulation of the postsynaptic neurons by presynaptic neurons produces an increase in synaptic

"efficacy," promoting what is termed, synaptic plasticity. A common translation of this perspective is "Cells that fire together become wired together." This type of learning is referred to as Hebbian learning [11].

Pulvermuller and Fadiga [6] nicely review this action-perception learning of speech sounds and the words which we speak. This requires a strong reciprocal connection between the superior temporal gyrus and the inferior frontal gyrus. These connections are especially robust in humans, as opposed to NHP. Although brain bulk need not index brain competence, the left hemisphere-related cerebral laterality of language is consonant with the fact that the arcuate fasciculus, instrumental in connecting WA to BA, is much larger in the left hemisphere than in the right hemisphere of human beings, possibly further explicating the action-perception circuit for speech (note this does not exclude other inputs).

When one utters a syllable, word, or phrase, these self-produced sounds stimulate the auditory cortex, residing in the superior temporal cortex. This activation occurs whether the sounds are uttered in a quiet environment or are whispered and concurrently masked by external noise. This activation increases with increasing speech rate, suggesting a strong relationship between motor output and auditory (input) activation during speech production.

Similarly, when one listens to sounds of speech that require a strong articulation activity, the motor system is similarly activated. Brain investigation studies, including fMRI, transcranial magnetic stimulation, and various neurometabolic paradigms, strongly suggest the inferior frontal premotor cortex and prefrontal cortex, cerebral areas that are very active during the motoric output of speech, are also active during identification and discrimination of speech sounds, as well as during routine ongoing perception of speech. In other words, areas of the brain involved in the production of speech are also very much involved in the perception of speech.

Models of Speech and Language

Many investigators study "functional models" of human speech. These include several perspectives, with some of the more exciting and heuristic models focusing upon oscillating systems within the brain, whereas others study a dual stream functional neuroanatomical paradigm, involving sound-to-articulation mapping and sound-to-meaning mapping.

Underlying all these avenues of modeling of language is the fact that speech is composed of multiple sounds, which in turn are composed of various, varying frequencies changing with time.

Analysis of different acoustic, neurophysiological, and psycholinguistic data pertaining to production and perception of speech/language suggests that underlying organizational principles and perceptual units of analysis can occur at different time scales. Thus, the "sounds" of speech can be divided into complexity of frequency (spectral analysis) and changes over time (temporal modulation). The spectral information is necessary for a listener to process the meaning of the word but, if

impoverished, the listener can still distill the meaning (e.g., "happy birthday" > "hpy brthdy" is still understandable). However, impoverishment of temporal modulation significantly compromises the perception of meaning and comprehension. Oscillation theory is an area of investigation that investigates the temporal analysis of speech from a bottom-up approach (reviewed in [12]), studying how the parceling of sounds can be put together to form meaning and understanding.

As noted above, perceptual units of speech can be analyzed at different time scales. Some of these units are aperiodic but sufficiently rhythmic to elicit regularities in the time domain. Thus, 30–50 Hz time frames are of short duration and high modulation frequency, and relate to phonemes and formant transitions (e.g., changes in spectral peaks of the sound spectrum, important for perception of different sounds and meanings). Further, 4–7 Hz relates to syllable rate and 1–2 Hz relates to lexical and phrase units, as well as intonation contour.

Giraud and Poeppel [12] nicely illustrate the theory of oscillation-based operations in the perception of speech. They maintain that an essential ingredient for understanding the meaning of connected speech lies within the infrastructure provided by intrinsic oscillations at rest of neurons and that these neurons are very well suited to dissect sound-related phenomenon that is especially sensitive to time domains. They argue for a robust relationship between time scales present in speech and the time constants (i.e., in electronics, the time required for the current or voltage in a circuit to rise or fall exponentially through approximately 63 % of its amplitude) underlying neuronal cortical oscillations and that this enables the brain to convert speech "rhythms" into meaningful linguistic segments.

These authors maintain that low-gamma (25–35 Hz), theta (4–8 Hz), and delta (1–3 Hz) bands provide a link between neurophysiology, neuronal computations, acoustics, and psycholinguistics. They contend there is a close correspondence between (sub)phonemic, syllabic, and phrasal processing, respectively, with gamma, theta, and delta oscillations as to how the brain processes the temporal data that underlie speech perception.

Other investigators of human language have a different, non-oscillatory perspective toward understanding human speech and language. Thus, psycholinguists traditionally do not focus upon the above-noted cerebral areas of processing but, rather, focus more upon language at an abstract level (reviewed in [13]). These investigators seek generalization of language at the level of phonemes (e.g., the smallest part of a word: "b," "oo," and "k" are phonemes within "book"), morphemes (e.g., the smallest semantic unit—a unit having meaning—in language), lemmas (e.g., the canonical/dictionary form of a set of words—i.e., "see" is the lemma of "seeing," "seen," "sees"), and phrasal units (e.g., prepositional phrases, major portions of sentences that carry meaning).

Researchers investigating motor control theory primarily are concerned with kinematic forces, movement trajectories, and control of feedback, often focusing upon lower-level articulatory control (e.g., voiced v. voiceless sounds—/z/ v. /s/, /v/ v. /f/; lingual sounds—/l/; labial sounds—/p/, /b/; voice-onset time differences—/pa/ v. /ba/, /ta/ v. /da/) and other similar parameters.

Ideally, since psycholinguistic and motor control theories investigate speech and language, although on different levels of organization and output, it would be ideal

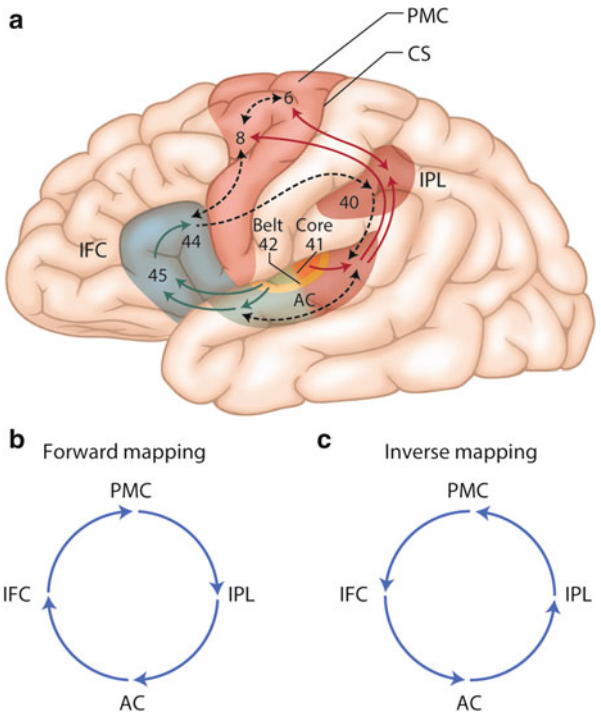


Fig. 1.4 Model of Rauschecker and Scott 2009 [4]. (**a**) Ventral (*blue*) and dorsal (*red*) streams and their relationship to the auditory cortex (AC). (**b**) The ventral stream initiates forward mapping between representations in the auditory cortex (AC), inferior frontal cortex (IFC, area 45), ventral premotor cortex (PMC, area 44) and inferior parietal lobule (IPL, and superior temporal cortex). (**c**) The dorsal stream initiates inverse mapping between AC representations, attention- or intention-dependent patterns in the IPL and context-dependent action programs in the PMC

to combine both perspectives. Within this context, recognizing that sensory representations, including auditory as well as visual inputs, are both within auditory and somatosensory cortex provides a hierarchy of targets for speech "gestures."

Auditory targets are predominantly syllabic and comprise a higher-level sensory goal; somatosensory targets represent lower-level goals that correspond loosely to phonemic-level targets. Movement plans that are coded in a corresponding cortical motor hierarchy are selected to hit the sensory targets. This involves an internal feedback control loop, including "forward prediction" and subsequent correction. The sensorimotor integration is achieved in the Sylvian fissure at the parietal-temporal boundary for the higher-level system and via the cerebellum for the lower-level circuit (reviewed in [13, 14]).

Rauschecker and Scott [4] provide a cogent, utilitarian model for scientific investigation as well as clinical paradigms, in which they propose interactions between streams of auditory processing systems and speech production. Their model proposes an anteroventral and posterodorsal auditory "stream" that originates in the auditory region of the superior temporal lobe cortex (Brodmann area #22); the latter interacts

posteriorly with the inferior parietal lobule (IPL), where a template of sensory event information rapidly can be compared to the predicted efferent motor output.

This model also provides forward mapping, in which object information (e.g., speech) is decoded in the anteroventral stream all the way to invariant categories within the IFC (Brodmann area #45) and subsequently be transformed into motor-articulatory representations (Brodmann area #44 and ventral premotor cortex (PMC)), the activation of which is then transmitted to the IPL and posterior superior temporal cortex as an efferent copy (Fig. 1.4).

In reverse direction, this model performs inverse mapping, wherein attention-related or intention-related changes in the IPL affect the context-dependent action programs in PFC and PMC. Just as Geschwind's models of speech and language (see above) provided experimental paradigms for testing normal human beings and patients with acquired cerebral compromise of language (e.g., those afflicted with fluent/dysfluent output compared to those with good/poor comprehension, and all possible interactions of these separate domains) and spawned progress in neuropsychology and language-related fields, models akin to that proposed by Rauschecker and Scott [4] similarly provide numerous avenues of scientific examinations, now focusing upon magnetoencephalography in humans, single-unit studies in monkeys, and numerous brain-imaging investigations (e.g., fMRI, PET) previously unavailable, now providing opportunity to dissect the components and ingredients of human language.

Conclusion

There is increasing understanding of the cerebral basis for human language, with expanding knowledge of how the human brain understands what it is we say and how the brain produces speech and language. Theories and models include computational analysis of sound waves, theories of motor production and perception, and various interactions of these perspectives.

References

1. Rosenfield DB, Barroso AO (2000) Difficulties with speech and swallowing. In: Bradley WG, Daroff RB, Fenichel GM, Marsden CD (eds) Neurology in clinical practice: principles of diagnosis and management, 3rd edn. Butterworth-Heinemann, Boston, pp 171–186
2. Schacter S, Devinsky O (1997) Behavioral neurology and the legacy of Norman Geschwind. Lippincott-Raven, Philadelphia
3. Mesulam M-M (2000) Principles of behavioral neurology, 2nd edn. Oxford University Press, New York
4. Rauschecker JP, Scott SK (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat Neurosci 12(6):718–724
5. Raphael LJ, Borden GJ, Harris KJ (2011) Speech science primer: physiology, acoustics, and perception of speech, 6th edn. Lippincott Williams and Wilkins, Baltimore
6. Pulvermuller F, Fadiga L (2010) Active perception: sensorimotor circuits as a cortical basis for language. Nat Rev Neurosci 11(5):351–360

7. Foundas AL, Bollich AM, Corey DM, Hurley M, Heilman KM (2001) Anomalous anatomy of speech-language areas in adults with persistent developmental stuttering. Neurology 57(2):207–215
8. Penfield W, Jasper H (1954) Epilepsy and the functional anatomy of the human brain. Churchill/Livingstone, London
9. Brown J (1972) Aphasia, apraxia and agnosia. Charles C. Thomas, Springfield
10. Muller-Preuss P, Jurgens U (1976) Projections from the “cingular” vocalization area in the squirrel monkey. Brain Res 103(1):29–43
11. Hebb DO (1949) The organization of behavior. Wiley, New York
12. Giraud AL, Poeppel D (2012) Cortical oscillations and speech processing: emerging computational principles and operations. Nat Neurosci 15(4):511–517
13. Hickok G (2012) Computational neuroanatomy of speech production. Nat Rev Neurosci 13(2):135–145
14. Hickok G, Poeppel D (2007) The cortical organization of speech processing. Nat Rev Neurosci 8(5):393–402

Chapter 2
Genetic Pathways Implicated in Speech and Language

Sonja C. Vernes and Simon E. Fisher

Abstract Disorders of speech and language are highly heritable, providing strong support for a genetic basis. However, the underlying genetic architecture is complex, involving multiple risk factors. This chapter begins by discussing genetic loci associated with common multifactorial language-related impairments and goes on to detail the only gene (known as *FOXP2*) to be directly implicated in a rare monogenic speech and language disorder. Although *FOXP2* was initially uncovered in humans, model systems have been invaluable in progressing our understanding of the function of this gene and its associated pathways in language-related areas of the brain. Research in species from mouse to songbird has revealed effects of this gene on relevant behaviours including acquisition of motor skills and learned vocalisations and demonstrated a role for Foxp2 in neuronal connectivity and signalling, particularly in the striatum. Animal models have also facilitated the identification of wider neurogenetic networks thought to be involved in language development and disorder and allowed the investigation of new candidate genes for disorders involving language, such as CNTNAP2 and FOXP1. Ongoing work in animal models promises to yield new insights into the genetic and neural mechanisms underlying human speech and language.

Keywords FOXP2 • Language genetics • Development • Speech and language • Specific language impairment • Transcription factor • CNTNAP2 • FOXP1

S.C. Vernes, D.Phil. (✉) • S.E. Fisher, D.Phil.
Max Planck Institute for Psycholinquistics, Wundtlaan 1,
Nijmegen 6525 XD, The Netherlands
e-mail: Sonja.Vernes@mpi.nl; simon.fisher@mpi.nl

S.A. Helekar (ed.), *Animal Models of Speech and Language Disorders*,
DOI 10.1007/978-1-4614-8400-4_2,

Heritability of Language and Language Disorders

There is considerable evidence to suggest that genes are important for directing developmental processes necessary for the normal use of speech and language. Furthermore, disorders that disrupt speech and language development have been shown to be highly heritable, providing strong support for a genetic basis for language impairments.

A number of well-studied neurodevelopmental disorders involve speech and/or language deficits as one part of a broader profile of symptoms. Examples include autism spectrum disorders (ASD, OMIM: 209850), fragile X syndrome (FXMR, OMIM: 300624) and Angelman syndrome (AS, OMIM: 105830). However, there are developmental disorders where the central or primary deficit is in the comprehension, processing and/or use (vocal or nonvocal) of language. These disorders, such as specific language impairment (SLI, OMIM: 606711), developmental dyslexia (DD, OMIM: 127700) and developmental verbal dyspraxia or childhood apraxia of speech (DVD or CAS, OMIM: 602081) can shed light not only on the genetic and developmental underpinnings of impairments but also on pathways involved in normal language development.

The first clues to the heritability of developmental disorders of speech and language came from observations of familial clustering. Children with language disorders are much more likely to have family members displaying speech, reading or language impairments than typically developing children [1–4]. The importance of genetic influences on language impairments was further illustrated by a study showing that children who had an affected parent but that had been adopted into a language rich environment were significantly more likely to suffer from language disorders than adopted children without a family history [5].

The magnitude of the genetic contribution to a disorder or trait can be investigated using heritability estimates. These can be calculated by comparing the rate of coinheritance of the disorder in monozygotic twins (considered to be genetically near identical) to that of dizygotic twins (who, like siblings, are ~50 % genetically similar). Assuming that mono- and dizygotic twins are subject to similar levels of shared environment during development and childhood, a genetically influenced disorder should co-occur more frequently in monozygotic twins than it does in dizygotic twins. Indeed, monozygotic twins show a higher concordance of language disorders, as well as more closely matched phenotypes within these disorders, as compared to their dizygotic counterparts [6]. Concordance has been reported to be near 100 % for monozygotic twins and between ~50 and 70 % for dizygotic twins, arguing strongly for a genetic component to language disorder [7, 8].

More recently, studies have focused on longitudinal measures in the normal range of abilities. These investigations found that while early language development (~2–5 years old) could largely be accounted for by environmental factors, linguistic skills showed higher heritability (54–60 %) between the ages of 7 and 12 years and very high heritability scores (~85 %) for long-term linguistic ability (up to age 18) [9–11].

Genetic Risk Factors for Complex Language Disorders

Evidence thus far suggests that the majority of language impairments are not caused by just a single gene acting in a Mendelian manner or by only a single region of the genome [7, 12]. Rather it appears likely that, in most cases, many different risk alleles spread around the genome make small contributions to the observed language phenotypes. A number of genetic mapping studies have attempted to define relevant regions of the genome and to pinpoint the key genes, but the complex multifactorial nature of the traits and the small effect sizes involved makes identification of the genetic risk factors challenging [13].

SLI is the most common form of language disorder, with approximately 7 % of school age children reported to meet diagnostic criteria [14, 15]. SLI is classified as the failure to develop normal speech and language skills in the absence of any environmental, medical or genetic impairments (e.g. hearing loss, mental retardation or other overt neurological disorders) [15]. In the first molecular investigations of common forms of SLI, researchers used DNA from multiple families to search through the genome for genetic markers whose inheritance may be linked to the trait (referred to as 'linkage analysis'). A genome-wide analysis for linkage to quantitative measures of language in 98 UK families identified two candidate regions: SLI1 (located at chr 16q, OMIM: 606711) and SLI2 (located at chr 19q, OMIM: 606712) [16]. These findings were replicated in a follow-up study of a further 86 UK families; in particular, the SLI1 region demonstrated highly significant linkage to deficits in non-word repetition (NWR), the ability to correctly repeat nonsense words, which has been proposed as a core feature of SLI [17]. High-density screening of the SLI1 region in an expanded set of the UK families and an independent population cohort identified association to two candidate genes, one encoding a calcium-transporting ATPase, ATP2C2, the other encoding c-maf-inducing protein, CMIP [18]. A third candidate region, SLI3 (located at chr 13q21, OMIM: 607134), was identified in an independent study of language impairment in 5 Canadian families [14], and linkage of this region to reading impairment was demonstrated in a follow-up study of 22 families from the USA [19].

With rapid advances in molecular technologies, more fine-grained and wide-ranging analysis of the genome has become possible, which will likely lead to the identification of further genomic regions and candidate genes contributing to language or language-related disorders. In order to make sense of these findings, it will be necessary to understand more about the phenotypes of the different language impairments and how they relate to each other. For example, it is likely that rather than being distinct syndromes, the spectrum of disorders that involve language impairments represent overlapping groups of syndromes that share endophenotypes (measurable components on the path between global phenotype and distal genotype), each of which might present along a distribution of severity. For example, a gene that is a risk factor for SLI may also be found to be a risk factor in some but not all individuals that meet the criteria for autism (a developmental disorder primarily affecting social interaction, verbal and non-verbal social communication and

repetitive, stereotyped behaviours) or dyslexia (an impairment of reading and spelling). How disorders are classified and how subjects are chosen for inclusion in studies will greatly influence our ability to detect shared or independent genetic factors underlying language and language impairment.

Studies investigating the underlying genetic factors contributing to dyslexia and autism are outside the scope of the current chapter, but there are a number of articles that comprehensively review this topic [20–24].

A Monogenic Speech and Language Disorder

As noted above, the vast majority of cases of language impairment are likely to have a complex genetic basis. However, in the late 1980s clinical geneticists came across an unusual large family showing an apparently simple inheritance pattern for their speech and language problems [25]. In this pedigree, known as the 'KE family', approximately half of the 30 family members, spread over three generations, suffered from a severe form of speech and language disorder [25]. The pattern of transmission observed in this family was consistent with simple autosomal dominant inheritance—highly suggestive that the disorder was monogenic that is due to disruption of just a single gene being passed from one generation to the next [25].

When researchers performed gene-mapping studies, they were able to formally demonstrate the monogenic nature of the disorder and pinpointed a small region of chromosome 7 (designated the SPCH1 locus) that was very likely to contain the causative gene [26]. The identification of an unrelated patient who had a highly similar speech and language disorder phenotype was key to determining which gene in the SPCH1 region was responsible [27]. This child (known as CS) carried a de novo translocation, involving a breakpoint in the SPCH1 region of chromosome 7. The investigators discovered that this breakpoint directly interrupted a previously unidentified gene known as *FOXP2*, and they hypothesised that disruption of this gene was responsible for the phenotype seen in CS. They went on to sequence the same gene in the KE family and found that all affected members carried a point mutation affecting a single nucleotide in the coding region of *FOXP2*, known as the R553H mutation (explained further below) [27]. This mutation was never found in unaffected individuals in the family or in the general population, and it was predicted to disturb the function of the gene [27]. Thus, both the CS case and the KE family carried disruptions to the FOXP2 gene, which were potentially causative of their speech and language problems.

The KE Family

The phenotype of the KE family has been studied in detail both at the behavioural/cognitive and the neuroanatomical levels in order to dissect out the core features of

this complex disorder. Early reports posited conflicting hypotheses that the impairment seen in the KE family was largely one of articulation or conversely a grammar-specific disorder [28, 29]. However, the reality is likely to lie somewhere between the two models.

The phenotype observed in the KE family involves severe developmental verbal dyspraxia (known as DVD or childhood apraxia of speech, CAS). DVD is characterised by problems coordinating sequences of mouth/face movements when speaking, such that speech is unintelligible to the naive listener [28]. However, in the KE family, additional severe impairments are also observed in multiple areas of expressive and receptive language, affecting both spoken and written modalities. In studies that assessed a range of abilities in the KE family, tests of nonsense-word repetition (NWR) provided the most reliable metric for distinguishing between affected and unaffected family members [30]. Receptive vocabulary, lexical decision making and verbal fluency, tense production, receptive syntax at word-order level and inflectional and derivational morphology were all found to be significantly impaired in the affected members of the KE family [30, 31]. Furthermore, the orofacial dyspraxia of affected members is not entirely specific to speech. Reduced performance has also been observed in complex and sequential non-verbal oral movements, although single simple movements were unaffected [32]. Rhythm was also affected in tests of both vocal and manual timing, similar to effects reported in some other language disorders [33, 34].

Neuroimaging studies of the KE family have uncovered functional and anatomical correlates of the disorder. Magnetic resonance imaging (MRI) in ten affected and seven unaffected family members observed no overt anatomical differences differentiating the two groups [31]. However, statistical analyses using voxel-based morphometry identified subtle bilateral changes in grey matter density in affected individuals for a number of brain regions implicated in speech and language processing. Significantly reduced grey matter was observed in Broca's area, the supplementary motor area, caudate nucleus of the striatum and the ventral cerebellum, while regions of significantly increased grey matter could be seen in the thalamus, angular gyrus and parts of the cortex, including the sensorimotor and temporal cortex [31, 35]. For most individuals, language is localised to the left hemisphere of the brain. In some cases, when the left hemisphere is damaged (e.g. due to a stroke), if the right hemisphere is unaffected, it can adapt to performing language-related tasks. This process, known as relocalisation, is not thought to be able to occur in the KE family due to the bilateral neuroanatomical changes in grey matter density, which may help to explain why the disorder is so severe and persistent [35].

Functional neuroimaging studies have demonstrated differences in brain activation patterns in affected versus unaffected family members that are suggestive of linguistic processing defects. Aberrant bilateral activation in affected family members during semantic retrieval and articulatory planning was observed by functional magnetic resonance imaging (fMRI) [36]. In one part of this study, imaging was performed during covert (unspoken) verb generation tasks so that any signal arising from articulatory defects could be excluded. Covert verb generation tasks in an unaffected individual (unaffected KE family member or unrelated normal

individual) typically result in activation of the inferior frontal gyrus (Broca's area) in the left hemisphere and (subcortically) the putamen in the striatum [36]. Despite all participants being able to successfully perform the task overtly outside the scanner, the affected KE family members demonstrated strikingly different patterns of activation during the covert task. Significant under-activation of Broca's area and the putamen was observed in affected individuals, accompanied by significant over-activation of diffuse regions of both hemispheres, including Wernicke's area and the precentral gyrus [36].

Thus, in the KE family, a complex speech and language disorder involving receptive and expressive language impairment and associated with anatomical and functional changes in the brain was directly related to a single mutation disturbing the *FOXP2* gene.

The FOXP2 Gene

The *FOXP2* gene is located on human chromosome 7q31, made up of 25 exons (i.e. the expressed parts of the gene) that span a locus of ~600,000 nucleotides of DNA (~600 kb) (Fig. 2.1a). This gene codes for a protein (called the FOXP2 protein) that is able to act as a transcription factor, meaning that it regulates the expression (switching on and/or off) of other genes. The main version of the FOXP2 protein is 715 amino acids long, but, as with most genes and proteins, differential processing (alternative splicing) can sometimes generate alternative versions that are longer or shorter than this.

The FOXP2 protein contains a number of functionally important regions, or 'domains' (Fig. 2.1b). Moving along the protein from one end (the N-terminus) to the other (the C-terminus), the following domains can be identified: a region containing a large number of glutamine (Q) residues (Q-rich), a zinc-finger/leucine zipper region (ZnF/LeuZ), a DNA-binding domain (FOX) and an acidic tail region (Acidic). The FOXP2 protein regulates gene expression by binding to regulatory regions of the genome, usually located close to the start site for the coding regions of genes, and thereby affecting the levels of transcription for these so-called 'target' genes (i.e. altering the amount of gene product that is made). FOXP2 can directly bind to these regulatory regions of DNA via the specialised section of the protein known as the forkhead-box DNA-binding domain (or FOX domain, for short) [37]. The FOX domain is a stretch of ~90 amino acids that folds into a three-dimensional structure which wraps itself around DNA [38]. The FOX domain does not wrap around just any section of DNA, but has a preference for specific sequences of nucleotide letters; thus, it binds only to particular regions of the genome, located within its target genes [39–41]. Following DNA binding, FOXP2 is able to activate (turn on) or repress (turn off) the expression of these target genes [42].

The point mutation identified in the KE family introduced an amino acid change at position 553 of the protein sequence, swapping the arginine (R) that is normally found at this position to a histidine (H) residue; thus, the mutation is known as

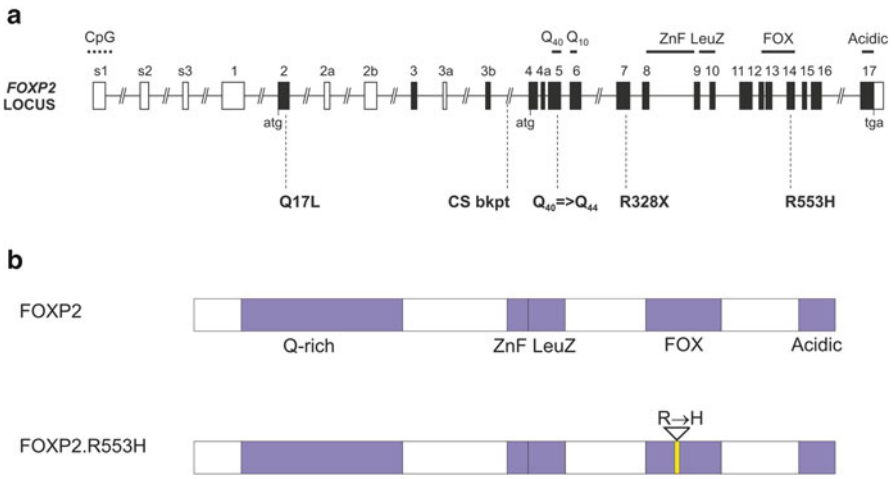


Fig. 2.1 (**a**) Schematic of the human FOXP2 locus, spanning >600 kb. *Black shading* indicates exons that are translated into protein; 'atg' and 'tga' denote start and end codons. Exon s1 overlaps with a type of regulatory region known as a CpG island. Additional information on features of this locus can be found in Fisher et al. [13]. Sites of coding variants reported in children with severe speech and language impairment are indicated below the locus schematic, including the R553H mutation initially identified in the KE family [27], and the three additional changes uncovered in a subsequent screening study of 49 other probands [61]. The figure also shows the site of the translocation breakpoint found in case CS, mapping between exons 3b and 4 [27]. Multiple additional translocation cases involving FOXP2 disruption have since been reported. *Adapted with permission from MacDermot et al. (2005)* [61]. (**b**) Schematic of the major form of the FOXP2 protein (encoded by exons 2–17) contains 715 amino acids, with polyglutamine tracts of 40 (Q40) and 10 residues (Q10) collectively known as the Q-rich region, a zinc-finger motif (ZnF), a leucine zipper (LeuZ), a forkhead domain (FOX) and an acidic C-terminus (Acidic). The location of the KE family mutation (R553H) in the protein is also shown. *Adapted from Vernes et al. (2006)*

R553H [27]. Crucially, position 553 is located in a key part of the FOXP2 DNA-binding domain (Fig. 2.1b) and an arginine is found at this location in every type of normal FOX protein that has been discovered so far (see next section), suggesting it must be important for protein function [27]. Indeed, in laboratory-based tests, introducing the R553H change into an isolated FOX domain, or into the full length FOXP2 protein, in each case abolished DNA binding to a known target sequence and affected the ability of the protein to regulate gene expression in cellular model systems [43]. These assays also suggested that the R553H form of the protein was able to interfere with the activity of normal FOXP2 protein present in the cells [43]. In addition, there was evidence of mis-localisation of the mutant protein within the cell. Normally, the FOXP2 protein is found in the nucleus of cells where it can access DNA in order to regulate gene expression [43]. However, the mutant version of the FOXP2 protein, carrying the R553H change, sometimes showed both nuclear and cytoplasmic localisation [43]. Thus, substitution of this single amino acid had severe and wide-ranging effects on the ability of the protein to function normally.

FOX Transcription Factors

FOXP2 is just one of a large group of transcription factors (known as FOX proteins), all of which carry a highly conserved version of the characteristic FOX DNA-binding domain. This group of proteins is classified into subgroups, based on comparing the sequences of the DNA-binding domain. To date 17 FOX subgroups have been identified, designated FOXA to FOXQ, in order of their discovery [44]. Accepted nomenclature for this gene family uses upper case for human genes (*FOXP*), lower case for mouse genes (*Foxp*) and upper and lower case for all other species (*FoxP*). Proteins are denoted by roman type (FOXP) [45].

FOXP2 falls into the 'P' subgroup of FOX proteins, which also includes FOXP1, FOXP3 and FOXP4 [46, 47]. The three most closely related members, FOXP1/FOXP2/FOXP4, show ~92 % similarity of amino acid sequence in their FOX domain [47], suggesting closely related function. These proteins have also been shown to interact with each other via their ZnF/LeuZ regions. Indeed, homodimerisation (interaction by two of the same FOXP molecules) and heterodimerisation (interaction between two different FOXP family members) appear to be required for efficient binding to target DNA [48]. It is also thought that the glutamine-rich regions of these transcription factors mediate interaction with other proteins to facilitate the assembly of protein complexes around gene regulatory regions. For the mouse homolog of FOXP1, the presence of the glutamine-rich region was found to be capable of fine-tuning the strength of regulation mediated by the protein [42].

FOXP1/2/4 display distinct but overlapping expression patterns in the brain. FOXP2 follows a highly restricted pattern of expression in a range of structures of the brain during development. During foetal development in humans, FOXP2 was shown to be expressed (at around 9–14 weeks gestation) in the thalamus; hypothalamus; developing striatum (caudate-putamen); areas of the cortex including the perisylvian regions and frontal, parietal and occipital cortices; the medulla; and the cerebellum [49, 50]. As development progresses, FOXP2 expression becomes restricted to specific subpopulations of neurons in these regions, for example, to the deep layers of the cortex (layers V/VI), the inferior olivary complex of the medulla and Purkinje cells of the cerebellum [49, 50].

Studies in rodents, which show the same expression patterns as seen in the human tissue tested, have illustrated the combinatorial expression of the Foxp subfamily in the brain. Foxp1 expression overlaps with Foxp2 in a number of regions as both genes are expressed in the developing cortical plate, striatum, thalamus and inferior olives of the medulla [49]. However, while expression overlaps in the developing cortex, in the mature (six-layer) cortex, Foxp2 is restricted to the deepest layers (layer V/VI), while Foxp1 is found throughout layers III–V [49]. By comparison, Foxp4 expression is spread throughout layers II–VI [51].

Unlike Foxp2, Foxp1 and Foxp4 can be found in the developing and adult hippocampus, and while Foxp2 is strongly expressed in the cerebellum, amygdala and olfactory bulb, Foxp1 is completely absent from these regions [49]. Foxp4 expression overlaps with Foxp2 in the developing striatum, olfactory bulb and Purkinje cells of the cerebellum [51, 52]. Foxp2 and Foxp4 are both expressed in the

amygdala but are largely found in different neuronal subtypes. Expression of Foxp2 and Foxp4 switches on earlier in development than Foxp1 and postnatally Foxp4 expression is severely downregulated in the forebrain, while Foxp1 and Foxp2 expression persists [51]. Foxp3 is not expressed in the brain. It appears that in the normal brain, a precisely controlled and coordinated pattern of expression is orchestrated, and the requirement of these closely related family members in different regions may give clues to the different functions they perform during brain development.

In addition to the contributions of FOXP2 to human language development, FOXP proteins have been shown to play functional roles in diverse processes ranging from organ development to tumorigenesis [44]. Both FOXP1 and FOXP2 have been implicated in cancer progression [53, 54]. Furthermore, studies of mouse models in which these genes are 'knocked out' have implicated Foxp1/2 in lung development and Foxp1/4 in heart development. Foxp1 has shown to be crucial for determining motor neuron identity in the spinal cord [55, 56]. Foxp4 is a key regulator of foregut development [57–59] and also appears to be important for Purkinje cell arborisation and connectivity [52]. FOXP3 is the most divergent member of the family in terms of sequence and function and has been shown to be important for immune system development [60].

FOXP2 Mutations in Other Cases of Language Disorder

Since the original identification of the point mutation in the KE family, further evidence has come forward, supporting the role of *FOXP2* in language disorders. A number of inherited and de novo mutations have been identified that disrupt the *FOXP2* locus in various ways. In some cases these mutations yield a speech and language disorder that matches very closely with the phenotype observed in the KE family. However, other cases display mixed phenotypes that incorporate elements of DVD, ASD, intellectual disability (ID) and other neurodevelopmental disorders. This heterogeneity can usually be attributed to the size of the mutation, as larger disruptions can often disturb multiple neighbouring genes on chromosome 7, in addition to their effect on *FOXP2*.

Point Mutations of *FOXP2*

The first report to identify independent mutations of *FOXP2* focused on a panel of 49 cases of DVD, chosen for their phenotypic similarity to the disorder observed in the KE family [61]. The study screened the entire coding region of *FOXP2* (Fig. 2.1a) and identified three novel heterozygous variants, in different cases of DVD. Two of these changes (Q17L and $Q_{40}\rightarrow_{44}$) were each found in an affected proband but not an affected sibling. As such, although these changes were not found when screening large numbers of control chromosomes, it was unclear if they represented functional mutations or merely rare coding variants [61]. The third variant

was a heterozygous C to T transition in exon 7 of *FOXP2*. This change was predicted to introduce an early stop codon into the FOXP2 protein sequence (R328X), severely truncating the protein such that most of the functional domains including the leucine zipper/zinc finger and DNA-binding domains were predicted to be lost. This variant was not observed in any of 252 control chromosomes tested, but was present in the other affected members of the proband's family. Specifically, the proband's sister had a similar diagnosis of DVD, while his mother suffered from expressive/receptive language difficulties and had shown speech delay in childhood; each carried one copy of the R328X mutation, whereas the phenotypically normal father did not [61]. Functional studies demonstrated that truncation of the encoded FOXP2 protein via introduction of the R328X mutation resulted in severe mis-localisation of the protein to the cytoplasm [43]. This early stop codon also appeared to result in nonsense-mediated decay and/or an unstable protein product, such that little or no protein could be detected [43]. Thus, the mutation found in this pedigree appears to be effectively a 'null' mutation, completely knocking out one copy of FOXP2.

Chromosomal Alterations Affecting the *FOXP2* Locus: Translocations

Many probands have also been identified that carry chromosomal rearrangements such as translocations or deletions involving the *FOXP2* locus. As described above, pivotal to the original identification of *FOXP2* was case CS who carried a balanced translocation of chromosome 7 that interrupted the coding region of the gene between exons 3b and 4 [27]. This proband displayed a phenotype that was highly similar to that observed in the KE family including severe DVD and substantial expressive and receptive language impairment [27].

A balanced translocation was also found in a mother and daughter with a mixed speech and language disorder with features of spastic dysarthria and DVD [62, 63]. The breakpoints of this translocation were located within the FOXP2 gene on chromosome 7 and the RFC3 gene on chr13 [63]. For both these genes, the translocation was predicted to introduce frameshift mutations resulting in early stop codons that would truncate the protein products. In fact the resulting FOXP2 fusion protein was predicted to be very similar to that produced by the R328X mutation and was thus suggested to be non-functional [43, 61, 63]. This study performed a detailed phenotypic comparison with the KE family and observed a similar impairment of speech, consistent with apraxia of speech (CAS/DVD) but also similar expressive and receptive language deficits—particularly in grammar. This provides supporting evidence for the effects of FOXP2 disruption on the normal development of language, in addition to motor impairment related to vocalisation [63].

Surprisingly, a balanced translocation of 7q31 and 10p14 that did not disrupt the FOXP2 coding region was also found in a patient displaying severe speech impairment and moderate mental retardation [64]. The authors suggest that although the coding region of FOXP2 is intact (FOXP2 is located ~500 kb from the breakpoint), the translocation may produce a position effect that changes the expression of FOXP2, but this hypothesis has not been functionally tested [64].

Chromosomal Alterations Affecting the *FOXP2* Locus: Deletions

A range of deletions spanning 7q31 of various sizes and associated phenotypes have been reported. Five patients with hemizygous deletions spanning the *FOXP2* locus (sized from 11 to 15 Mb) and displaying a phenotype that included (but was not limited to) DVD were reported by Feuk and colleagues [65]. In addition to DVD, four of these patients also displayed symptoms of ASD or developmental delays. The additional phenotypic features observed in these patients are likely to be related to the large deletions observed in which multiple genes in addition to *FOXP2* were also lost.

All five of these individuals carried deletions affecting the paternal copy of 7q31 (i.e. on the chromosome inherited from the father of the proband). In addition, this study reported a further seven probands that inherited two copies of chromosome 7 from their mother, instead of a copy from each parent (a phenomenon known as maternal uni-parental disomy of chromosome 7 or matUPD7) and who presented with features of DVD and Silver-Russell Syndrome (SRS). SRS is a developmental disorder characterised by intrauterine and postnatal growth retardation, craniofacial dysmorphy and musculoskeletal abnormalities. It was observed that FOXP2 expression levels were significantly lower in lymphoblast cells from patients with 7q31 deletions or with matUPD7, as compared to cells from unaffected controls. The researchers proposed that the reduced expression of the *FOXP2* gene in individuals with matUPD7 may be due to a 'parent-of-origin effect' [65]. This kind of effect has been observed for other genetic loci where only one copy of the gene is normally expressed (in this case hypothesised to be the paternally derived copy of the gene), and the other copy is normally 'imprinted' or silenced (here hypothesised to be the maternally derived allele). Under this hypothesis, loss of the maternally derived copy should not severely affect gene expression or phenotype, but loss of only the paternally derived version would be expected to produce a severe effect, similar to that observed when one or both copies of the gene are lost.

In keeping with this parent-of-origin theory, a paternally inherited 16 Mb deletion of 7q31 spanning the *FOXP2* locus was later identified in a proband with a severe expressive and receptive communication disorder including DVD, dysmorphic features and mild developmental delay [66]. In addition, this proband carried a separate inversion of 7q11, although this inversion did not interrupt the coding region of any genes.

However, subsequent reports have identified that deletions of maternal *FOXP2* also cause speech and language problems and thus call into question the parent-of-origin hypothesis. One proband was reported with DVD, expressive/receptive language disorder, language delay, dysmorphic features and moderate mental retardation, due to a maternally inherited 9.1 Mb deletion spanning 7q31.1–7q31.31 [67]. Moreover, members of two additional pedigrees were reported to carry 8.3 and 6.5 Mb deletions of 7q31 [68]. Family members carrying the deletion demonstrated speech problems in addition to a range of other defects, including developmental delay, some autistic features and dysmorphic features. Importantly, one of these families included independent cases of maternal and paternal transmission within the same pedigree, with no discernable difference in the severity of phenotype,

suggesting no parent-of-origin effect for the deletion, in contrast to the findings of Feuk et al. [65].

The smallest *FOXP2*-spanning deletion identified thus far was found in a pedigree in which a 1.57 Mb deletion was transmitted from mother to son, both of whom were affected with DVD [69]. This submicroscopic deletion encompassed only three genes: *FOXP2*, *MDFIC* and *PPPIR3A*. Dysmorphic features were not observed in the family, and they did not meet the criteria for ASD, although the mother and son were classified with pervasive developmental disorder-not otherwise specified (PDD-NOS). The proband displayed severe DVD, expressive language impairment and motor planning defects, while his mother presented with a more moderate phenotype. Thus, in this pedigree, the maternally inherited deletion (in the proband) produced a very severe phenotype of DVD, in contrast to the suggestion of parent-of-origin effects [65].

Finally, a proband was described showing the first example of mosaicism for a *FOXP2* deletion [70]. This was a large de novo deletion of 14.8 Mb, spanning multiple genes, which was only observed in ~50 % of (blood) cells. Despite this, a strong phenotype was observed, including severe DVD, mild mental retardation and language disorder. Thus, a 7q31 deletion in only ~50 % of cells appears sufficient to produce the severe phenotype usually associated with complete loss of one copy of FOXP2 in all cells. It is worth noting that this deletion involves a number of additional genes that may be affecting the phenotype and that the level of mosaicism in the brain may not reflect the pattern observed in blood cells [70]. At present, little is known regarding patterns of mosaicism in different regions of the brain and how accurately this is represented by the mosaicism observed in blood cells.

In sum, the large number of unrelated individuals in which FOXP2 disruptions are associated with a language-related phenotype provides strong support for the involvement of FOXP2 in language disorder. Although only accounting for a small percentage of cases worldwide, it is likely that as DNA technologies advance, smaller (submicroscopic) deletions, copy number variants (CNVs) and further mutations affecting *FOXP2* expression and/or function will be identified. In addition, more precise phenotype definitions and standardised testing will be valuable in unravelling the different genetic causes of speech and language disorders.

Evolution of FOXP2

Although it is clearly involved in speech and language, a human-specific trait, the *FOXP2* gene, is not exclusive to humans. The gene is found in many vertebrate species throughout the animal kingdom, and ancestral forms of *FOXP2* have also been identified in the genomes of invertebrates. Furthermore, orthologues of the FOXP2 protein found in species such as chimpanzee, mouse or songbird are remarkably similar to the protein produced in humans.

The common ancestors of humans and mice diverged over ~65 million years ago, but the versions of FOXP2 protein observed in these two species only differ by three amino acids (equating to ~99.5 % similarity) [71]. This makes FOXP2 one of the most highly conserved proteins shared by these two species [71, 72]. Interestingly, these three amino acid changes are found in exon 3 (E80D) and exon 7 (T303N and N325S) and thus outside the known functional domains of the protein [71]. The fact that the DNA-binding domain is identical in the different species suggests that mouse and human FOXP2 protein are capable of binding to the same target DNA sequences. Moreover, consistent patterns of expression observed in the mouse and human brain suggest similar functions during brain development in both species [49, 50]. Sequence conservation remains high when making comparisons with more distant species; only 8 amino acid changes are found between the human and zebra finch proteins (99 % similarity) and again none of these changes are located in the DNA-binding domain. It is not until one compares human FOXP2 protein to the corresponding fish orthologue that amino acid conservation drops to ~75 % similarity [73]. Even some invertebrates, such as the fruit fly (*D. melanogaster*), worm (*C. elegans*) or sponge (*A. queenslandica*), have an orthologous ancestral protein. However, unlike vertebrates, where FOXP2 is a member of a subgroup of 4 proteins (FOXP1–4), invertebrates have so far only been found to carry a single FoxP molecule that displays ~62–67 % amino acid similarity with the human FOXP family [74]. Given the high degree of conservation of the FOXP2 protein, it follows that model organisms will be highly beneficial in helping us understand how this gene contributes to neural development and function, particularly at a molecular level.

Remarkably, against this background of little change in the protein over millions of years of evolution, two amino acid substitutions in FOXP2 occurred on the human lineage, after splitting from the chimpanzee lineage, at some point within the last 6 million years. The evolutionary time separating humans from chimpanzees is less than a tenth of that separating human and mouse. Yet, in this short period, two of the three amino acid changes that distinguish the human and mouse orthologues arose in the human FOXP2 protein sequence [71]. This rapid fixation of amino acid substitutions on the human lineage is thought to be due to positive selection and may point to altered functions for FOXP2 in the human brain that are subtly different from that in other closely related species [71]. As noted above, the strict constraints on FOXP2 protein sequence over long periods of evolution argue for important role(s) in brain development across a wide range of species. How can we reconcile this observation with its demonstrated impact on complex spoken language, a human-specific phenotype? Human communication involves coordination of a range of sensorimotor, auditory and cognitive components and it is likely that the capacity for language evolved from existing systems in the brain, rather than as a completely novel system [75]. Thus, FOXP2 may have been involved in directing the development of aspects of the ancestral brain that have been later co-opted to subserve language processing during human evolution. If this is true, we can learn a great deal about the neurological basis of language by studying such systems in animal models.

Mouse Models of FOXP2 Mutations

A number of different mouse models have been generated to investigate functions of FOXP2. These are providing complementary insights into neural mechanisms that are normally mediated by FOXP2 as well as the effects of aetiological mutations that cause human disorder (reviewed by [76]). Current mouse models include animals (a) with a complete loss of the protein, (b) carrying changes that mimic the aetiological mutations implicated in speech/language disorder and (c) engineered with evolutionary substitutions that are specific to the human protein (i.e. a mouse that is partially 'humanised' at this locus).

Groszer and colleagues generated two mouse lines (*Foxp2-S321X* and *Foxp2-R552H*) that carried distinct point mutations in *Foxp2* akin to those found in humans with *FOXP2*-related speech and language disorder [77]. The *Foxp2-S321X* allele introduced an early stop codon that results in a truncated protein product highly similar to that observed for a small pedigree segregating verbal dyspraxia (*FOXP2-R328X*) [61]. *Foxp2-R552H* mimics the aetiological missense mutation originally found in the KE family (*FOXP2-R553H*) [27]. Note that although the amino acid numbering system of the human and mouse proteins is slightly different, the *Foxp2-R552H* change in mouse matches exactly the *FOXP2-R553H* change in humans, yielding an arginine-to-histidine substitution at the same position in the DNA-binding domain. As such, mice carrying the S321X or R552H mutations were assessed for phenotypic abnormalities that might shed light on pathways that go awry in the human disorder. Homozygous mutant mice (carrying two mutant copies) were smaller, showed abnormal motor function (e.g. in tests of 'righting reflex'—the ability of a mouse to regain its footing when laid on its back) and survived ~3–4 weeks postnatally, before dying for unknown reasons [77]. The only gross brain abnormality that could be observed was a disproportionately small cerebellum with reduced foliation, indicative of a delayed maturation of this structure [77].

The heterozygous mice (one mutant copy, one normal copy—as in humans with FOXP2-related speech disorders) appeared to be overtly normal, showing none of the developmental delays or reduced viability observed in the homozygotes [77]. These mice did however display subtle phenotypes that point to abnormalities in Foxp2-related areas of the brain. Despite normal baseline motor abilities, the heterozygous S321X and R552H mice demonstrated significantly impaired motor-skill learning on voluntary running wheel and accelerating rotarod tasks. Furthermore, altered synaptic plasticity was observed in two key areas of Foxp2 expression that are already established to be important for motor-skill learning, the striatum and the cerebellum; in particular, there was a dramatic reduction of long-term depression (LTD) in corticostriatal circuits [77]. More recently, in vivo electrophysiological recordings in awake behaving mice have shown Foxp2-mediated effects on striatal plasticity while these mice are actively acquiring a motor skill [78]. In heterozygous R552H mice, the normally low resting firing rate of medium spiny neurons (MSNs)—thought to be important for normal action selection and movement—was elevated. During learning trials on an accelerated rotarod, MSN firing rate typically

increases in wild-type mice, but by contrast showed negative modulation in Foxp2 heterozygous mutants, in a manner that could not be explained by the increased resting rate. There was also a clear reduction in plasticity of MSN firing during training sessions. Finally, the temporal coordination of striatal input was observed to be different between the wild-type and heterozygous mutant mice [78].

Taken together, these results highlight the importance of Foxp2 for the activity and function of the neural circuits in which it is expressed, particularly those involving neuronal subpopulations of the striatum and cerebellum. Interestingly, Groszer et al. [77] and French et al. [78] observed that heterozygous R552H mice showed greater disruptions of motor learning than their heterozygous S321X counterparts. This is consistent with data from in vitro human studies that observed a potential dominant negative effect for the FOXP2-R553H protein in functional cell-based assays, beyond a simple loss of function [43]. Other studies have uncovered additional subtle differences between R552H and S321X heterozygous mice. Although gross hearing appears normal in both mouse models, sound-evoked auditory brainstem responses in S321X heterozygotes did not differ from wild-type littermates, whereas those from R552H mice showed some small but systematic alterations, suggesting potential roles for Foxp2 in auditory processing and auditory system development [79]. To our knowledge, detailed audiometry has not been described for the KE family, so it is not known if there are subtle alterations in the auditory system of the affected individuals and whether this contributes to the severity of their speech and language disorder.

Given the importance of human FOXP2 for spoken language, a capacity that obviously involves vocal output, several studies have assessed the impact of Foxp2 disruptions on mouse vocalisation. A complete homozygous knockout of Foxp2 yields a lack of ultrasonic isolation calls, the innately specified cries that mouse pups make when they are separated from their mother [80, 81]. However, it has been argued that the absence of isolation calls may be a secondary consequence of the severe general developmental and motoric impairments that these homozygous mice suffer from [77, 82]. Crucially, vocalisations made by heterozygous S321X and R552H mouse pups (in the absence of general developmental delay) are produced with similar frequency to wild-type littermates and have largely normal acoustic properties [77, 82]. The use of innate pup vocalisations of mice as a proxy for human speech and language is problematic at best. Innate mouse pup calls are relatively simple and produced without any requirement for voluntary control or auditory feedback (they begin before the animal is able to hear); these vocalisations are more akin to the crying of a baby than to human speech. Furthermore there is evidence from primate studies that innate and learned vocalisations utilise different neural pathways [82, 83].

Moving beyond models of gene dysfunction, a mouse line was engineered to carry certain human-specific changes in the FOXP2 gene, to explore the functional significance of evolutionary differences between human and chimpanzee FoxP2 proteins. This mouse model (which is sometimes referred to as a ‘humanised’ line) carries two amino acid changes (T303N and N325S), encoded by exon 7, which distinguish the human FOXP2 protein from its chimpanzee counterpart (see also

earlier section on evolution of the gene) [84]. For clarity, we refer to the 'humanised' form as FoxP2 as it does not completely match the human or mouse protein. It should be noted that an N325S change also occurred independently during evolution of carnivores and thus this substitution is not unique to humans [85]. Furthermore, a third amino acid difference between human and mouse (E80D, found in exon 3) was left unchanged in this mouse model, so the encoded protein is only partially humanised, and the regulatory regions that control its expression were also unaltered. Nevertheless, intriguing phenotypic differences could be observed in this mouse model as compared to wild-type littermates. Although a large phenotypic screen observed no gross effects on FoxP2 expression, or on anatomy or physiology for any of the tissues tested, including the brain, the partially humanised mice displayed reduced exploratory behaviour and reduced dopamine levels [84]. Since FoxP2 is not expressed in dopaminergic neurons, the authors hypothesised that this transcription factor may indirectly regulate dopamine levels, possibly via its expression in striatal MSNs, which are major targets of dopaminergic neurons [84]. In these mice, the synaptic plasticity of MSNs was also found to be altered, with significantly increased LTD [84]; this contrasted with prior observations in mice carrying disruptive mutations, where LTD was significantly reduced [77]. Lastly, the human-specific amino acid changes were found to have an effect on the length of dendrites in certain FoxP2-expressing areas of the brain. In the striatum, the thalamus and bipolar cells from deep layers of the cortex, dendrites were longer in the 'humanised' mouse than in wild-type littermates [86, 87]. In other FoxP2-expressing areas, such as Purkinje cells of the cerebellum or pyramidal cells from deep layer cortical regions, no significant effects could be observed [87]. In combination, these data suggest that the human form of the FOXP2 protein contributes to the connectivity and function of corticostriatal circuitry, in ways that are subtly different from the murine version [87].

It is not yet clear whether the predominant mode of action for Foxp2 in the brain is developmental or if there is a continued requirement for the protein in circuits of the mature CNS. Studies of songbirds have provided evidence for the importance of FoxP2 in the postnatal brain. Haesler and colleagues selectively knocked down FoxP2 expression in a key song-related nucleus of the juvenile zebra finch brain. Remarkably, this resulted in inaccurate and incomplete imitation of tutor songs, with significantly lower accuracy per song motif, indicating a generalised lack of copying precision compared to controls [88]. The generation of a conditional knockout mouse in which Foxp2 expression can be selectively disrupted at specific developmental time points (or in particular regions of the brain) will also allow investigation of the continued requirement for Foxp2 in the mouse brain [89].

Molecular Networks Underlying Speech and Language

Although *FOXP2* is the most well-defined and extensively studied gene contributing to human speech and language, the molecular mechanisms underlying language development in the brain are likely to involve complex interactions between large

numbers of genes, potentially acting in a range of different neural circuits, and at varying developmental time points. *FOXP2* has been proposed as a 'molecular window' by which we can gain a better understanding of these networks [76]. Indeed, studies exploiting the role of FOXP2 in regulating the expression of downstream target genes have allowed identification of a large number of genes for further investigation.

FOXP2-Related Molecular Networks

The first high-throughput screens for FOXP2 target genes searched for regions of the genome bound by this protein in human neuronal cell lines and human foetal brain tissue [90, 91]. More than 300 predicted targets of FOXP2 were identified in each study, with highly significant overlap observed between the two reports [90]. When a subset of targets were assayed individually, the effect of FOXP2 binding to these promoter regions could be observed. In cell-based assays, the transcription factor typically acted to reduce the expression of the majority of targets tested, although there were some genes that increased their expression in response to the presence of FOXP2. Thus, it appears that FOXP2 largely acts as a repressor but in a small proportion of cases is able to activate gene expression [90, 91]. Given the large number of FOXP2 target genes, it was possible to get an indication of the types of processes that FOXP2 is involved in by understanding the previously identified functions (also known as 'gene ontology') of these target genes. By looking for functional categories that are significantly overrepresented in the target list, it was hypothesised that FOXP2 regulates pathways including the growth and guidance of axons, signalling pathways important for brain development such as 'Wnt/notch signalling' as well as organ morphogenesis [90, 91].

One of the identified FOXP2 targets, the *uPAR* gene, caught the attention of researchers working on other disorders involving disrupted language [90, 92]. The uPAR protein (also known as PLAUR) forms a complex with the protein encoded by *SRPX2*, a gene mutated in epilepsy of the rolandic speech areas of the brain [92]. SRPX2 mutations may also produce symptoms of DVD and/or perisylvian polymicrogyria—a disorder involving abnormal cortical development associated with motor control deficits, cognitive impairment and in some cases seizures and/or language disorder [92]. Given the shared endophenotypes and neurobiological features of syndromes involving *FOXP2* and *SRPX2* mutations, researchers hypothesised that the molecular pathways might be linked. Indeed, functional cell-based assays demonstrated that the FOXP2 protein can bind to the promoter regions of both *uPAR* and *SRPX2* to downregulate their expression [92]. Interestingly, when these same assays were carried out using a mutant version of FOXP2, carrying the R553H substitution from the KE, there was a loss of repression for both target genes. This led Roll and colleagues to screen *FOXP2* in people with disorders of the speech cortex, similar to those caused by *SRPX2* mutations. A screen of 32 patients identified a heterozygous missense mutation of *FOXP2* (M406T) in a proband displaying focal epilepsy, polymicrogyria of the left rolandic operculum and cognitive and

speech defects [92]. Although this change was also observed in one other family member, with no neurological problems, functional assays demonstrated that the amino acid substitution affected the normal activity of the FOXP2 protein. The M406T change resulted in increased mis-localisation of FOXP2 to the cytoplasm and reduced its ability to regulate the *SRPX2* target gene, while uPAR regulation remained unaffected [92]. These data suggest that the mutation of FOXP2 in this patient contributes to the observed phenotype but that it is not directly causative. Further genomic analysis of this proband may uncover mutations in other genes that contribute to the penetrance of the disorder.

Mouse models of Foxp2 have facilitated more in-depth investigation of the crucial molecular networks, using methods that would be difficult to apply to human cases of speech/language disorder. Foxp2 targets in the embryonic mouse brain were inferred from experiments assessing promoter binding across the entire genome. These efforts were coupled with whole genome expression analysis in the developing striatum, a region of high Foxp2 expression that has been implicated in speech and language-related networks in humans and which shows altered function in people with language disorders [93]. The data from this study implicated Foxp2 in a range of developmental processes including cell migration, G-protein-coupled receptor signalling, dopamine signalling, neuron projection morphogenesis and, as before, wnt signalling and axon guidance [90, 93].

Neurite outgrowth and axon guidance are functional categories that reflect the ability of neurons to connect to each other by developing and directing the growth of cellular projections (known as axons and dendrites). These categories were consistently observed across multiple independent FOXP2 studies [84, 90, 93] prompting researchers to investigate this pathway in more detail. A number of putative targets of Foxp2 that were known to be involved in neurite outgrowth were shown to be regulated by Foxp2 in vivo in the developing mouse brain and/or in neuron-like cells in vitro [93]. Furthermore in cultured primary neurons taken from the developing mouse striatum, the loss of functional Foxp2 significantly affected the growth of neurites [93]. Cells expressing normal Foxp2 showed significantly longer neurites with more branch points than the cells expressing mutated Foxp2, suggesting that in the developing brain, Foxp2 may contribute to setting up neural networks in language-related areas of the brain by affecting their connectivity [93].

Another interesting finding from this study was that Foxp2 could regulate the expression of microRNA (miRNA) molecules [93]. miRNAs are short (~22 nt) non-coding RNA molecules that mediate post-translational regulation of gene expression [94]. Mature miRNAs recognise target mRNA molecules via complementary base pairing with a target site in the 3′-UTR of genes and this process generally results in inhibition of translation and/or degradation of mRNA [94]. MicroRNAs such as mir-137, mir-9 and mir-216 that have previously been implicated in brain development and neuronal differentiation were shown to be directly regulated in the developing mouse brain, suggesting that Foxp2 may fine tune gene expression during brain development via the control of miRNA levels [93].

New Candidate Disease Genes: CNTNAP2

A striking proof of principle of the utility of FOXP2 as a molecular window into wider networks of genes involved in language development and disorder came with the identification of *CNTNAP2* as a directly regulated FOXP2 target [95]. In early, low-throughput studies of FOXP2 transcription factor activity, a site within the first intron of the *CNTNAP2* locus was identified as being bound by the FOXP2 protein in human neuronal models [95]. CNTNAP2 expression was also significantly repressed by FOXP2 [95]. *CNTNAP2* is a large gene that encodes Caspr2, a member of the neurexin superfamily. This protein is localised to the axon initial segment (AIS) and juxtaparanodal regions of myelinated nerve fibres and is involved in regulating the clustering of potassium channels in this region [96, 97]. Given that CNTNAP2 had previously been implicated in language-related disorders, such as ASD, cortical dysplasia and focal epilepsy (CDFE) with language regression and Tourette's syndrome [98–101], its function in neuronal recognition and adhesion [101, 102] and its enriched expression in language-related circuitry [103], this presented an excellent candidate gene for language development and disorder [95]. Analysis of SLI families with quantitative measures of SLI endophenotypes demonstrated significant association between 'non-word repetition' and a cluster of genetic markers (single nucleotide polymorphisms, or SNPs) towards the end of the *CNTNAP2* coding region (between exons 13–15) [95]. Some of these same SNP alleles had also previously shown association with a different language-related measure, 'age at first word', in a cohort of autistic children [100]. Since any individuals displaying features of ASD were excluded from the SLI cohort, these findings suggests that similar susceptibility factors at the *CNTNAP2* locus may influence language-related endophenotypes of these different disorders [95]. This work demonstrated that knowledge of a rare Mendelian disorder (speech/language disorder caused by high penetrance *FOXP2* mutations) could inform the genetic basis of more complex language phenotypes (such as SLI or ASD) to highlight shared neurogenetic pathways (FOXP2-CNTNAP2) between clinically distinct syndromes.

New Candidate Disease Genes: FOXP1

FOXP1 was initially considered a good candidate gene underlying language pathways given that it is the most closely related gene to *FOXP2* in the genome. The two protein products have very high amino acid similarity and show conserved and overlapping patterns of expression in regions of the brain, such as the striatum, thalamus and developing cortical plate [49]. Furthermore, the FOXP1 and FOXP2 proteins have been shown to interact to form heterodimers and cooperatively regulate target gene expression [48, 57]. Studies in songbird models have also pointed to a functional role for FOXP1 in the brain as the songbird orthologue of human FOXP1

(FoxP1) shows sexually dimorphic expression in neural structures involved in song learning and production [104].

The first study to screen language-related disorders for potential *FOXP1* mutations sequenced the coding region of this gene in a panel of 49 verbal dyspraxia probands [105]. This was the same panel in which *FOXP2* mutations had previously been identified [61] but no pathological coding changes of *FOXP1* were detected [105]. However, following this, several cases of people with deletions or mutations of the *FOXP1* locus have been reported, associated with complex neurodevelopmental disorders involving multiple symptoms, which often include severe disruptions of speech and language.

A single child was identified with hypertonia and contractures of the hands and feet, blepharophimosis, intermittent muscle spasms and speech delay. This child carried a de novo deletion of 3p14.1 that encompassed four genes: *FOXP1*, *EIF4E3*, *PROK2* and *GPR27* [106]. Subsequently, another patient was identified with a deletion of 3p14.1, but in this case the deleted region only spanned the coding region of a single gene, *FOXP1* [107]. Despite only directly affecting a single gene locus, the patient again showed a complex phenotype that included gross motor delay, Chiari I malformation, epileptiform discharges and limited verbal output [107]. Given that FOXP1 is known to play a key role in motor neuron development and connectivity [55, 56, 108], it might be expected that FOXP1 mutations would yield abnormal motor development and related phenotypes. The challenge lies in determining if the observed speech problems in these patients are due to aberrant development of specific speech-related pathways in the central nervous system or simply a consequence of global motor defects.

Two studies searched for alterations to the *FOXP1* locus in patients with Intellectual Disability (ID) and ASD or speech delay [109, 110]. Screening of 80 ASD and 30 ID probands identified a single patient with a de novo deletion encompassing only the *FOXP1* coding region [109]. A second patient was identified with a de novo nonsense mutation altering the FOXP1 protein to produce a shorter protein that lacked part of the DNA-binding domain [109]. The mutated protein, FOXP1 R525X, was no longer able to regulate gene expression [109]. Both patients displayed global developmental delays coupled with severe language impairments, but no deficits in oromotor coordination were observed [109]. While both patients also displayed autistic features, only the patient carrying the nonsense mutation R525X met clinical criteria for an ASD diagnosis [109].

FOXP1 mutations were also found in an independent screen of patients with moderate ID, general developmental delay, reduced expressive and receptive vocabulary and general speech delay [110]. In a large genome-wide screen for CNVs in 1,523 patients, three cases were identified with deletions that only affected the FOXP1 locus. However, it should be noted that a single deletion affecting four genes (*FOXP1*, *EIF3E3*, *PROK2* and *GPR27*) similar to that observed by Pariani et al. was also found when screening a panel of 4,104 control DNA samples. Horn and colleagues also identified 5 FOXP1 missense mutations in patients with ID that were not observed in control panels [110]. Although no functional analysis was

performed to determine the effects of these mutations, it was predicted that the changes might contribute to the observed phenotype.

A recent study that sequenced the exomes (the entire coding region of an individual's genome) of sporadic autism patients demonstrated four parent child trios with potentially causative mutations [111]. One of the trios was of particular interest as the autistic proband carried a de novo mutation of the *FOXP1* gene. This mutation resulted in an early stop codon in the protein sequence that produces a severely truncated protein product that lacks the key functional domains of FOXP1 [111]. However, this proband carried a deleterious mutation in another gene, *CNTNAP2*, that was inherited from his mother (who is not autistic) and was also passed on to the unaffected sibling. Thus, this *CNTNAP2* mutation did not segregate with the disorder and could not, on its own, be considered to be causative. Functional assays investigating the effects of both the FOXP1 and CNTNAP2 mutations gave some intriguing findings. As noted above, *CNTNAP2*/Caspr2 had previously been implicated in autism and had been shown to be regulated by FOXP2—the most closely related protein to FOXP1. O'Roak et al. demonstrated that the presence of normal FOXP1 is able to downregulate the expression of *CNTNAP2*, but that when the patient identified FOXP1 mutant protein was introduced, CNTNAP2 expression levels were no longer repressed; they were in fact massively increased compared to controls [111]. This could potentially represent a 'two-hit disease model' in which the FOXP1 mutation not only has a direct phenotypic effect but also produces further effects by increasing the expression of the deleterious form of CNTNAP2 [111].

Thus, the potential contributions of *FOXP1* to speech and language functions make a more complex story than that seen for *FOXP2*. It appears that *FOXP1* disruptions are not a major or specific cause of language disorder but that rare mutations of this gene yield susceptibility to complex disorders involving ASDs, ID, generalised developmental and motor delays, often accompanied by speech and language deficiencies. It has been suggested that these data reveal a more global impact on brain development resulting from FOXP1 disruption than is observed from FOXP2 mutations, despite the close homology and overlapping expression patterns [109]. Key to understanding the different effects of FOXP1 and FOXP2 mutations may lie in not only understanding the differences in the pathways they regulate but also understanding the specific neuronal subtypes where these genes are required. Given their close homology it has been suggested that these genes are able to functionally compensate to some degree for each other when genetic disruptions occur. However, this may not be equally true for all target genes or in all types of neurons. Furthermore, there are some regions of the brain where the expression of these genes are mutually exclusive, such as the hippocampus (FOXP1 is present) or the amygdala and cerebellum (FOXP2 is present). Thus, in order to understand the contributions of FOXP1 and FOXP2 to language development, it will be necessary to understand how the functions of these transcription factors overlap and which functions are specific to each family member (as well as when and where they are required).

Perspectives: Language Genetics and Animal Models

This chapter has aimed to present a snapshot of the current knowledge of genetics underlying both normal and disrupted language development. Much of this information was initially obtained from studies of individuals with language-related disorders, which provided the identity of several critical risk genes. However, once these genes or risk factors are identified, new questions arise. What are the normal functions of these genes? How are molecular pathways or neural circuitry in the brain affected when their sequence is altered? In order to begin answering these types of in-depth functional questions, it becomes necessary to move into model systems. Although human model systems are possible, in the form of post-mortem tissue samples or in vitro cell cultures, such models provide only limited options for investigating a trait as complex as language. In addition to the scarcity of human tissue samples, particularly for individuals with well-defined language disorders, the use of post-mortem tissue restricts the range of experimental techniques that can be used to assess gene function and there is no way to manipulate the genetic background of the tissue.

Immortalised human cell lines allow researchers to circumvent some of the above problems since they will grow in the laboratory, can be used for an array of live functional analyses and can be manipulated to alter the sequence and/or expression of particular genes. However, the conclusions that can be drawn from such studies are restricted by the artificial nature of these cells. Typically, these cell lines are derived from tumour biopsies and thus have been altered during the progression of the cancer, often displaying many differences with neurons, including multiple chromosomal abnormalities. They do not represent any particular neuronal subtype, rather they are classified as ‘neuron-like’ cells. Furthermore, the cells are a homogenous population of a single cell type, existing in a monolayer or in suspension. They experience few of the interactions with other cell types or external signals that a normal neuron would have in the brain, essential for directing the complex molecular programmes that distinguish different cellular subpopulations. This is particularly relevant when investigating language as many cell types make up the distributed neural circuits that are thought to underlie human speech and language.

Animal models therefore provide researchers with an excellent tool to study the role of genes in the context of a functioning brain with evolutionary ties to our own. Animal models allow us to manipulate gene expression and observe the effects at multiple levels, from DNA and protein to functional or neurobiological analysis. And although we cannot directly assess language in an animal model, we can look at behaviours that are related to aspects of brain function necessary for language use, such as learning and memory.

Studies of FOXP2 homologues in animal models have provided much of the key information regarding the role of this gene. Studies in mouse models with mutated versions of Foxp2 have identified molecular and functional networks regulated by Foxp2 in the developing brain; demonstrated that Foxp2 is important for neurite outgrowth, synaptic plasticity and motor-skill learning [77, 93]; and highlighted

evolutionary differences by observing the effect of having a humanised version of the protein present in the mouse brain [84, 87]. Chimpanzee cells and tissue have also been used in an attempt to understand how FOXP2 regulatory networks have evolved to contribute to language-related processes [112].

Studies in zebra finch have illustrated the importance of FoxP2 for learned vocalisations, since reduced levels of FoxP2 in the brain affect the ability of the songbirds to correctly imitate and learn song from a tutor [88]. Most recently, sophisticated bioinformatic analyses have also demonstrated a range of FoxP2-related gene networks that appear to be differentially regulated during singing in the zebra finch, suggesting some activity-dependent regulation of these networks [113].

The recent emergence of methods to measure brain activity in living animals will doubtlessly greatly enhance our understanding of the genetic underpinnings of speech and language. Already, studies have been able to measure brain activity in awake, behaving mice and demonstrated neurological differences in how normal brains behave during motor learning compared to Foxp2 mutant brains [114]. The rapidly developing field of optogenetics provides further ways in which the activity and connectivity of neural networks can be probed to determine how genes contribute to processing in language-related structures of the brain. Optogenetics involves introducing light sensitive molecules (opsins) into subsets of cells in the brain. When the cells are exposed to light of specific wavelengths, the neurons that carry these opsins can be either activated or silenced. In this way, it is possible to control and measure the activity through specific neural circuitry, and by combining this with mutant mouse models, it is possible to observe the role of specific genes on the functioning of these circuits [115]. In the future, optogenetics is likely to not only provide insight into the evolution and function of language-related networks in the brain but help us to understand the genetic mechanisms underlying their development.

In summary, to understand the genetic basis of speech and language pathways in the brain, it will be necessary to integrate information gained from clinical studies in human patients with the elegant genetic and behavioural manipulations that can be performed in animal models. Only in this way will it be possible to understand how the faculty for language evolved in the brain and the genetic, molecular and neural mechanisms underlying this most complex human trait.

References

1. Neils J, Aram DM (1986) Family history of children with developmental language disorders. Percept Mot Skills 63(2 Pt 1):655–658
2. Lewis BA, Ekelman BL, Aram DM (1989) A familial study of severe phonological disorders. J Speech Hear Res 32(4):713–724
3. Tallal P, Ross R, Curtiss S (1989) Familial aggregation in specific language impairment. J Speech Hear Disord 54(2):167–173
4. Tomblin JB (1989) Familial concentration of developmental language impairment. J Speech Hear Disord 54(2):287–295

5. Felsenfeld S, Plomin R (1997) Epidemiological and offspring analyses of developmental speech disorders using data from the Colorado Adoption Project. J Speech Lang Hear Res 40(4):778–791
6. Lewis BA, Thompson LA (1992) A study of developmental speech and language disorders in twins. J Speech Hear Res 35(5):1086–1094
7. Bishop DV, North T, Donlan C (1995) Genetic basis of specific language impairment: evidence from a twin study. Dev Med Child Neurol 37(1):56–71
8. Tomblin JB, Buckwalter PR (1998) Heritability of poor language achievement among twins. J Speech Lang Hear Res 41(1):188–199
9. Hayiou-Thomas ME, Dale PS, Plomin R (2012) The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. Dev Sci 15(2):233–249
10. Kovas Y, Hayiou-Thomas ME, Oliver B et al (2005) Genetic influences in different aspects of language development: the etiology of language skills in 4.5-year-old twins. Child Dev 76(3):632–651
11. Hoekstra R, Bartels M, Boomsma D (2007) Longitudinal genetic study of verbal and nonverbal IQ from early childhood to young adulthood. Learn Individ Diff 17(2):97–114
12. Stromswold K (2001) The heritability of language: a review and metaanalysis of twin, adoption, and linkage studies. Language 77(4):647–723
13. Fisher SE, Lai CSL, Monaco AP (2003) Deciphering the genetic basis of speech and language disorders. Annu Rev Neurosci 26:57–80
14. Bartlett CW, Flax JF, Logue MW et al (2002) A major susceptibility locus for specific language impairment is located on 13q21. Am J Hum Genet 71(1):45–55
15. Tomblin JB, Records NL, Buckwalter P et al (1997) Prevalence of specific language impairment in kindergarten children. J Speech Lang Hear Res 40(6):1245–1260
16. The SLI Consortium (2002) A genomewide scan identifies two novel loci involved in specific language impairment. Am J Hum Genet 70(2):384–398
17. The SLI Consortium (SLIC) (2004) Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. Am J Hum Genet 74(6):1225–1238
18. Newbury DF, Winchester L, Addis L et al (2009) CMIP and ATP2C2 modulate phonological short-term memory in language impairment. Am J Hum Genet 85(2):264–272
19. Bartlett CW, Flax JF, Logue MW et al (2004) Examination of potential overlap in autism and language loci on chromosomes 2, 7, and 13 in two independent samples ascertained for specific language impairment. Hum Hered 57(1):10–20
20. Poelmans G, Buitelaar JK, Pauls DL, Franke B (2011) A theoretical molecular network for dyslexia: integrating available genetic findings. Mol Psychiatry 16(4):365–382
21. Scerri TS, Schulte-Körne G (2010) Genetics of developmental dyslexia. Eur Child Adolesc Psychiatry 19(3):179–197
22. Abrahams BS, Geschwind DH (2010) Connecting genes to brain in the autism spectrum disorders. Arch Neurol 67(4):395–399
23. Holt R, Monaco AP (2011) Links between genetics and pathophysiology in the autism spectrum disorders. EMBO Mol Med 3(8):438–450
24. El-Fishawy P, State MW (2010) The genetics of autism: key issues, recent findings, and clinical implications. Psychiatr Clin North Am 33(1):83–105
25. Hurst JA, Baraitser M, Auger E, Graham F, Norell S (1990) An extended family with a dominantly inherited speech disorder. Dev Med Child Neurol 32(4):352–355
26. Fisher SE, Vargha-Khadem F, Watkins KEE et al (1998) Localisation of a gene implicated in a severe speech and language disorder. Nat Genet 18(2):168–170
27. 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
28. Vargha-Khadem F, Watkins K, Alcock K, Fletcher P, Passingham R (1995) Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. Proc Natl Acad Sci U S A 92(3):930–933

29. Gopnik M (1990) Feature-blind grammar and dysphagia. Nature 344(6268):715
30. Watkins KE, Dronkers NF, Vargha-Khadem F (2002) Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. Brain 125(Pt 3):452–464
31. Vargha-Khadem F, Watkins KE, Price CJ et al (1998) Neural basis of an inherited speech and language disorder. Proc Natl Acad Sci U S A 95(21):12695–12700
32. Alcock KJ, Passingham RE, Watkins KE, Vargha-Khadem F (2000) Oral dyspraxia in inherited speech and language impairment and acquired dysphasia. Brain Lang 75(1):17–33
33. Alcock KJ, Wade D, Anslow P, Passingham RE (2000) Pitch and timing abilities in adult left-hemisphere-dysphasic and right-hemisphere-damaged subjects. Brain Lang 75(1):47–65
34. Tallal P, Sainburg RL, Jernigan T (1991) The neuropathology of developmental dysphasia: behavioral, morphological, and physiological evidence for a pervasive temporal processing disorder. Read Writ 3(3–4):363–377
35. Belton E, Salmond CH, Watkins KE, Vargha-Khadem F, Gadian DG (2003) Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia. Hum Brain Mapp 18(3):194–200
36. Liegeois F, Baldeweg T, Connelly A et al (2003) Language fMRI abnormalities associated with FOXP2 gene mutation. Nat Neurosci 6(11):1230–1237
37. Weigel D, Jurgens G, Kuttner F, Seifert E, Jackle H (1989) The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the Drosophila embryo. Cell 57(4):645–658
38. Clark KL, Halay ED, Lai E, Burley SK (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. Nature 364(6436):412–420
39. Overdier DG, Porcella A, Costa RH (1994) The DNA-binding specificity of the hepatocyte nuclear factor 3/forkhead domain is influenced by amino-acid residues adjacent to the recognition helix. Mol Cell Biol 14(4):2755–2766
40. Pierrou S, Hellqvist M, Samuelsson L, Enerback S, Carlsson P (1994) Cloning and characterization of seven human forkhead proteins: binding site specificity and DNA bending. EMBO J 13(20):5002–5012
41. Kaufmann E, Hoch M, Jackle H (1994) The interaction of DNA with the DNA-binding domain encoded by the Drosophila gene fork head. Eur J Biochem 223(2):329–337
42. Wang B, Lin D, Li C, Tucker P (2003) Multiple domains define the expression and regulatory properties of Foxp1 forkhead transcriptional repressors. J Biol Chem 278(27):24259–24268
43. Vernes SC, Nicod J, Elahi FM et al (2006) Functional genetic analysis of mutations implicated in a human speech and language disorder. Hum Mol Genet 15(21):3154–3167
44. Katoh M, Katoh M (2004) Human FOX gene family (Review). Int J Oncol 25(5):1495–1500
45. Kaestner KH, Knochel W, Martinez DE (2000) Unified nomenclature for the winged helix/forkhead transcription factors. Genes Dev 14(2):142–146
46. Shu W, Yang H, Zhang L, Lu MM, Morrisey EE (2001) Characterization of a new subfamily of winged-helix/forkhead (Fox) genes that are expressed in the lung and act as transcriptional repressors. J Biol Chem 276(29):27488–27497
47. Lu MM, Li S, Yang H, Morrisey EE, Min M (2002) Foxp4: a novel member of the Foxp subfamily of winged-helix genes co-expressed with Foxp1 and Foxp2 in pulmonary and gut tissues. Gene Expr Patterns 2(3–4):223–228
48. Li S, Weidenfeld J, Morrisey EE (2004) Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions. Mol Cell Biol 24(2):809–822
49. Ferland RJ, Cherry TJ, Preware PO, Morrisey EE, Walsh CA (2003) Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain. J Comp Neurol 460(2):266–279
50. Lai CSL, Gerrelli D, Monaco AP, Fisher SE, Copp AJ (2003) FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. Brain 126(Pt 11):2455–2462
51. Takahashi K, Liu F-C, Hirokawa K, Takahashi H (2008) Expression of Foxp4 in the developing and adult rat forebrain. J Neurosci Res 86(14):3106–3116

52. Tam WY, Leung CKY, Tong KK, Kwan KM (2011) Foxp4 is essential in maintenance of Purkinje cell dendritic arborization in the mouse cerebellum. Neuroscience 172:562–571
53. Campbell AJ, Lyne L, Brown PJ et al (2010) Aberrant expression of the neuronal transcription factor FOXP2 in neoplastic plasma cells. Br J Haematol 149(2):221–230
54. Banham AH, Beasley N, Campo E et al (2001) The FOXP1 winged helix transcription factor is a novel candidate tumor suppressor gene on chromosome 3p. Cancer Res 61(24):8820–8829
55. Dasen JS, De Camilli A, Wang B, Tucker PW, Jessell TM (2008) Hox repertoires for motor neuron diversity and connectivity gated by a single accessory factor, FoxP1. Cell 134(2):304–316
56. Rousso DL, Gaber ZB, Wellik D, Morrisey EE, Novitch BG (2008) Coordinated actions of the forkhead protein Foxp1 and Hox proteins in the columnar organization of spinal motor neurons. Neuron 59(2):226–240
57. Shu W, Lu MM, Zhang Y et al (2007) Foxp2 and Foxp1 cooperatively regulate lung and esophagus development. Development 134(10):1991–2000
58. Li S, Zhou D, Lu MM, Morrisey EE (2004) Advanced cardiac morphogenesis does not require heart tube fusion. Science 305(5690):1619–1622
59. Wang B, Weidenfeld J, Lu MM et al (2004) Foxp1 regulates cardiac outflow tract, endocardial cushion morphogenesis and myocyte proliferation and maturation. Development 131(18):4477–4487
60. Wildin RS, Ramsdell F, Faravelli F et al (2001) The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3 (Brief Communications). Group 27:20–21
61. MacDermot KD, Bonora E, Sykes N et al (2005) Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. Am J Hum Genet 76(6):1074–1080
62. Shriberg LD, Ballard KJ, Tomblin JB et al (2006) Speech, prosody, and voice characteristics of a mother and daughter with a 7;13 translocation affecting FOXP2. J Speech Lang Hear Res 49(3):500–525
63. Tomblin JB, O'Brien M, Shriberg LD et al (2009) Language features in a mother and daughter of a chromosome 7;13 translocation involving FOXP2. J Speech Lang Hear Res 52(5):1157–1174
64. Kosho T, Sakazume S, Kawame H et al (2007) De-novo balanced translocation between 7q31 and 10p14 in a girl with central precocious puberty, moderate mental retardation, and severe speech impairment. Clin Dysmorphol 17(1):31–34
65. Feuk L, Kalervo A, Lipsanen-Nyman M et al (2006) Absence of a paternally inherited FOXP2 gene in developmental verbal dyspraxia. Am J Hum Genet 79(5):965–972
66. Zeesman S, Nowaczyk MJM, Teshima I et al (2006) Speech and language impairment and oromotor dyspraxia due to deletion of 7q31 that involves FOXP2. Am J Med Genet A 140(5):509–514
67. Lennon PA, Cooper ML, Peiffer DA et al (2007) Deletion of 7q31.1 supports involvement of FOXP2 in language impairment: clinical report and review. Am J Med Genet A 143(8):791–798
68. Zilina O, Reimand T, Zjablovskaja P et al (2012) Maternally and paternally inherited deletion of 7q31 involving the FOXP2 gene in two families. Am J Med Genet A 158A(1):254–256
69. Rice GM, Raca G, Jakielski KJ et al (2012) Phenotype of FOXP2 haploinsufficiency in a mother and son. Am J Med Genet A 158A(1):174–181
70. Palka C, Alfonsi M, Mohn A et al (2012) Mosaic 7q31 deletion involving FOXP2 gene associated with language impairment. Pediatrics 129(1):e183–e188
71. Enard W, Przeworski M, Fisher SE et al (2002) Molecular evolution of FOXP2, a gene involved in speech and language. Nature 418(6900):869–872
72. Zhang J, Webb DM, Podlaha O (2002) Accelerated protein evolution and origins of human-specific features: Foxp2 as an example. Genetics 162(4):1825–1835
73. Itakura T, Chandra A, Yang Z et al (2008) The medaka FoxP2, a homologue of human language gene FOXP2, has a diverged structure and function. J Biochem 143(3):407–416

74. Santos ME, Athanasiadis A, Leitão AB, DuPasquier L, Sucena E (2011) Alternative splicing and gene duplication in the evolution of the FoxP gene subfamily. Mol Biol Evol 28(1):237–247
75. Fisher SE, Marcus GF (2006) The eloquent ape: genes, brains and the evolution of language. Nat Rev Genet 7:9–20
76. Fisher SE, Scharff C (2009) FOXP2 as a molecular window into speech and language. Trends Genet 25(4):166–177
77. Groszer M, Keays DA, Deacon RMJ et al (2008) Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. Curr Biol 18(5):354–362
78. French CA, Jin X, Campbell TG, et al. An aetiological Foxp2 mutation causes aberrant striatal activity and alters plasticity during skill learning. Mol Psychiatry 2011. doi:10.1038/mp:1-9
79. Kurt S, Groszer M, Fisher SE, Ehret G (2009) Modified sound-evoked brainstem potentials in Foxp2 mutant mice. Brain Res 1289:30–36
80. Shu W, Cho JY, Jiang Y et al (2005) Altered ultrasonic vocalization in mice with a disruption in the Foxp2 gene. Proc Natl Acad Sci U S A 102(27):9643–9648
81. Fujita E, Tanabe Y, Shiota A et al (2008) Ultrasonic vocalization impairment of Foxp2 (R552H) knockin mice related to speech-language disorder and abnormality of Purkinje cells. Proc Natl Acad Sci U S A 105(8):3117–3122
82. Gaub S, Groszer M, Fisher SE, Ehret G (2010) The structure of innate vocalizations in Foxp2-deficient mouse pups. Genes Brain Behav 9(4):390–401
83. Jürgens U (2009) The neural control of vocalization in mammals: a review. J Voice 23(1):1–10
84. Enard W, Gehre S, Hammerschmidt K et al (2009) A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. Cell 137(5):961–971
85. Webb DM, Zhang J (2005) FoxP2 in song-learning birds and vocal-learning mammals. J Hered 96(3):212–216
86. Enard W (2011) FOXP2 and the role of cortico-basal ganglia circuits in speech and language evolution. Curr Opin Neurobiol 21(3):415–424
87. Reimers-Kipping S, Hevers W, Pääbo S, Enard W, Paabo S (2011) Humanized Foxp2 specifically affects cortico-basal ganglia circuits. Neuroscience 175:75–84
88. Haesler S, Rochefort C, Georgi B et al (2007) Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus Area X. PLoS Biol 5(12):e321
89. French CA, Groszer M, Preece C et al (2007) Generation of mice with a conditional Foxp2 null allele. Genesis 45(7):440–446
90. Vernes SC, Spiteri E, Nicod J et al (2007) High-throughput analysis of promoter occupancy reveals direct neural targets of FOXP2, a gene mutated in speech and language disorders. Am J Hum Genet 81(6):1232–1250
91. Spiteri E, Konopka G, Coppola G et al (2007) Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. Am J Hum Genet 81(6):1144–1157
92. Roll P, Vernes SC, Bruneau N et al (2010) Molecular networks implicated in speech-related disorders: FOXP2 regulates the SRPX2/uPAR complex. Hum Mol Genet 19(24):4848–4860
93. Vernes SC, Oliver PL, Spiteri E et al (2011) Foxp2 regulates gene networks implicated in neurite outgrowth in the developing brain. PLoS Genet 7(7):e1002145
94. Chen K, Rajewsky N (2007) The evolution of gene regulation by transcription factors and microRNAs. Nat Rev Genet 8(2):93–103
95. Vernes SC, Newbury DF, Abrahams BS et al (2008) A functional genetic link between distinct developmental language disorders. N Eng J Med 359(22):2337–2345
96. Poliak S, Salomon D, Elhanany H et al (2003) Juxtaparanodal clustering of Shaker-like K+channels in myelinated axons depends on Caspr2 and TAG-1. J Cell Biol 162(6):1149–1160

97. Traka M, Goutebroze L, Denisenko N et al (2003) Association of TAG-1 with Caspr2 is essential for the molecular organization of juxtaparanodal regions of myelinated fibers. J Cell Biol 162(6):1161–1172
98. Arking DE, Cutler DJ, Brune CW et al (2008) A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. Am J Hum Genet 82(1):160–164
99. Verkerk AJ, Mathews CA, Joosse M et al (2003) CNTNAP2 is disrupted in a family with Gilles de la Tourette syndrome and obsessive compulsive disorder. Genomics 82(1):1–9
100. Alarcon M, Abrahams BS, Stone JL et al (2008) Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. Am J Hum Genet 82(1): 150–159
101. Strauss KA, Puffenberger EG, Huentelman MJ et al (2006) Recessive symptomatic focal epilepsy and mutant contactin-associated protein-like 2. N Engl J Med 354(13):1370–1377
102. Rasband MN (2004) It's "juxta" potassium channel! J Neurosci Res 76(6):749–757
103. Abrahams BS, Tentler D, Perederiy JV et al (2007) Genome-wide analyses of human perisylvian cerebral cortical patterning. Proc Natl Acad Sci U S A 104(45):17849–17854
104. 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
105. Vernes SC, MacDermot KD, Monaco AP, Fisher SE (2009) Assessing the impact of FOXP1 mutations on developmental verbal dyspraxia. Eur J Hum Genet 17(10):1354–1358
106. Pariani MJ, Spencer A, Graham JM, Rimoin DL (2009) A 785kb deletion of 3p14.1p13, including the FOXP1 gene, associated with speech delay, contractures, hypertonia and blepharophimosis. Eur J Med Genet 52(2–3):123–127
107. Carr CW, Moreno-De-Luca D, Parker C et al (2010) Chiari I malformation, delayed gross motor skills, severe speech delay, and epileptiform discharges in a child with FOXP1 haploinsufficiency. Eur J Hum Genet 18(11):1216–1220
108. Palmesino E, Rousso DL, Kao T-J et al (2010) Foxp1 and lhx1 coordinate motor neuron migration with axon trajectory choice by gating Reelin signalling. PLoS Biol 8(8):e1000446
109. Hamdan FF, Daoud H, Rochefort D et al (2010) De novo mutations in FOXP1 in cases with intellectual disability, autism, and language impairment. Am J Hum Genet 87(5):671–678
110. Horn D, Kapeller J, Rivera-Brugués N et al (2010) Identification of FOXP1 deletions in three unrelated patients with mental retardation and significant speech and language deficits. Hum Mutat 31(11):E1851–E1860
111. O'Roak BJ, Deriziotis P, Lee C et al (2011) Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. Nat Genet 43(6):585–589
112. Konopka G, Bomar JM, Winden K et al (2009) Human-specific transcriptional regulation of CNS development genes by FOXP2. Nature 462(7270):213–217
113. Hilliard AT, Miller JE, Fraley ER, Horvath S, White SA (2012) Molecular microcircuitry underlies functional specification in a basal ganglia circuit dedicated to vocal learning. Neuron 73(3):537–552
114. French CA, Jin X, Campbell TG, Gerfen E, Groszer M, Fisher SE, Costa RM (2012) An aetiological Foxp2 mutation causes aberrant striatal activity and alters plasticity during skill learning. Mol Psychiatry 17(11):1077–1085
115. Rein ML, Deussing JM (2012) The optogenetic (r)evolution. Mol Genet Genomics 287(2):95–109

Part II
Songbird Model of Vocal Learning

Chapter 3
Time Scales of Vocal Learning in Songbirds

Ofer Tchernichovski and Daniel Margoliash

Abstract Songbirds are excellent animal models for studying developmental vocal learning. This developmental process, as well as pathologies that might be associated with it, can be studied in songbirds under tight experimental control: One can control the when and what of vocal learning, record an entire vocal development, and measure neuronal activity in auditory and vocal brain centers while developmental learning takes place. Here we review recent findings about the time scales of vocal changes, and of patterns in neuronal activity, which are associated with vocal learning. We focus on four time scales of vocal change: The first and longest time scale is of the transition between subsong and plastic song. As in speech development, birdsong emerges from a highly variable vocal babbling called subsong. Subsong is followed by plastic song, with the first emergence of distinct syllable types and structured song syntax, which then gradually becomes more elaborated. This transition might be related to large-scale changes in the contributions of basal ganglia pathways to direct control (driver) of song motor output. The second and shortest time scale of vocal learning is the regulation of exploratory variability, which is controlled at time scales of milliseconds, such that different song elements (corresponding to vocal gestures), even just several milliseconds apart, can be learned independently. The third time scale is of learning combinatorial sequences, which occurs at longer time scales and cannot be explained based on the dynamics of trial-and-error learning alone. Across these time scales, it appears that the strongest vocal changes are not induced by singing but by sleep. This fourth time scale occurs over diurnal/nocturnal intervals, involving "offline" changes in neuronal activity, which are thought to be related to consolidation of trial-and-error learning.

O. Tchernichovski, D.V.M, Ph.D. (✉)
Hunter College, 695 Park Avenue, Room 621 HN, New York, NY 10065, USA
e-mail: tchernichovski@gmail.com

D. Margoliash, Ph.D.
Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL, USA

S.A. Helekar (ed.), *Animal Models of Speech and Language Disorders*,
DOI 10.1007/978-1-4614-8400-4_3,

Studying mechanisms of developmental learning across such diverse time scales in songbirds might be useful for understanding developmental and acquired speech pathologies in humans.

Keywords Speech • Language • Development • Timing • Birdsong • Sleep • Vocal motor control • Tutoring

Introduction

Understanding basic mechanisms of developmental vocal learning can be useful for making advances in the treatment of speech and learning pathologies. Some speech pathologies, such as stuttering and apraxia, are primarily developmental disorders. Other language disorders, such as aphasia, are often acquired in adulthood due to brain damage caused by stroke or physical trauma, but here too, the recovery is a developmental process [1].

It is difficult to study mechanisms of speech development directly: Speech development progresses slowly over years, and the neuronal machinery that governs vocal changes is mostly inaccessible to research for both ethical and practical reasons. About 50 years ago, song learning in birds emerged as an animal model for studying basic mechanisms of vocal development [2, 3]. Studying vocal learning is much easier in songbirds: Song learning takes place over weeks instead of years, and with modern techniques it is easy to experimentally control the learning and track it [4]. Furthermore, brain structures involved in vocal learning are much more localized compared to mammalian brains [5, 6]. Consequently, much progress has been made in understanding mechanisms of vocal learning in songbirds, from identifying genes associated with vocal development [7–9], identifying specific brain circuitries that contribute to different aspects of vocal learning [10–12], and in quantifying the behavioral changes that occur in fine time scales [13, 14]. However, little progress has been made so far on the translational front. In this review, we present recent findings about mechanisms and time scales of vocal learning, and then discuss the challenge of bridging this gap, and present evidence suggesting that vocal learning research in songbirds has reached a level of maturity where useful translational studies are becoming possible.

We will first present a brief review of the brain song system [15], focusing on neuronal timing, followed by an updated view of stages in song learning, describing results obtained by continuous recording of an entire vocal development. Then, we will focus on specific time scales of song development, ranging from weeks to milliseconds, attempting to bridge across them. For example, we will describe the effect of sleep on developmental song learning and attempt to relate time scales of circadian changes to the process of matching the developing song to the song model in time scales of milliseconds. Finally, we will raise questions about how looking at those diverse time scales of vocal development could help understand speech pathologies.

The work described here focuses on a few songbird species, with the bulk of the focus on zebra finches. There are roughly 3,500 species of oscine passerine birds [16] and 1,000 suboscine passerine birds [17] and a vast variety of patterns of adult song. This represents a rich diversity of naturally occurring neuronal and behavioral dynamics at all scales of times, the seed for future comparative analyses [18–20]. Finally, towards our goals of relating birdsong and speech, it is valuable to highlight commonalities between the avian and mammalian telencephalons, and the song system and speech and language systems. Many scientists not engaged in comparative studies may not be aware that the old ideas of avian forebrain organization were wrong, and have been overturned even to the extent that a new terminology for avian forebrain has been adopted. This seismic shift in understanding arises from a revolution in identifying homologous relations between mammalian neocortex with what is now recognized as avian cortex. These data include molecular, neurochemical, hodological (connectional), functional, and behavioral observations. We recommend that readers unaware of these developments consult references [21–26].

Timing in the Song System

Song is a whole-animal behavior that involves interaction with and coordination between many systems, including respiration and posture, sleep-related plasticity, and social interactions mediated in part by a host of brainstem and basal forebrain modulatory systems, moment-to-moment feedback arising from ascending auditory and somatosensory systems, an intimate relation with the auditory system in relation to song memories, and more. Thus there is no definition of a "song system" that includes all relevant nuclei associated with any one behavior, and the term should be understood to be colloquial. It traditionally refers to the two principal cortical pathways, a motor pathway and a corticobasal ganglia pathway, and the immediate thalamic, midbrain, and brainstem nuclei associated with the cortical pathways.

There are recent books [27] and numerous recent reviews of the song system, general reviews that attempt to be comprehensive [28, 29] and those that emphasize specific aspects such as theories of learning, molecular mechanisms [30], system developmental mechanisms, the role of the basal ganglia in song learning [31], and the role of sleep in song learning [32]. Here we briefly review the sources of timing information in the song system at multiple time scales.

A traditional view of the song system is a top-down view (Fig. 3.1), which emphasizes the idea that the cortical pathways regulate singing behavior from moment to moment. The motor pathway includes the two cortical nuclei (HVC and RA) necessary for singing (as judged by lesion and electrical stimulation studies): pre-HVC forebrain nuclei that are sources of auditory input or otherwise participate in regulating song production and brainstem nuclei involved in controlling syringeal and respiratory muscles. Neurons in the motor pathway, especially in HVC and RA, show exquisite precision in timing of activity from song bout to bout, in relation to features of song structure. Activity in HVC is sensitive to larger time scales than is

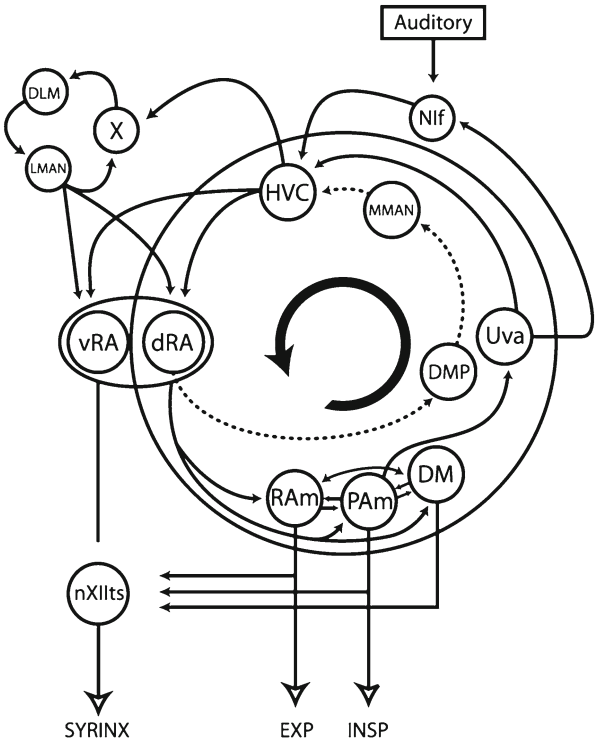


Fig. 3.1 *A bottom-up view of song system organization*. The importance of interaction with respiration during singing and the concomitant role of brainstem respiratory nuclei is emphasized in this figure. The respiratory nucleus PAm, driving inspiration, sends feedback to the thalamic nucleus Uva. Along with feedback from the dorsal RA to the thalamic nucleus DMP, this gives rise to the concept that the song system operates with information flow from brainstem to the forebrain as well as the forebrain to the brainstem, hence, the large circular pathway (*middle*). *Arrows* indicate the direction of connectivity between structures. *DLM* dorsolateral nucleus of the medial thalamus, *DMP* dorsomedial nucleus of the posterior thalamus, *DM* dorsal medial nucleus, *dRA* dorsal subdivision of the robust nucleus of the arcopallium, *EXP* expiratory output, *HVC* (acronym is the proper name), *INSP* inspiratory output, *lMAN* lateral magnocellular nucleus of the anterior nidopallium, *mMAN* medial magnocellular nucleus of the anterior nidopallium, *NIf* interfacial nucleus of the nidopallium, *nXIIts* nucleus XII, tracheosyringeal part *PAm* para-ambiguus, *RAm* retroambiguus, *Uva* nucleus uvaeformis, *vRA* ventral subdivision of the robust nucleus of the arcopallium, *X* Area X. *Adapted from Ashmore RC*, *Wild JM*, *Schmidt MF*. *Brainstem and forebrain contributions to the generation of learned motor behaviors for song*. *J Neurosci 2005*;*25*(37):*8543–54* [55]

the activity in RA, as judged by cases where birds sing syllables that are distinct yet share common notes. The structure of RA activity is related to nearby features of song; the structure of HVC activity can be affected by more distant features of song. At the same time, HVC projection neurons burst at zero, one or a few specific moments in song [33], whereas RA projection neurons burst at multiple moments in song. In this sense, in the motor pathway during singing, integration across song is first

observed in the activity of RA, not HVC projection neurons, so there may not be a simple functional hierarchy between these structures.

The corticobasal ganglia thalamocortical pathway is an alternate pathway from HVC to RA but via the projection of HVC to Area X, to DLM, and to LMAN [34]. This "anterior forebrain pathway" (AFP) is also implicated in regulation of song output timing. At short time scales, at least in estrildid finches such as zebra finches, the regularity of spiking activity in the AFP depends on social context, and this is reflected in time scales of social interaction during singing behavior (e.g., courtship singing vs. undirected singing). A further elaboration of AFP activity is that whereas HVC is the main driver of singing activity in RA during adult singing, early in development it is the output of the AFP (nucleus LMAN) that is the main driver of RA activity during singing [35]. The switch between the two sources of drive is related to a large time scale aspect of song development (see below).

An alternate view of the song system (Fig. 3.1) emphasizes the idea that bottom-up activation and interhemispheric control may also be highly relevant to song control [36]. In this view, activity in brainstem nuclei gives rise to feedback signals that impinge upon and regulate ongoing motor control in the forebrain nuclei. It remains unresolved if feedback signals arise directly from the syringeal muscles, but certainly there are proprioceptive signals arising from respiratory effort associated with singing. This perspective "puts the syrinx back into the song system," that is, it emphasizes that the song system must encode and tightly control some aspects of the dynamics and kinematics of the periphery.

Several theories have been proposed to explain the precise timing of forebrain activity that is associated with singing. One prominent theory addressed the timing of HVC neurons projecting to RA (HVC_{RA}). In zebra finches, the activity of these neurons during singing is remarkable, with each neuron emitting a short (~10 ms) [33] burst of one to a few spikes at some precise moment in song, each time the song passes through that moment, and otherwise remaining silent. (Some HVC_{RA} do not burst at all during singing.) This suggests that song is encoded by sequential temporal activation of small groups of HVC_{RA} that are recruited throughout the song.

In the original study identifying HVC_{RA} during singing [33], no correlation was found between time events of singing behavior (such as onsets or offsets of syllables) and events of HVC_{RA} bursts. Thus it was conjectured that bursts encode a pure time signal. The progression of activity across the population of HVC_{RA} would represent a time base for song, each burst representing a 10 ms "tick" and the population of HVC_{RA} activated as in a synfire chain. By this hypothesis, HVC acts a timekeeper for song production, with the activity of individual HVC_{RA} not directly encoding specific features of motor control. Motor encoding would only emerge at RA, which would translate the timing into a pattern of muscle activation. Although this hypothesis arises from a conjecture, absent any direct experimental support, it is theoretically attractive and has been influential. The hypothesis represents a strong top-down view of song motor control.

A strong challenge to the timer model of HVC activity has recently been raised [37]. This work results in a model that fully incorporates peripheral dynamics into song system. First, under the conditions of sufficient air sac pressure to generate

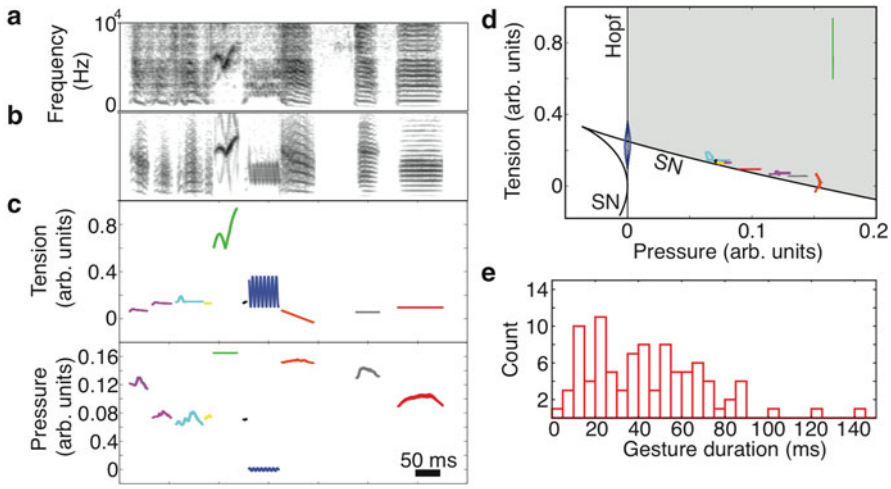


Fig. 3.2 *Gesture model of birdsong production*. (**a**) Spectrograph of a bird's own song and (**b**) spectrograph of a model synthetic song. (**c**) A dynamical systems model of the syrinx and upper vocal tract (see text) identifies distinct movements in subsyringeal air sac pressure and syringeal tension as the basic units of song production. Each of these vocal gestures is color coded, and the sequence of these gestures produced the synthetic song (**b**). It is a hypothesis that the gestures predicted by the model are the ones the bird used to create his own song (**a**). Examining the parameter space of pressure versus tension identifies two types of bifurcations, Hopf and Saddle Node in Limit Cycle (SN) (**d**). Note that most gestures occur in the region of phonation (*grey* region) near the bifurcations. This means that the bird can rapidly change the quality of the vocal output with small changes in pressure or tension. (**e**) The distribution of gestures durations. This distribution is biased towards shorter duration gestures but has a long tail. This feature of the distribution helps to distinguish between two models of encoding of song by HVC neurons (see text). *Adapted from Amador A*, *Perl YS*, *Mindlin GB*, *Margoliash D. Elemental gesture dynamics are encoded by song premotor cortical neurons*. *Nature 2013*;*495*(*7439*):*59–64* [37]

labial vibrations in the syrinx, a model of the dynamics of the syrinx and filtering properties of the upper vocal tract was developed [38, 39]. The modeling effort progressed in tight conjunction with experiments to record measures of peripheral activity, relating those measures to model variables. The result of that extensive effort was a new definition of song production. In this framework, song is described by a nonlinear dynamical system, a set of differential equations. One practical result is that from a microphone recording of a bird's song (Fig. 3.2a), the model permits synthesis of an artificial song (Fig. 3.2b). The model expresses two important conceptual results. First, there are just two time-varying parameters required to control song, subsyringeal air sac pressure and syringeal labial tension. Second, song comprises a sequence of vocal "gestures," with each gesture being a coordinated movement in pressure and tension (Fig. 3.2c). The rapid dynamics of singing is reflected by the fact that for most gestures, pressure and tension variables are maintained in the region of phonation close to lines of bifurcation (Fig. 3.2d).

These peripheral modeling results can also be viewed as hypotheses regarding song system organization. To test these hypotheses, two sets of experiments were

conducted, focused on critically assessing the functional organization of HVC [37]. In one, birds were presented acoustic playback of recordings of the bird's own song (BOS) and the synthesized model BOS (mBOS). The recordings were made in sleeping birds, a condition that gives rise to extreme selectivity for responses to BOS over any other conspecific song [40]. Using this approach it was possible to resolve the value of remaining unfixed static parameters of the model.

At the same time, a remarkable relation was observed between the timing of HVC activity and the timing of gestures. HVC projection neurons burst exclusively at the onsets or offsets of gestures or the moments when pressure or tension achieved a unique maximum value within a gesture. HVC interneurons (that fire tonically throughout song) showed local minima in spike rate functions exclusively at those moments in time. These are experimental results that are inconsistent with the timer model of HVC activity. Given that the primary data for the former model had been collected in singing birds, to further test the gesture model, recordings were also conducted in singing birds. Analysis of those recordings also demonstrated a clear relation between HVC activity and the timing of significant moments in gestures as described above. These results were confirmed by extensive statistical analysis.

This represents a new model of HVC activity, with important implications for how time is represented in the song system. Instead of "ticks" of activity with fixed duration, there are gestures whose duration varies depending on the motor act (Fig. 3.2e). The durations of gestures are short, no longer than a note (elemental unit of a syllable), and the times between significant moments within individual gestures are even shorter. Yet these minute units of behavior, which can be as short as a few ms or as long as >100 ms (in the case of a simple harmonic stack), are represented in output of HVC to the rest of the song system. It is therefore likely that both RA and the AFP receive such information from HVC during singing. Since the HVC_X neurons in HVC that project to Area X (the first nucleus in the AFP) can burst multiple times per song, then perhaps this represents greater integration over song than for HVC_{RA}, which burst only once per song. In terms of theory, the synfire style models need to be modified so that the time base for sequential activity is informed of the time base for motor production. We next discuss how larger units such as syllables, motifs, and song may be represented in the song system, and how a bird learns gestures, notes, syllables, motifs, and song over the time course of development.

Stages in Song Learning

Classical studies [41–46] documented three stages in song development: The first precursor of the song is the subsong. Subsong is a low-amplitude vocal sound, with syllables of broad power spectrum, high-frequency modulation, and variable duration (Fig. 3.3a). A closer look at subsong often reveals a large variety of sounds: high-frequency whistles, buzzing sounds, clicks, etc. But there is no regular pattern of repetition, and instead, the subsong drifts stochastically from one vocal state to another. The second stage is called "plastic song." At that stage, one can already

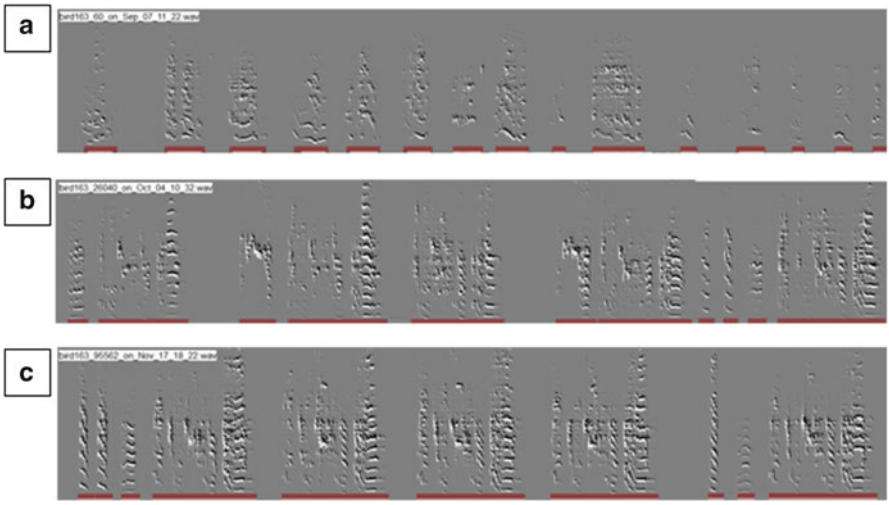


Fig. 3.3 *Song development in a zebra finch*. Panels show spectral derivatives, as in a traditional sonogram but with higher frequency resolution. (**a**), Subsong with rich acoustic structure but no distinct syllable types (graded signal); (**b**), plastic song, which is not fully stereotyped yet, but syllable types can be recognized; (**c**), crystallized song of highly stereotyped syllables repeated in a fixed order, forming song motifs denoted by *red outlines*

identify structured syllables that often resemble those of the model song, although their structure is still variable, and sequential order of syllables is somewhat irregular (Fig. 3.3b). The last stage is called crystallization, when during a relatively short period (e.g., a few days), the song becomes highly stereotyped (Fig. 3.3c).

Physiological observations provide some insight into these largest scales in the transitions in singing patterns during song learning. The roles of HVC and the AFP change over the time course of song development. Whereas HVC provides the main drive to RA in adult birds singing crystallized songs, the AFP provides the main drive to RA during subsong [47]. Thus, subsong can be thought of as a state where song output fully expresses the variability in pattern imposed by the AFP. The transition between subsong and plastic song has recently been investigated [48]. The transition (as assessed by Wiener entropy variance) is not gradual but occurs relatively quickly and is strongly correlated with measures of auditory responsiveness in RA. Lesion experiments (in adult animals) demonstrate that the auditory input to RA arises from HVC. Collectively, these data suggest that birds enter plastic song at a relatively distinct moment in time when HVC gains functional control of RA [35]. This raises the interesting hypothesis that HVC_X are providing feedback to the AFP as it learns how to induce fundamental features of song control ("protogestures") in RA that are prerequisite to sequencing. The same activity in HVC must also be structuring HVC_{RA} activity that soon will start driving RA.

Continuous recording of an entire vocal development made it possible to observe the stages of song learning in detail, leading to a slightly modified view of the stages

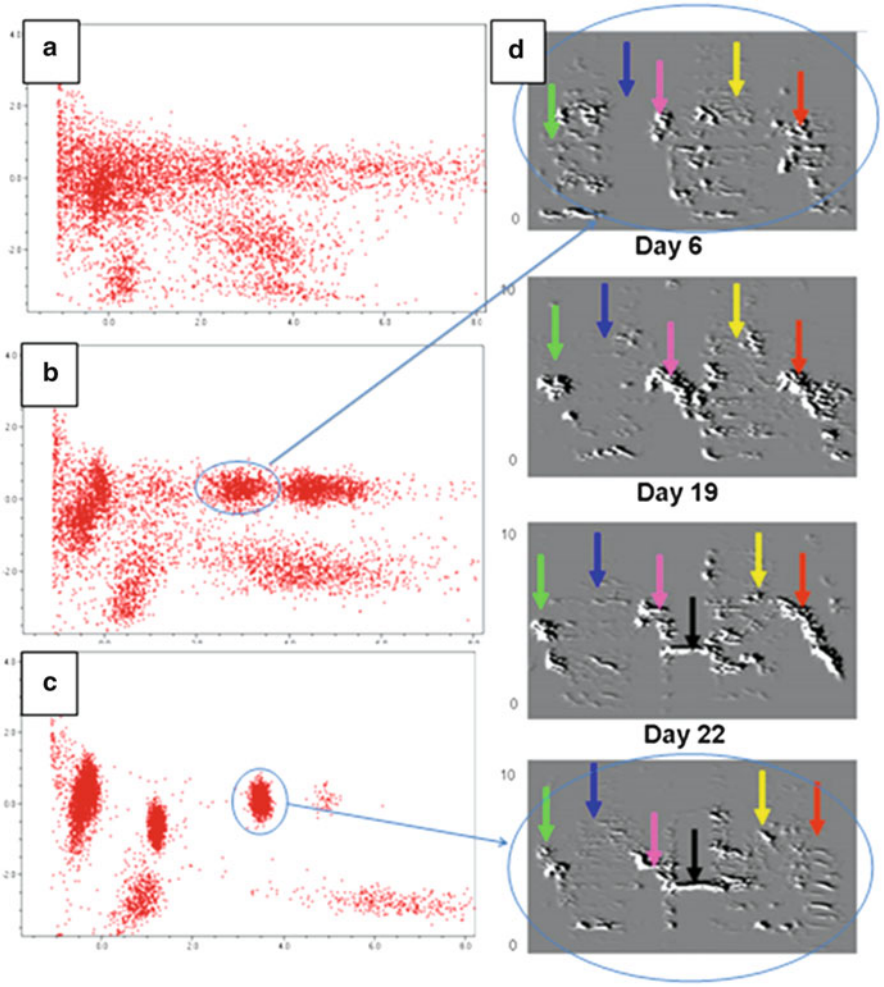


Fig. 3.4 (**a–c**) *Scatterplots of syllable features and analysis of the development of subasyllabic structure*. (**a–c**) Each *dot* represents two features of an occurrence of a syllable, duration versus frequency modulation. The distribution of syllables over one day of singing is presented during subsong (**a**), plastic song (**b**), and crystallized song (**c**) stages. (**d**) Once a cluster (syllable type) is formed, the intra-syllabic structure evolves and differentiated over several days, as distinct mini-clusters emerge within each clusters. Arrows denote the locations of intra-syllabic events that are automatically identified minima of Wiener entropy time courses. *Modified from Ravbar P, Lipkind D, Parra LC, Tchernichovski O. Vocal exploration is locally regulated during song learning. J Neurosci 2012;32(10):3422–32* [13]

of song development: Looking at continuous changes in the distribution of syllable features, e.g., characterizing each syllable by its duration, mean pitch, etc., stages in song development can be observed in nearly real time by looking at scatterplots of syllable features (Fig. 3.4a–b). In Fig. 3.4a, we see an unstructured scatterplot

during the subsong stage, and in Fig. 3.4b we can see how clusters start emerging during the plastic song stage. Those clusters correspond to syllable types. The emergence of those clusters is a structured process: First, we often see only one or two clusters (proto-syllables). Other clusters appear later, either de novo or by differentiation of existing clusters (Fig. 3.4c). Zooming in on one cluster, one can often detect several sub-syllabic structures called notes. Looking at the development of intra-syllabic structure (Fig. 3.4d) reveals a similar process: Initially the cluster is internally unstructured, that is, there are no distinct notes in it. Within a few days, transitions in acoustic state within the syllable become structured and stereotyped, and those can be identified as shown in Fig. 3.4d [13]. The "clusters within the cluster" which correspond to the emergence of notes within the syllable are therefore representing the formation of fine structure on top of an earlier process, where coarse structures (syllable types) are formed. Therefore, in addition to describing a global transition from subsong to plastic song, we like to think about multiple transitions, at smaller time scales, that may occur asynchronously in developmental time and hierarchically in song time.

Multiple Time Scales of Song Development: The Effect of Sleep

As described in the section above, although the transition from unstructured to structured sound may be an abrupt event (which may sometimes unfold within a few hours), thereafter, song structure emerges gradually, and relatively smoothly, first by the formation of coarse structure (distinct clusters indicating syllable types) and then by the formation in intra-syllable structure (Fig. 3.4d). The level of song structure can be quantified by assessing the diversity of song features [4], and indeed, feature diversity increases monotonically over development when looking at daily averages. However, zooming in to, say, hourly levels of song structure, it immediately becomes obvious that the developmental trajectory is strongly non-monotonic [49]. Song structure appears "deteriorated" after night sleep, and structure recovers after about 3 h of intense morning singing. Figure 3.5a shows the trajectory in the structure of one song syllable over time.

As shown, the structure of the syllable (as captured by a feature called Wiener entropy variance) increases slowly over development, with very strong oscillations after night sleep. Interestingly, in addition to the deterioration in song structure, the variability of intra-syllabic structure increases [13]. The two effects appear to be distinct: On one hand, the different acoustic states (notes) within the syllable during plastic song become less distinct from each other, resulting in a syllable that sounds more "flat," but at the same time, each of those states becomes more variable.

These changes in song structure are strongly associated with changes in neuronal activity in the premotor song system. The ongoing activity of RA was assessed by recording during the subjective night in sleeping birds. In young birds raised in isolation of tutor song exposure, RA activity was intermittent, with relatively low

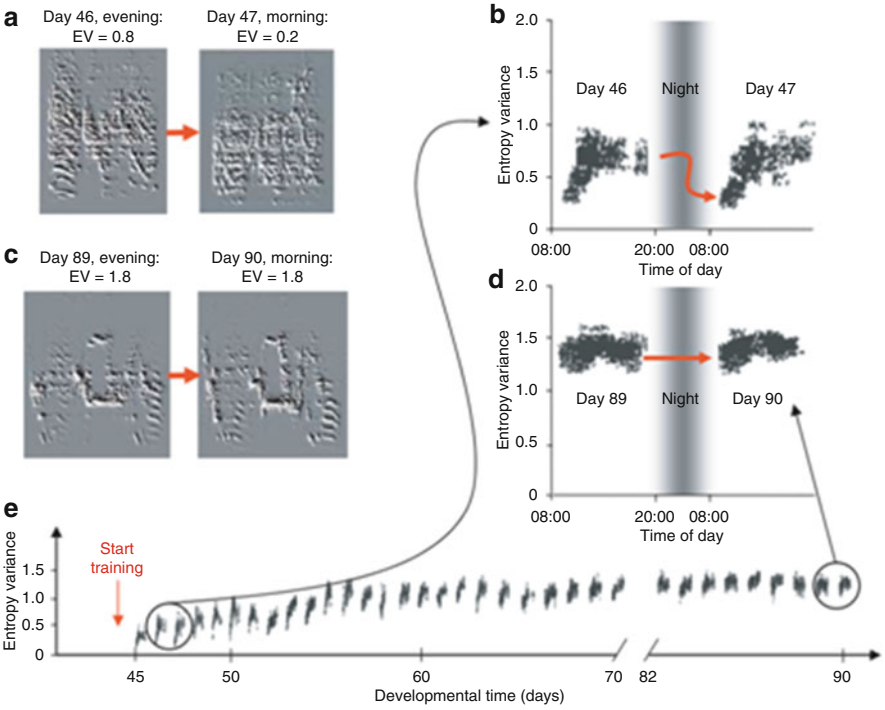


Fig. 3.5 *Recovery of syllable structure after night sleep*. After night sleep, the spectral structure of developing song syllables deteriorates and then recovers within 3 h of intense morning singing. Developmental change in the structure of the same syllable is captured by the variance of Wiener entropy (EV), which captures the diversity of spectral shapes within the syllable. (**a**) Changes in the structure of that syllable during night sleep. Tracking EV values continuously shows a decrease in EV values after the night sleep of day 46 (**a**, **b**) but not after the night sleep of day 89 (**c**, **d**). (**e**) Tracing EV values continuously during development shows oscillations between days 45 and 60. EV values have been smoothed with a running median (period ¼ 40 data points). *Modified from Deregnaucourt S*, *Mitra PP*, *Feher O*, *Pytte C*, *Tchernichovski O*. *How sleep affects the developmental learning of bird song*. *Nature 2005*;*433*(*7027*):*710–6* [49]

average rates and little bursting. On the day such birds were first exposed to a tutor song, there was no discernible change in singing behavior. Yet that night, recordings from RA were as if from a different part of the brain. Neurons were much more active and had much more activity at high frequency (i.e., protobursts). Thus, exposure to the tutor song fundamentally restructured physiological activity in the song system, presumably through a pathway of HVCx on to the AFP, and on to RA. This is the transition into subsong [50].

It is noteworthy that the changes in RA, on the first night after tutor song exposure, preceded the first observed changes in singing patterns associated with night sleep, on the day after the first day of tutor song exposure. This implicates a causal relation between activity in RA during sleep and song learning. A host of experiments in humans and animals implicates sleep and learning, but the physiological

observations are particularly compelling in the song system. After tutor song exposure, RA neurons increase their protobursting but also this high-frequency activity is structured. Even examining simple second order statistics (inter-spike intervals), the distributions in each animal depended on which tutor song the bird was exposed to. The expression of these tutor song-dependent distributions also required that the animal have access to his auditory feedback. Thus, the ongoing RA activity at night is shaped by the two most salient features of learning—the song model and auditory feedback. There is also strong evidence in adult birds that changes in RA activity at night help sculpt the daytime song, contributing to song maintenance, which is a form of learning.

One explanation of the diurnal oscillations from variable performance to more structured performance in young birds is that there is an inherent tension between plasticity and consolidation of structure that the bird has to cope with during vocal learning. The more song structure gets consolidated, the more difficult it might become to add more structure or to undo errors. One can therefore think about those oscillations as a possible mechanism for allowing learning to reconsolidate periodically to finally achieve perfect performance of the song.

The space of parameters or models explored by oscillations during the day and consolidated at night might vary during different phases of song learning. During subsong, only (or principally) the AFP acts on RA, providing both drive and variation of drive onto RA. We suspect that this induces global changes in singing structure during the day, with concomitant global changes in representations of song each night. In contrast during plastic singing, HVC provides the drive and the AFP provides the variation, and we wonder if the interaction between the two is necessary for restructuring the RA network at night. The evidence is that birds express song patterns shaped by learning very early after first exposure to tutor songs, and they don't easily change those patterns [4]. We wonder if syntax and phonology (which in humans is related to semantics) are not independent; perhaps they are constrained by related dynamics of peripheral mechanisms. If true, once these patterns are first established in subsong, this may limit the scope of changes in RA during the period of plastic singing, perhaps focusing changes on those components of song that have undergone the most recent or most substantial changes in the preceding day. This hypothesis could explain, to some extent, why smooth, highly localized changes of limited scope in singing may dominate plastic singing, as we now describe.

Multiple Time Scales of Song Development: Vocal Exploration

The tension between plasticity and consolidation of structure can be seen also at the "localized" changes in vocal sounds from moment to moment during singing: Vocal sounds are highly variable during development across all vocal learners. As with learning to play a musical instrument, one would expect variability to stem, at least in part, from difficulties in gaining control over the instrument. However, recent

studies showed that variability is actively generated and injected into the song structure by AFP [35]. There is direct evidence that this variability promotes vocal learning. What are the natural time scales of this variability? In adulthood, song variability depends on behavioral state. For example, when a male zebra finch is courting a female, he sings a female-directed song, with low variability, while when singing alone (undirected song), variability is significantly higher [51].

Interestingly, during development, the consolidation of structure, as measured by the decrease in variability, changes very locally in song time. To allow an experimental evaluation of localized vocal changes of different type, such as creating a new syllable type, matching the pitch of a particular syllable to a model syllable pitch, swapping the order of syllables, and inserting a new syllable into a string, Dina Lipkind developed an experimental approach that we call "altered-target training" [52]. As shown in Fig. 3.6a, the bird is trained in two stages. First, we train the bird with one song (e.g., AAAA). Once the bird performs a recognizable imitation of that song, we switch the playbacks to another song (e.g., ABAB). In this case, the two songs were designed to separate between the imitations of syllable A and syllable B and also to present the bird with a syllable rearrangement task of inserting a syllable into a string.

As shown in Fig. 3.6, a newly learned syllable type is much more variable than syllables learned earlier [13]. Further, even within a syllable, sub-syllabic units consolidate their structure independently of each other. The rate in which variability decreases is correlated with the localized vocal error. That is, as the performance of any vocal sound approaches the features of the song model that the bird is imitating, its variability decreases.

Therefore, there is evidence for two putative mechanisms for coping with the tension between plasticity and consolidation of structure: One is of periodically oscillating the structure of each vocal sound (a process that requires sleep) and the other is a process of regulating variability locally for each vocal element, restricting it to those parts of the song that require further learning.

Whereas a quantitative analysis has yet to be made, the time scale of the local features of songs that birds learn is seemingly well matched to the time scale of gestures as define by the biophysical model and activity of adult HVC neurons. Are birds learning gestures one at a time? Imagine that learning a gesture requires activation of certain patterns during the day and then reactivation at night. Consolidation of the pattern is reflected in recruitment of a subset of HVC neurons in a local, sequential order. Once this chain develops, it can connect to other chains. A HVC chain would represent a local minimum and would be relatively immune to change upon subsequent reactivation. This would restrict modification to those parts of song and those parts of HVC that require further learning.

To test for constraints on learning to rearrange syllables and presumably connect HVC neuronal chains to each other tail-to-head, Lipkind et al. presented a learning task for birds, which required them to swap syllable order (Fig. 3.7) [53]. As shown, pairwise vocal transitions were acquired, one by one, sparsely over development. A similar effect was then confirmed in human infants during babbling. Therefore, the gestures and neural chain model are supported by behavioral data across species

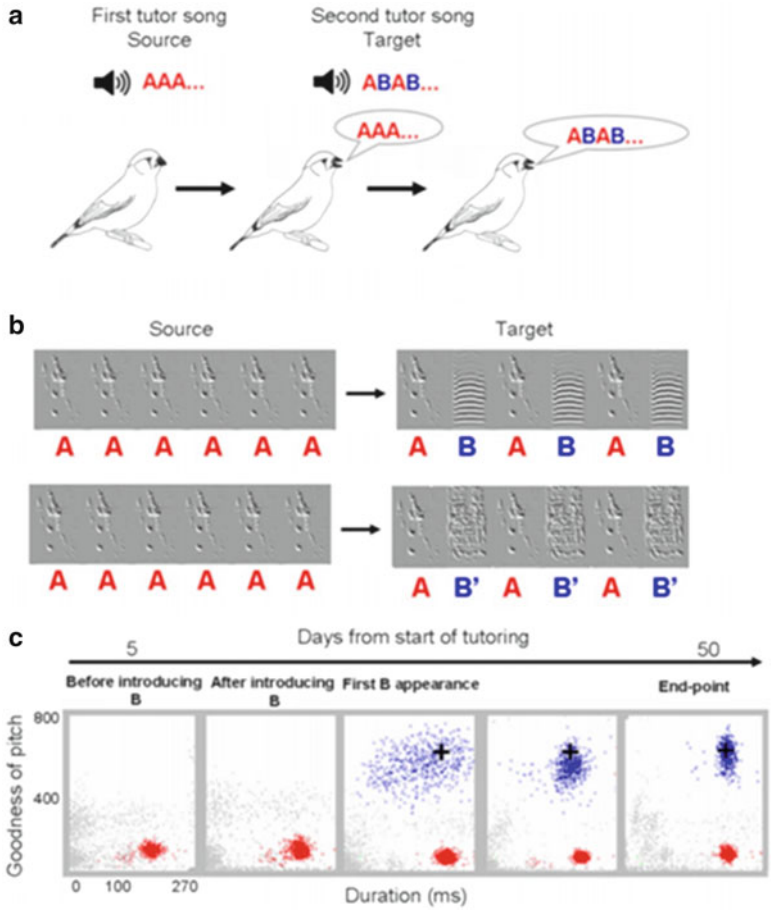


Fig. 3.6 *Vocal exploration is confined to newly added syllables*. (**a**), Birds were first trained with a single-syllable song (AAA…) and then with a two-syllable song (ABAB). (**b**), Spectral derivatives (sonograms) showing the source and target song models. (**c**), Scatterplots of syllable features (goodness of pitch vs. duration and Wiener entropy vs. duration in the same bird). The *red* cluster corresponds to syllable (**a**) and the *blue* cluster to syllable (**b**) (unmodulated version). The + symbol indicates the position of the target syllable (**b**). As shown, the variability of (**a**) cluster does not change significantly after the appearance of (**b**), whereas variability of B drops significantly. *Modified from Ravbar P*, *Lipkind D*, *Parra LC*, *Tchernichovski O*. *Vocal exploration is locally regulated during song learning*. *J Neurosci 2012*; *32*(*10*):*3422–32* [13]

as distant as songbirds and humans. Note, however, that as opposed to the time scales of "birdsong phonology," at the time scale of learning combinatorial abilities, the bird cannot inject combinatorial noise at the early stage of learning. Instead, combinatorial abilities develop slowly, constraining transitions in song syntax.

Such a description across time scales and levels of organization nicely maps the language of behavior to the language of neurophysiology. But the main questions

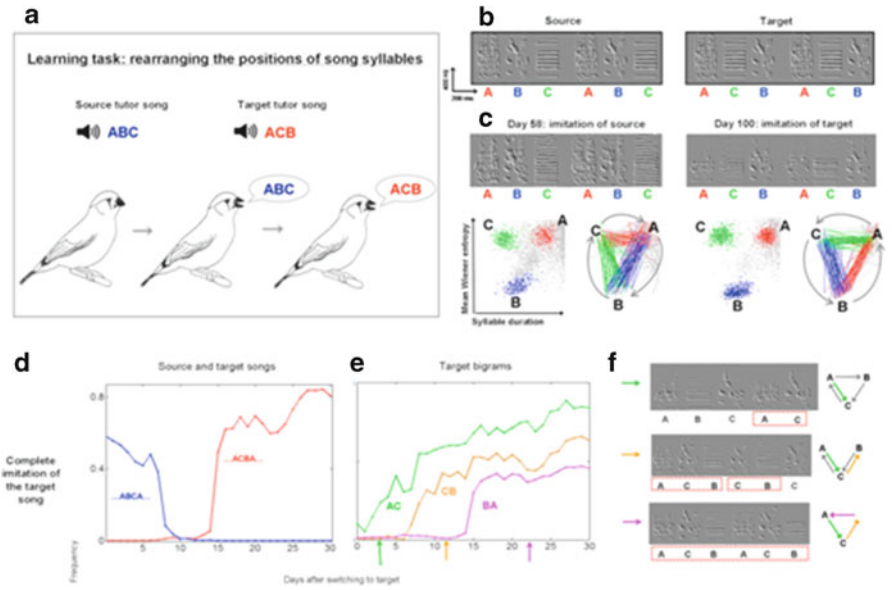


Fig. 3.7 *Stepwise acquisition of syllable transitions during song development*. (**a**) sequential training with two songs ABC → ACB, a task that require changing of syllable order. (**b**, **c**) Song examples (*top*) and scatterplots of syllable features (*bottom*) after source and after learning in one bird. Clusters represent syllable types and lines represent transitions (colors represent transition end syllable). (**d**) Daily frequencies (in one bird) of source song (ABC-ABC) and of target song (ACB-ACB) during development. Note that the source song frequencies decrease to near zero values before the target song appears, indicating intermediate steps. (**e**) Daily frequencies of target-song pairwise transitions AC, CB, and BA. Note that those pairwise transitions are acquired independently and sparsely over development. (**f**) Examples of those steps. *Modified from Lipkind D, Marcus GF, Bemis DK, Sasahara K, Jacoby N, Takahasi M, Suzuki K, Feher O, Ravbar P, Okanoya K, Tchernichovski O. Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. Nature 2013;498(7452):104–8* [53]

remain unanswered. In physiological terms, we seek to learn how the oscillation of activation across day and night consolidates chains, and how those chains are constrained to encode the dynamics of gestures.

From Song Learning to Speech and Language Pathologies

The continuous analysis of vocal learning in songbirds shows that variability has a fundamental role in both guiding and constraining vocal learning and that the regulation of this variability across different time scales is one of the strongest developmental effects that correlate with vocal learning. Let us consider the implications of this to vocal learning in humans. We do not know if and to what extent the mechanisms are conserved, but if they are, would it be possible to figure it out given the

existing data? With the exception of combinatorial abilities, where at least at the behavioral level there are striking behavioral similarities between song development and early speech development, the answer is unfortunately negative [53]. There are no available data of speech development in time scales that can match those of song learning in birds. However, speech development takes place also in adult people who suffer from aphasia following strokes and brain trauma. Very little acoustic analysis of their speech pathologies has been performed, and none of these analytical efforts have looked at variability at any level, phonological, syntactic, or circadian. Our group and others are now performing continuous recordings of therapy sessions in aphasia patients to test if diversity of vocal features and their oscillations might predict the outcome of treatment.

This analysis can be facilitated by a deeper understanding of brain representations of the speech signal. Recently a topography of sensorimotor ventral premotor (vPM) responses was observed in humans, based on the phonetics of the speech signals presented [54]. Speech is a dynamic process, so there must be additional information in the activity in the vPM to represent those dynamics. The gesture model for birdsong can be extended to speech production. This could give insight into possible speech dynamical representations encoded in vPM activity. More generally, the behavioral and physiological observations presented here strongly constrain any model of birdsong learning. It would be valuable to test those constraints on human speech and language acquisition.

References

1. Gitterman MR, Goral M, Obler LK (eds) (2012) Aspects of multilingual aphasia (Edition illustrated). Communication disorders across languages, vol 8. Multilingual Matters, Clevedon. ISBN 1847697542, 9781847697547
2. Marler P, Tamura M (1962) Song “dialects” in three populations of white-crowned sparrows. The Condor 64:368–377
3. Konishi M (1965) The role of auditory feedback in the control of vocalization in the white-crowned sparrow. Z Tierpsychol 22:770–783
4. Tchernichovski O, Mitra PP, Lints T, Nottebohm F (2001) Dynamics of the vocal imitation process: how a zebra finch learns its song. Science 291(5513):2564–2569
5. Nottebohm F, Stokes TM, Leonard CM (1976) Central control of song in the canary Serinus canarius. J Comp Neurol 165(4):457–486
6. 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
7. Hilliard AT, Miller JE, Fraley ER, Horvath S, White SA (2012) Molecular microcircuitry underlies functional specification in a basal ganglia circuit dedicated to vocal learning. Neuron 73(3):537–552
8. 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
9. Warren WC, Clayton DF, Ellegren H, Arnold AP, Hillier LW, Kunstner A et al (2010) The genome of a songbird. Nature 464(7289):757–762
10. Bottjer SW, Miesner EA, Arnold AP (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. Science 224(4651):901–903

11. 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(30):12518–12523
12. Mooney R, Prather JF (2005) The HVC microcircuit: the synaptic basis for interactions between song motor and vocal plasticity pathways. J Neurosci 25(8):1952–1964
13. Ravbar P, Lipkind D, Parra LC, Tchernichovski O (2012) Vocal exploration is locally regulated during song learning. J Neurosci 32(10):3422–3432
14. Tumer EC, Brainard MS (2007) Performance variability enables adaptive plasticity of "crystallized" adult birdsong. Nature 450(7173):1240–1244
15. Nottebohm F, Stokes T, Leonard C (1976) Central control of song in the canary Serinus canarius. J Comp Neurol 165(4):457–486
16. Kroodsma D (2005) The singing life of birds: the art and science of listening to birdsong. Houghton Mifflin Harcourt, Boston
17. Kroodsma DE, Konishi M (1991) A suboscine bird (eastern phoebe, Sayornis phoebe) develops normal song without auditory feedback. Anim Behav 42(3):477–487
18. Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC (2006) Recursive syntactic pattern learning by songbirds. Nature 440(7088):1204–1207
19. Feher O, Wang H, Saar S, Mitra PP, Tchernichovski O (2009) De novo establishment of wild-type song culture in the zebra finch. Nature 459(7246):564–568
20. Fortune ES, Rodriguez C, Li D, Ball GF, Coleman MJ (2011) Neural mechanisms for the coordination of duet singing in wrens. Science 334(6056):666–670
21. Reiner A, Perkel DJ, Bruce LL, Butler AB, Csillag A, Kuenzel W et al (2004) Revised nomenclature for avian telencephalon and some related brainstem nuclei. J Comp Neurol 473(3):377–414
22. 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
23. 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
24. Doupe AJ, Perkel DJ, Reiner A, Stern EA (2005) Birdbrains could teach basal ganglia research a new song. Trends Neurosci 28(7):353–363
25. Luo M, Perkel DJ (1999) Long-range GABAergic projection in a circuit essential for vocal learning. J Comp Neurol 403(1):68–84
26. Mello CV, Clayton DF (1994) Song-induced ZENK gene expression in auditory pathways of songbird brain and its relation to the song control system. J Neurosci 14(11 Pt 1):6652–6666
27. Zeigler H, Marler P (eds) (2008) Neuroscience of birdsong. Cambridge University Press, Cambridge
28. Bolhuis JJ, Gahr M (2006) Neural mechanisms of birdsong memory. Nat Rev Neurosci 7(5):347–357
29. Mooney R, Prather J, Roberts T (2008) Neurophysiology of birdsong learning. In: Eichenbaum H, Byrne JH (eds) Learning and memory: a comprehensive reference. Memory systems, vol 3. Elsevier, Oxford
30. Clayton DF (2000) The genomic action potential. Neurobiol Learn Mem 74(3):185–216
31. Fee MS, Goldberg JH (2011) A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. Neuroscience 198:152–170
32. Margoliash D, Schmidt MF (2010) Sleep, off-line processing, and vocal learning. Brain Lang 115(1):45–58
33. 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
34. Bottjer SW, Altenau B (2010) Parallel pathways for vocal learning in basal ganglia of songbirds. Nat Neurosci 13(2):153–155
35. 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
36. 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
37. Amador A, Perl YS, Mindlin GB, Margoliash D (2013) Elemental gesture dynamics are encoded by song premotor cortical neurons. Nature 495(7439):59–64

38. Mindlin G, Laje R (2005) The physics of birdsong. Springer, New York
39. Perl YS, Arneodo EM, Amador A, Goller F, Mindlin GB (2011) Reconstruction of physiological instructions from Zebra finch song. Phys Rev E Stat Nonlin Soft Matter Phys 84(5 Pt 1): 051909
40. Dave AS, Yu AC, Margoliash D (1998) Behavioral state modulation of auditory activity in a vocal motor system. Science 282(5397):2250–2254
41. Marler P (1970) A comparative approach to vocal learning: song development in white-crowned sparrows. J Comp Physiol Psychol 71:1–25
42. Payne RB (1981) Song learning and social interaction in indigo buntings. Anim Behav 29:688–697
43. Marler P, Peters S (1982) Developmental overproduction and selective attrition: new processes in the epigenesis of birdsong. Dev Psychobiol 15(4):369–378
44. Baptista LF, Petrinovich L (1984) Social interaction, sensitive phrases and the song template hypothesis in the white-crowned sparrow. Anim Behav 32:172–181
45. Nottebohm F (1970 Feb 13) Ontogeny of bird song. Science 167(3920):950–956
46. Nelson DA, Marler P (1994) Selection-based learning in bird song development. Proc Natl Acad Sci U S A 91(22):10498–10501
47. Aronov D, Andalman AS, Fee MS (2008) A specialized forebrain circuit for vocal babbling in the juvenile songbird. Science 320(5876):630–634
48. Lusignan ME, Margoliash D, (2009). Sleep mechanisms in birdsong learning. *Soc Neurosci SFN meeting Abstr*
49. Deregnaucourt S, Mitra PP, Feher O, Pytte C, Tchernichovski O (2005) How sleep affects the developmental learning of bird song. Nature 433(7027):710–716
50. Shank SS, Margoliash D (2009) Sleep and sensorimotor integration during early vocal learning in a songbird. Nature 458(7234):73–77
51. Hessler N, Doupe A (1999) Social context modulates singing-related neural activity in the songbird forebrain. Nat Neurosci 2(3):209–211
52. Lipkind D, Tchernichovski O (2011) Quantification of developmental birdsong learning from the subsyllabic scale to cultural evolution. Proc Natl Acad Sci U S A 108(3):15572–15579
53. Lipkind D, Marcus GF, Bemis DK, Sasahara K, Jacoby N, Takahasi M, Suzuki K, Feher O, Ravbar P, Okanoya K, Tchernichovski O (2013) Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. Nature 498(7452):104–108
54. Bouchard KE, Mesgarani N, Johnson K, Chang EF (2013) Functional organization of human' sensorimotor cortex for speech articulation. Nature 495(7441):327–332
55. Ashmore RC, Wild JM, Schmidt MF (2005) Brainstem and forebrain contributions to the generation of learned motor behaviors for song. J Neurosci 25(37):8543–8554

Chapter 4
The Songbird Auditory System

Sarah M.N. Woolley

Abstract The songbird auditory system offers a unique opportunity to investigate the role of auditory processing in vocal learning and communication. Vocal learning is defined here as imitating the vocalizations of others for communication. While all vertebrates have auditory systems, few use them for learned vocal communication because the ability to learn vocal gestures is rare. Humans, some cetaceans, and three clades of birds (parrots, hummingbirds, and songbirds) are scientifically confirmed vocal learners. Other animals produce and perceive unlearned vocalizations. Currently, the songbird is the only animal model of auditory processing and perception in vocal learners. Extensive studies of songbird vocal behavior, including its dependence on learning, reveal numerous parallels between speech and birdsong. Examples include (1) developmental sensitive periods for vocal learning, (2) learning through imitation of adult models, (3) a dependence of vocal behavior on auditory feedback, (4) the use of unique vocal signals to recognize individuals, and (5) lateralized vocalization processing in sensory and sensorimotor brain regions. Because of these parallels and the evolutionary conservation of auditory circuitry among vertebrates, our understanding of how the songbird auditory system encodes the information in vocal signals and decodes that information into social messages can provide valuable hypotheses for how human auditory processing forms the sensory foundation of speech.

Keywords Speech • Vocal learning • Neural coding • Cortex • Midbrain • Experience

S.M.N. Woolley, Ph.D. (✉)
Department of Psychology, Columbia Univesity, 1190 Amsterdam Avenue,
406 Schermerhorn Hall, New York, NY 10027, USA
e-mail: sw2277@columbia.edu

S.A. Helekar (ed.), *Animal Models of Speech and Language Disorders*,
DOI 10.1007/978-1-4614-8400-4_4,

Background

The songbird auditory system offers a unique opportunity to investigate the role of auditory processing in vocal learning and communication. Vocal learning is defined here as imitating the vocalizations of others for communication. While all vertebrates have auditory systems, few use them for learned vocal communication because the ability to learn vocal gestures is rare. Humans, some cetaceans, and three clades of birds (parrots, hummingbirds, and songbirds) are scientifically confirmed vocal learners [1]. Other animals produce and perceive unlearned vocalizations. Currently, the songbird is the only animal model of auditory processing and perception in vocal learners. Extensive studies of songbird vocal behavior, including its dependence on learning, reveal numerous parallels between speech and birdsong [2]. Examples include (1) developmental sensitive periods for vocal learning [3], (2) learning through imitation of adult models [3, 4], (3) a dependence of vocal behavior on auditory feedback [5, 6], (4) the use of unique vocal signals to recognize individuals [7, 8], and (5) lateralized vocalization processing in sensory and sensorimotor brain regions [9–12]. Because of these parallels and the evolutionary conservation of auditory circuitry among vertebrates [13, 14], our understanding of how the songbird auditory system encodes the information in vocal signals and decodes that information into social messages can provide valuable hypotheses for how human auditory processing forms the sensory foundation of speech.

Studies on the auditory coding of birdsong and the coding properties of songbird auditory regions that parallel human auditory regions have increased dramatically in the last decade. In the zebra finch (*Taeniopygia guttata*) in particular, we now have a general understanding of the spectral and temporal tuning properties of neurons in the auditory midbrain and multiple, higher forebrain (cortical) regions. Comparing tuning properties to the acoustics of song allows us to generate hypotheses about the tuning properties of human auditory neurons as well as how those neurons may encode speech. Furthermore, our understanding of how neural representations of songs transform along the auditory pathway and the role of experience in this process can suggest auditory coding principles that are applicable to speech processing. Here, I describe the songbird auditory system and what we know about auditory coding of learned songs in the different processing stages of the auditory pathway, with a focus on single neuron responses. Where possible, I also describe connections between song and speech processing. Lastly, I suggest hypotheses regarding speech processing in humans, with the goal of using knowledge of auditory processing in the songbird to better understand how the human auditory system subserves speech learning and perception.

Similarities Between Song and Speech

Many birdsongs resemble speech in that they are acoustically complex sequences of stereotyped vocal gestures, lasting from seconds to minutes. Also like speech, song is composed of hierarchically organized acoustic units. Figure 4.1 shows spectrograms

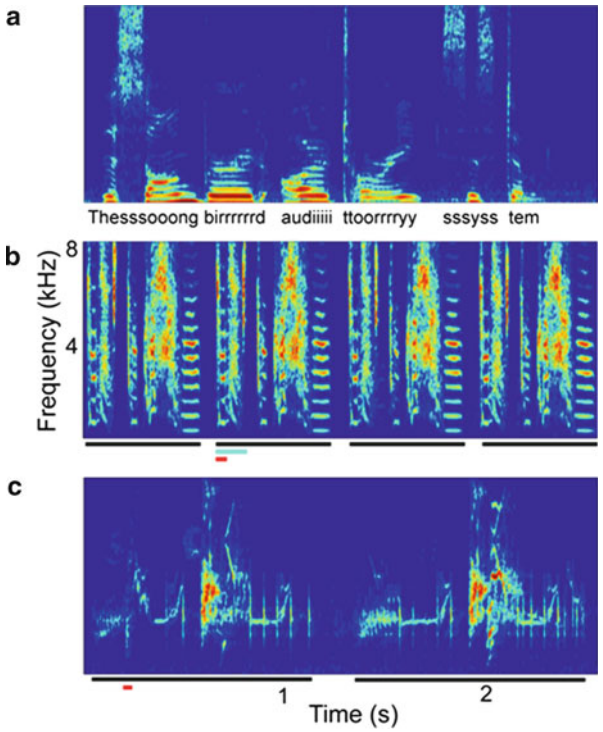


Fig. 4.1 Spectrograms (frequency over time plots) show the frequency and temporal features of speech and song. (**a**) Speech. (**b**) Zebra finch song. (**c**) European starling song. Gray scale indicates intensity. Below the song spectrograms, lines mark example song notes, a syllable, and motifs

(frequency as a function of time, with color indicating intensity) of speech, zebra finch song, and European starling (*Sturnus vulgaris*) song. The hierarchically organized acoustic units of song are indicated with lines below the spectrograms. The smallest individual acoustic elements in songs are "notes." These sounds are equivalent to speech phonemes, the smallest acoustic units that convey meaning in a particular language. Song notes and speech phonemes are grouped together in time into "syllables." In birdsong, a series of syllables repeated in a predictable sequence is a song "phrase" or "motif." A specific combination of syllables or phrases that occurs consistently as a unit is a motif "type," also called a song type. The spectral, temporal, and joint spectrotemporal features of syllables, words, and sentences in speech and syllables and motifs in birdsong are important for communicating information. While speech conveys specific messages and information about the sender's sex, age, identity, and emotional state, song conveys information about species, sex, age, identity, and reproductive fitness. Birdsong lacks speech's limitless capacity to convey different messages by combining words according to grammatical rules. However, the shared capacity to learn and use vocal gestures that convey social

information via acoustic features provides us with an invaluable animal model to examine the neural processing and perception of learned vocalizations.

Songs are species-specific in acoustics and complexity of notes, syllables, and motifs. Some species sing only one song type while others sing hundreds of song types [1]. Because of these differences, some species are more appropriate than others for comparative study of song and speech. For example, based on acoustic features such as spectral structure (e.g., harmonics) and temporal modulations, zebra finch song is similar to speech [15–17]. Figure 4.1a, b show the spectrograms of speech and zebra finch song for comparison. On the other hand, based on higher level structure such as syntax and vocal plasticity, starling song is a good comparison to speech [18, 19]. Starlings learn to produce new phrases in adulthood and sing variable combinations of syllables and motifs, making their songs structurally malleable over time, like speech [20]. Zebra finches and starlings are highly social birds that live well in the laboratory and represent opposite ends of the spectrum in terms of song complexity. Research on the songbird auditory system has converged on these two species because most of the comparative questions regarding auditory coding of learned vocalizations can be addressed using one or the other.

Hearing and the Ear

Hearing is the perceptual outcome of processing in the entire auditory system and the ear performs only the initial stage of sound processing. It may therefore seem paradoxical to discuss hearing and the ear together, saving neural processing for later. But the impact of the ear on basic hearing abilities is so profound that it is useful to describe them together. The ear filters, parses, and amplifies the frequency and temporal components of complex sounds; the transformation of the auditory world performed by the ear determines the sound information that reaches the brain. When considering general hearing abilities such as audible frequencies and frequency-specific sensitivity, peripheral processing is closely related to hearing.

Audible frequencies and frequency-specific sensitivity are measured using behavioral report or large-scale electrophysiological recording of auditory brainstem activity during presentation of pure tones. These measurements yield audibility curves, which plot frequency sensitivity as a function of sound intensity (Fig. 4.2). Songbird and human audibility curves share a similar U shape, with maximum sensitivity at a middle frequency range and sensitivity worsening gradually at divergent frequencies (Fig. 4.2). Maximal sensitivity in this context is defined as the frequencies that can be heard at the lowest intensities, or those at the bottom of the U shape on the audibility curve. Maximal sensitivity in songbirds and humans is similar and well matched to the frequencies in song and speech (see above). Both songbirds and humans hear best between 2 and 5 kHz, unlike other animal models of auditory processing (e.g., mice). Songbird hearing is less sensitive in general and limited in frequency range compared to human hearing, however. Songbirds hear frequencies as high as 10 kHz [21–23], while humans hear frequencies as high as 20 kHz [24].

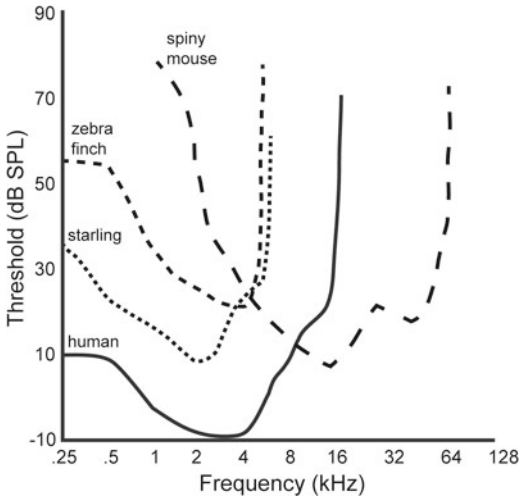


Fig. 4.2 Behavioral audibility curves for human, spiny mouse, and two species of songbird, the zebra finch and the European starling. Songbirds and humans hear best at similar frequencies. Human data are taken from [24]. Mouse data are taken from [173]. Songbird data are taken from [23]

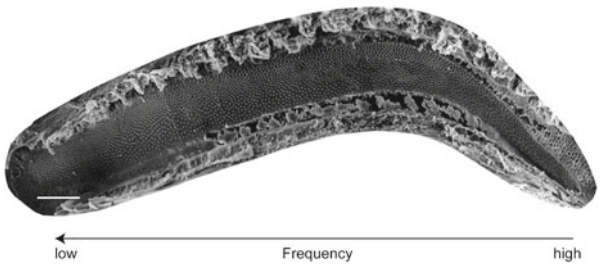


Fig. 4.3 Scanning electron micrograph of a Bengalese finch cochlea. The bone encasing the papilla on the hair cell side and the tectorial membrane have been removed so that the epithelium composed of a mosaic of hair cells is visible. *White dots* on the surfaces of hair cells are stereocilia bundles. *White scale bar* is 100 μm

Figure 4.2 shows example average songbird, human, and mouse audibility curves. Curves are shown for the zebra finch and the European starling because most of the research on the songbird auditory system uses these species.

In birds and mammals, the cochlea performs a spectral decomposition of complex sounds into their component frequency bands; basal hair cells encode high frequencies with short wavelengths and apical hair cells encode low frequencies, with hair cells in the middle of the epithelium encoding the mid-frequencies. Because the auditory epithelium is tonotopically organized, the smaller frequency range of avian hearing corresponds to the shorter length of the cochlea compared to mammals [25]. For example, the sensory epithelium of the Bengalese finch (*Lonchura striata domestica*) cochlea is 2 mm long [26] while that of the mouse is 6 mm long [27]. The avian cochlea (called the basilar papilla) is curvilinear rather than spiraled as it is in the mammal. It is a simpler structure than the mammalian cochlea and it has no tunnel of Corti. Still, bird and mammal cochleae share a basic design and many structures [28]. Figure 4.3 shows the surface structure of the

Bengalese finch auditory epithelium where mechanical vibration of the sheet of receptor cells called hair cells and the overlying tectorial membrane (removed in the figure to visualize the hair cells) results in the release of neurotransmitter from hair cells onto afferent nerve terminals. This is the same basic mechanism for sound transduction as in the mammal cochlea. The lower sensitivity of songbird hearing compared to human hearing (Fig. 4.2) can be partially explained by the structural differences in the avian and mammalian outer ear (birds do not have pinnae), middle ear bones (birds have one rather than three), and the lack of outer hair cells in the avian cochlea. These hair cells amplify peak vibrations of the sensory epithelium in the mammal cochlea [28].

In humans, at least some of the peak hearing sensitivity between 2 and 5 kHz is due to the filtering properties of the ear canal [29]. In the songbird ear, maximal sensitivity in this range may be partly due to the physical properties of the hair cells that encode those frequencies. Hair cell damage and regeneration studies indicate that frequencies above 2 kHz are encoded by the hair cells along the basal half of the songbird cochlea [26, 30, 31]. This region of the avian epithelium has a high relative density of short hair cells, which have larger stereocilia bundles than do tall hair cells and therefore presumably larger numbers of sensory transduction channels [28]. In summary, the similarities in vocal acoustics and hearing sensitivity between songbirds and humans suggest that central auditory neurons in the two systems may exhibit coding similarities since both process complex vocalizations with rich spectrotemporal structure (Fig. 4.1) and hear best at the same frequencies (Fig. 4.2).

Hearing onset in songbirds is not sufficiently studied. It is generally thought that they do not hear until after hatching, in contrast to precocial birds such as chickens and quail [32, 33]. Songbirds hatch in an extremely undeveloped state, another similarity to humans. In both zebra finches and humans, adult-like auditory brainstem responses do not develop until well into the juvenile period [34, 35]. In agreement with this estimate, studies suggest that the impact of auditory experience on song development begins weeks after hatching [36, 37].

Early Auditory Experience and Song Perception

As in humans [38], the perception of vocal communication sounds in songbirds is shaped by early auditory experience of vocalizations. This is the case for both sexes, even in species in which females do not sing. One clear way in which developmental exposure to song influences behavior is the memorization of individual songs heard early in life, including but not restricted to songs that are copied. Both males that learn to sing and females that do not sing remember their fathers' songs as adults [39, 40]. While the father/tutor song memory is obviously important for song development in birds that learn to sing, its significance for non-singing females is less clear. Females may sexually imprint on their fathers' songs to guide their attraction to particular song features or regional dialects during mate choice [40].

Developmental experience of song can also influence general perceptual preferences. Adults of both sexes show preferences for conspecific (same species) songs

over heterospecific (different species) songs that depend at least in part on early experience. Birds that are raised and tutored by conspecific adults are more attracted to conspecific song than to the songs of other species. On the other hand, birds that are raised and tutored by heterospecific adults show either reduced or no preference for conspecific song [41–46]. Such species-specific song preferences may begin with innate biases [21] and be either reinforced or counteracted by early song experience.

The Auditory Nerve

In both birds and mammals, auditory information travels from the cochlea to hindbrain cochlear nuclei via the eighth cranial nerve. Nerve fibers maintain the cochlea's frequency decomposition of complex sounds. Each fiber responds to a limited range of frequencies and is most sensitive to a specific frequency called the best frequency (BF). Response sensitivity decreases to frequencies above and below BF and no sensitivity exists to frequencies that are far from BF [47, 48]. Because of this frequency tuning, the peaks in spectral energy in speech and song may be most important for understanding how the nerve encodes vocal sounds [48, 49]. Examples of spectral energy peaks include the formants in speech vowels (Fig. 4.1a) and the harmonics in zebra finch song syllables (Fig. 4.1b). Because each nerve fiber is frequency tuned, sound information is conveyed to the brain as separate frequency bands, with different groups of fibers representing the information in each band. Based on nerve recordings in response to tones at varying intensities, songbird and mammal auditory nerve fibers differ somewhat in the relationship between frequency tuning and intensity. Frequency tuning (the range of frequencies evoking a response) widens as tone intensity increases in both bird and mammal nerve fibers, though to a much larger extent in mammals than birds [47]. This difference may be important for vocalization coding because it indicates that frequency sensitivity in the songbird auditory nerve is intensity invariant compared to that in the mammal auditory nerve.

The Central Auditory System

Overview

The central auditory systems of birds and mammals are organized similarly up to the higher auditory forebrain (cortex) where significant differences among birds, rodents, cats, and primates are evident. For comparative purposes, I refer here to forebrain regions above the thalamus as cortex. From the hindbrain to the primary cortex, auditory processing regions in birds and mammals have comparable connectivity, molecular markers, and shared functional properties [13, 14, 50–52]. Figure 4.4 shows a circuit diagram of the major auditory pathways in songbirds and

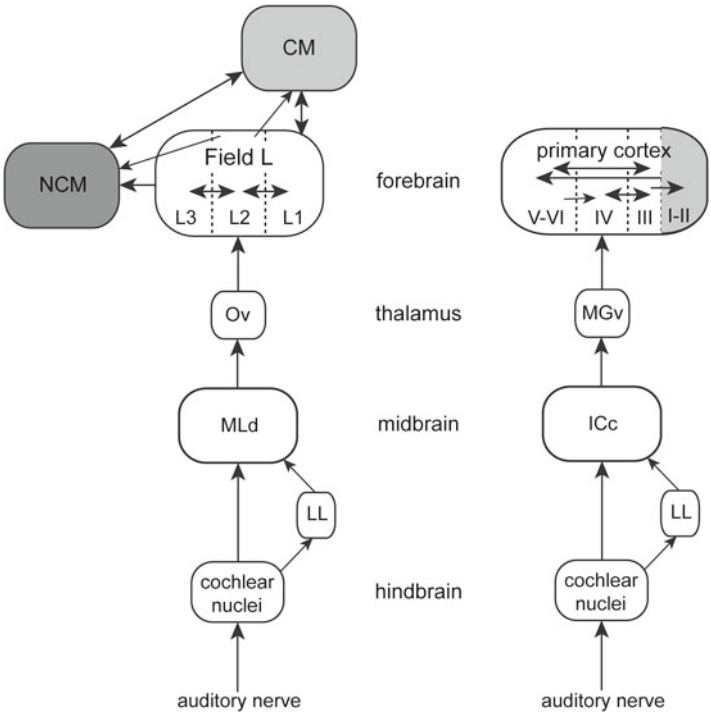


Fig. 4.4 Schematic diagram shows the primary pathways of the songbird auditory (*left*) and mammal auditory (*right*) systems. *Gray-shaded regions* in the songbird and mammal cortex are proposed to be homologous based on connectivity and tonotopy [51]. The *dark gray-shaded region* (NCM) has no known homologous region in mammalian cortex

mammals and cortical homologies as proposed by Karten and colleagues [13, 14]. Auditory information travels from the cochlea to hindbrain cochlear nuclei via the auditory nerve. From the hindbrain, projections lead either directly to the auditory midbrain [53] or to lateral lemniscal nuclei [54], which then project to the midbrain [55]. The avian auditory midbrain is traditionally called the lateral dorsal mesencephalon (MLd) because of its anatomical location, but this nucleus is homologous and functionally similar to the mammalian central nucleus of the inferior colliculus (ICc). Like the ICc, MLd integrates multiple parallel brainstem pathways and provides the primary input to the ascending thalamocortical pathway. It projects to the auditory thalamus (nucleus ovoidalis, Ov), which in turn projects to field L. Field L can be divided into subregions based on connectivity patterns and cytoarchitectural differences [56]. The subregion field L2 is considered homologous to layer 4 of mammalian primary auditory cortex [51, 56–59]. Descending projections from the vocal motor control system innervate MLd, the surrounding region, and the shell around Ov [60]. Interestingly, these projections are found in songbirds but not other birds.

The connections to and among the auditory forebrain regions are complex and reciprocal [58, 60–62]. Subregions L1 and L3 receive input from L2, and these in

turn project to two large auditory regions that surround field L, the caudal medial nidopallium (NCM), and caudal mesopallium (CM) [58, 60]. Field L and CM have the potential to transmit auditory information that influences song motor production through their projections to the motor nucleus HVC and/or its underlying shelf region [62, 63]. Understanding the auditory feedback information carried by neurons in these regions may help us to explain how auditory-motor integration occurs during song learning and maintenance [64]. Below, I describe the neural coding of song and other sounds in the hindbrain, midbrain, and different regions of the cortex as well as discuss how understanding song processing in birds can lead to hypotheses about speech processing in humans.

Hindbrain

The auditory hindbrain is composed of three main nuclei. Nucleus magnocellularis (NM) and nucleus angularis (NA) are innervated by auditory nerve fibers, and nucleus laminaris receives input from NM and the superior olivary nucleus [54, 65]. To date, no studies have examined the responses of songbird cochlear nucleus neurons to songs, but the responses to tones can inform our understanding of their tuning properties [22, 66–68]. To understand how vocal sounds are coded along the entire ascending auditory system, it is useful to describe the frequency tuning and temporal response properties of NM and NA neurons measured from tone-evoked responses. Auditory coding principles that are seen in higher auditory regions begin at this level and demonstrate the similarity between frequency tuning and hearing.

The frequency selectivity of auditory nerve fibers extends to the frequency-selective tuning in cochlear nucleus neurons in both birds and mammals [22, 66–69]. As in nerve fibers, each neuron is most sensitive to a specific BF. Depending on the species, the lowest BFs are around 0.2 kHz (below frequencies found in songs) and the highest BFs are around 9 kHz [22, 66]. In each cochlear nucleus, BF is anatomically mapped, and the range of BFs found in NA maps onto behavioral audiograms [22, 66, 68]. The highest BFs are lower than the highest song frequencies, suggesting that high-frequency information in song is irrelevant for communication.

At very low intensities, only the BF evokes neuronal firing rates that significantly exceed spontaneous firing in cochlear neurons. As intensity increases, the range of excitatory frequencies can increase to include frequencies around BF. Based on responses to tones, it is reasonable to assume that loud vocalizations evoke responses in a larger number of hindbrain neurons than do quiet vocalizations. Another important feature of frequency tuning is that frequencies just above and/or below excitatory frequencies can be inhibitory, meaning that they suppress a neuron's firing to below the spontaneous rate (Fig. 4.4) [67]. This "sideband inhibition" is functionally similar to lateral inhibition in the visual system in that it sharpens tuning. For auditory neurons in general, the frequency dependence of neural excitation and/or inhibition means that broadband vocalizations or those that have energy peaks at multiple points along the frequency axis (Fig. 4.1) can both excite and inhibit a single neuron

[70]. This means that the ability to predict the response of a neuron to broadband song requires a detailed understanding of the extent to which frequencies present in the song excite or suppress that neuron's firing. In contrast, narrowband vocalizations such as whistles present in some birds' songs will excite those neurons tuned to the whistle frequency but may inhibit neurons tuned to adjacent frequencies or have no impact on the neurons tuned to very different frequencies. The tuning properties described above are also observed in mammal cochlear nucleus neurons [71, 72]. Therefore, at the level of the hindbrain, frequency tuning in the songbird resembles that in the mammals, suggesting that understanding "low-level" song coding in the bird brain may be applicable to understanding the encoding of speech in the human hindbrain.

Responses of cochlear nucleus neurons also depend on sound intensity and neurons' inherent temporal response properties. As in mammals, most songbird cochlear nucleus neurons respond more to an excitatory tone as the tone intensity increases. However, after a certain point, increases in intensity either have no effect on firing rate or evoke smaller responses [67]. The intensity coding observed in most cochlear nucleus neurons suggests that at behavioral sound levels, neural responses to vocalizations scale with intensity.

Temporal response patterns evoked by tones differ among anatomically distinct cell types in mammalian cochlear nuclei [69, 73]. The most common response pattern is a phasic burst of spikes at tone onset followed by a tonic, lower firing rate response that declines gradually for the duration of the tone. Bushy cells produce these responses, which are called "primary-like" because they resemble the response patterns of auditory nerve fibers [73, 74]. In contrast, octopus cells respond only to tone onsets. Other neurons in mammalian cochlear nuclei produce "pauser" responses. These responses are characterized by a strong onset response followed by a drop and then a gradual increase in firing rate. Songbird cochlear nucleus neurons also produce these three temporal response patterns [67, 68]. We know nothing about how such physiological cell types map onto anatomical cell types in songbirds, however. In summary, songbird and mammal auditory hindbrains exhibit the same fundamental sound coding properties of frequency selectivity at the level of a single neuron, tonotopic maps at the level of the nucleus, and three main types of temporal response patterns.

Midbrain

As in mammals, the songbird midbrain is a major site of integration for the flow of auditory information along the ascending pathway. Although the boundaries and functions of its subregions are debated [75–77], the midbrain is where multiple parallel brainstem pathways carrying auditory information converge; projections from NA, NL, the superior olivary nucleus, and the lateral lemniscal nuclei all meet in the midbrain [50, 53]. Interestingly, inputs from NA and NL have overlapping terminations in the songbird MLd [53], an aspect of auditory circuitry that is shared

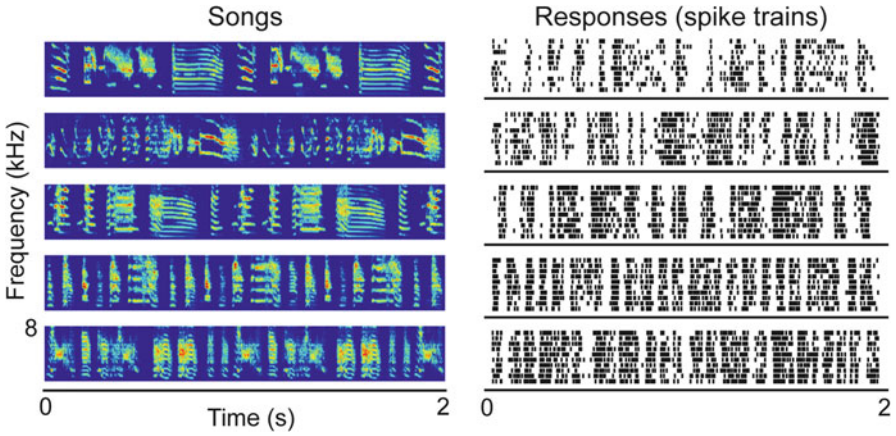


Fig. 4.5 Single neuron responses to songs are temporal patterns of spikes that are consistent within a song and distinctive between the songs of different birds. *Left*, spectrograms of five songs presented to a bird while recording the responses of a single midbrain neuron. *Right*, raster plots showing the spike trains evoked by presentation of the songs. Raster plots show reliable responses to ten presentations of the same song (one large row of ten small rows with ticks showing spike times) and unique responses to different songs (different tick patterns across large rows)

with mammals [78] but not with other birds [79–81]. Neurons projecting out of MLd provide the primary input to the ascending thalamocortical pathway. The medial portions of MLd and the region that surrounds it receive descending input from the cup of the rostral arcopallium (RA), a vocal motor control nucleus [60]. This input may send feedback information about song motor output to the auditory system.

The songbird auditory midbrain is a focus for studying subcortical auditory processing of song. Unlike neurons in some higher processing regions, midbrain neurons respond robustly to a wide variety of sounds including tones [70, 82, 83], noise [17, 70, 84, 85], and song [70, 85–88]. The responses of single midbrain neurons are reliable, meaning that they produce highly similar responses to the same stimulus presented multiple times (Fig. 4.5). This is in contrast to the context-sensitive and habituating (decreasing over time) responses of neurons in some cortical regions. The robust and reliable responses of midbrain neurons to different types of sounds allow the direct comparison of neural responses to song and to other sounds in the same single neurons. Such comparisons reveal complex interactions of excitatory and suppressive acoustic features in the subcortical coding of natural sounds [70].

In parallel with the mammalian IC, midbrain responses to tones show that the majority of neurons have classic V-shaped tuning similar to the tuning in cochlear nuclei, while others have more complex frequency and intensity tuning [70, 82, 83]. Nearly all midbrain neurons are most sensitive to one specific frequency (BF) which is anatomically mapped such that neurons in the dorsal midbrain are most sensitive to low frequencies and ventral neurons are most sensitive to high frequencies [83]. The reliable responses of midbrain neurons can be used to discriminate among the songs of different birds based on the temporal pattern of the response to each song

[86]. The spike trains evoked by one song presented multiple times have highly similar patterns while the spike trains evoked by the unique songs of different birds have dissimilar patterns. Figure 4.5 shows spectrograms of the unique songs of five male zebra finches (left) and raster plots of one midbrain neuron's responses to those songs (right). When the same song is presented ten times, the neuron fires at roughly the same points in the song each time; the neuron's responses are reliable (Fig. 4.5, compare spike trains within a row). The responses to one song have a specific temporal pattern because the neuron responds to short timescale acoustic features that occur at specific points in the song. For example, the spikes shown in Fig. 4.5 align with the timing of the song syllables and are dense when the song contains significant energy at 4 kHz. In contrast, the neuron fires at different points in response to each song because the acoustic feature to which it is responsive occurs at different points in different songs (compare spike trains across rows). At this level of the auditory system, the temporal patterns of spike trains evoked by different songs vary more than does the average number of spikes evoked by different songs [86]. The similarity of spike trains evoked by numerous presentations of the same song and differences in the spike trains evoked by presentation of different songs can be quantified to estimate the "neural discrimination" of songs [86, 89–91]. In the zebra finch midbrain, spike trains of a single neuron can be used to predict which song a bird heard [86]. The same neural discrimination of songs based on spike train temporal patterns is observed in many neurons in field L [91, 92]. This coding principle may provide a basis for the discrimination between songs sung by different males and potentially contribute to individual recognition during social interactions.

When the responses of neurons that produce similar spike trains to the same song are combined, the accuracy of song neural discrimination improves [86]. In addition, the responses of populations of midbrain neurons can be used to reconstruct the spectrograms of songs that evoked those responses [93]. As a result, the combined responses of individual midbrain neurons tuned to different acoustic features in songs can represent the complete song. Similarly, population activity in the human auditory cortex can be used to accurately identify presented speech segments [94].

Spectrotemporal tuning properties largely explain the unique responses of midbrain neurons to song [70, 82, 85–87, 95]. Each single neuron has a specific combination of spectral, temporal, and intensity tuning [83, 84]. These response properties determine the features of songs that each neuron encodes, specifically whether and how much it fires over the time course of the song. Thus, the different responses of each neuron to the same song and the different responses of the same neuron to different songs are both due to a neuron's spectrotemporal tuning properties and to the specific acoustic patterns comprising each song.

The spectrotemporal tuning properties of a midbrain neuron can be estimated from the responses of that neuron to a large set of songs. The acoustic features that drive a neuron are measured by determining exactly when the neuron fires during the presentation of many songs (Fig. 4.5) and averaging the acoustic features that consistently precede spiking (Fig. 4.6). With this approach, the acoustic features that reliably evoke firing are estimated and a calculated spectrotemporal receptive field (STRF) depicts the neuron's tuning during song processing (Fig. 4.6a, b) [96].

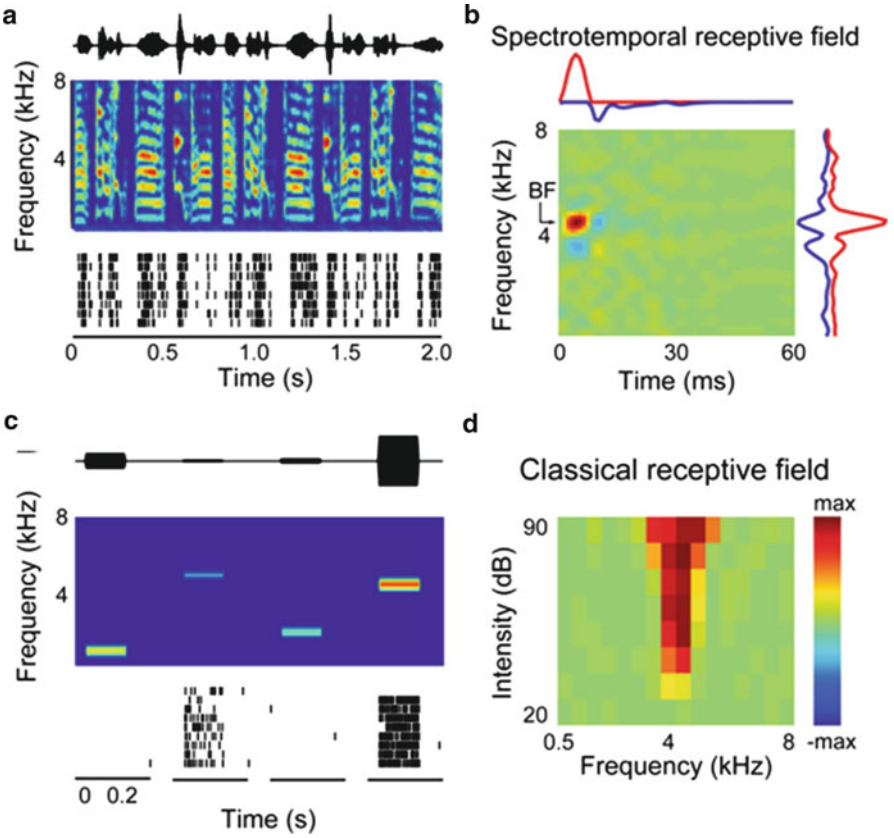


Fig. 4.6 Tuning of a single neuron can be compared between the processing of single tones and songs. (**a**) Amplitude waveform (*top*) spectrogram (*middle*) and spike rasters showing spike times for ten presentations of the song (*bottom*). (**b**) The spectrotemporal receptive field (STRF) for the neuron, calculated from responses to 20 unique zebra finch songs. *Dark* indicates strong power and light indicates weak power. The temporal power profiles for excitation and suppression are above the STRF. The spectral power profiles for excitation and suppression are to the right of the STRF. The STRF shows that the neuron is tuned to brief bursts of sound at ~4.5 kHz and is suppressed by co-occurring sound at ~3.5 kHz. This neuron is driven by a harmonic stack with an energy peak at 4.5 kHz and an energy trough at 3.5 kHz and a song note with energy only at 4.5 kHz. The neuron does not respond to a song note with energy restricted to 3.5 kHz (see (**a**)). (**c**) Amplitude waveform, spectrogram, and raster plots of the neuron's responses to pure tones that vary in frequency and intensity. These stimuli are used to obtain the neuron's tuning curve and classical receptive field. (**d**) Frequency response area plot showing the classical receptive field of the neuron, defined as the tuning measured from responses to single pure tones. The best frequency (BF) is 4.5 kHz and the range of frequencies evoking a response increases slightly as tone intensity increases

The STRF in Fig. 4.6b shows tuning for a neuron that is maximally excited by a brief acoustic feature (e.g., a song note) that contains energy at 4.5 kHz and no energy immediately below that frequency. The neuron will fire ~7 ms after the presentation of that feature. The tuning properties of midbrain neurons are well

characterized by such linear-nonlinear tuning models [95]. While it is straightforward to measure tuning by analyzing responses to simple sounds such as tones (Fig. 4.6c, d), the STRF provides a method for analyzing spectral and temporal tuning of a neuron from its responses to complex, natural sounds such as song and speech [70, 85, 87, 95–97]. Because we ultimately want to understand how the auditory system encodes vocal sounds rather than just tones and/or noise, the use of STRFs is crucial for advancing our understanding of how song is represented by the firing patterns of neurons along multiple stages of the auditory pathway.

Auditory tuning mechanisms that facilitate the coding of vocalizations are evident at the level of the midbrain. This does not surprise those who are familiar with the literature on midbrain coding of vocal signals in mammals [98–100] and frogs [101, 102]. But it does surprise those who think of cortex as the origin of functional specializations for the coding of behaviorally relevant signals. Complex vocalizations like speech and song are characterized by modulations in energy across frequency (spectral modulations; e.g., harmonics) and time (temporal modulations) [15–17]. The spectral, temporal, and combined spectrotemporal modulations in speech are important for intelligibility; removal of low-frequency spectral modulations and midrange temporal modulations significantly lowers the comprehension of speech sentences in human listeners [103]. The dependence of speech perception on spectrotemporal modulation frequencies suggests that auditory neurons are tuned to modulation frequencies that characterize behaviorally salient sounds such as vocalizations. In the songbird midbrain, auditory neurons are tuned for the spectral and temporal modulations that characterize song [17]. When a noise stimulus containing spectral and temporal modulations that are in song and modulations that are not in song is presented, neurons selectively respond to noise segments containing modulations that match those in song. These selective responses to modulations in noise that match song modulations suggest a tuning mechanism that facilitates the coding of acoustic information in song and filters out other acoustic features. Similar tuning for spectrotemporal modulations that match frequency-modulated sweeps in vocalizations has been found in mammal IC neurons [98, 99]. A recent study on human perception of spectrotemporal modulations supports the hypothesis that human auditory neurons are tuned for sounds composed of specific spectrotemporal modulations. Sabin et al. [104] found that training subjects to discriminate modulation depth for a stimulus consisting of one spectral and one temporal modulation frequency led to improved discrimination on that stimulus but the learned improvement did not generalize to stimuli with different spectrotemporal modulation frequencies, even if the spectral modulation frequencies were shared. This provides behavioral evidence that the human auditory system contains neurons that are tuned to specific spectrotemporal modulations. Together, these studies suggest that modulation tuning is a conserved mechanism for the neural coding of spectrotemporally complex vocalizations.

Many songbird midbrain neurons are sensitive to statistical differences between classes of complex sounds. The tuning properties of these neurons differ during the processing of different sound classes such as vocalizations and noise [17, 85]. This phenomenon is called stimulus-dependent tuning [85] and is found in neurons of

other animals and other sensory systems [97, 105–107]. Tuning differences during the processing of complex sounds such as song and noise that differ in spectrotemporal structure are measured by comparing the STRFs of a single neuron calculated separately from responses to two sound classes. Stimulus-dependent tuning based on statistical differences among stimuli may maximize the mutual information between stimulus and response [108–113], facilitate neural discrimination of natural stimuli [17, 85], and correlate with changes in perceptual sensitivity [114, 115].

Midbrain neurons also exhibit "extra-classical" receptive fields (RFs) consisting of sideband excitation and sideband inhibition [70]. This tuning property serves as the underlying mechanism of stimulus-dependent tuning. Classical RFs are determined by responses to single tones that vary in frequency and intensity (Fig. 4.6c, d). Responses to frequencies within a neuron's classical RF are modulated by frequencies that fall outside of the classical RF. This makes neurons with extra-classical tuning sensitive to the structure of spectrally correlated sounds such as the noisy bursts of sound and harmonic stacks that characterize vocalizations such as zebra finch song and speech. Midbrain neurons, therefore, exhibit a simple nonlinearity that can account for the stimulus dependence of receptive fields estimated from the responses to sounds with natural and nonnatural statistics.

The spectrotemporal tuning properties of individual midbrain neurons cluster into functional groups based on how the neurons encode specific acoustic features important for perceptual qualities such as pitch, rhythm, and timber [87]. The four major functional groups are distinguished by the combined spectral and temporal properties of sounds to which they respond. For example, spectrally and temporally narrowband neurons encode brief sound segments that contain power at a specific frequency, possibly serving to encode pitch information. On the other hand, neurons with broad spectral tuning and narrow temporal tuning are sensitive to the onsets of sound segments at a large range of frequencies, possibly contributing to the encoding of rhythm.

The responses of auditory midbrain neurons to songs and other sounds show that tuning mechanisms to encode the special acoustic properties of song exist even at this early level in the auditory processing stream. The tuning of midbrain neurons for spectrotemporal modulations found in vocal sounds, the stimulus dependence of receptive fields, the neural discrimination of songs by temporal response patterns, and the functional grouping of spectrotemporal tuning among neurons demonstrate neural mechanisms for encoding communication vocalizations. These mechanisms subserve the formation of neural representations that can be used by downstream neurons to decode information such as species identity and individual identity.

The importance of subcortical processing in speech perception [35, 116, 117] indicates that understanding how the midbrain encodes vocalizations will be important for future prevention and remediation of speech and language disorders. One hypothesis is that human midbrain neurons, like songbird midbrain neurons, cluster based on spectrotemporal tuning and that these groups form the initial basis for the perceptual features that are well described in humans [118]. Based on this premise, we can also hypothesize that humans with impaired speech perception have poorly developed functional tuning groups or missing groups. Human brain imaging techniques could be used to test these hypotheses. For example, sound

stimuli can be designed to maximally excite neurons in putative functional tuning groups that encode acoustic features important for speech perception. These stimuli could then be presented during brain imaging in order to test for the presence of functional tuning groups in the human auditory system. Observed differences in imaging results between humans with and without normal speech perception could then provide organizational principles for the role of subcortical auditory processing in speech perception.

Thalamus

The avian auditory thalamus, nucleus ovoidalis (Ov) is homologous to the ventral division of the mammalian medial geniculate body (MGv; Fig. 4.4) and is tonotopically organized like midbrain and hindbrain nuclei [119, 120]. It is a small region of densely packed neurons organized into a core and a shell, with the core receiving the majority of the projections from the auditory midbrain [121] and providing the major input to the primary auditory forebrain [122]. Like the borders of the auditory midbrain, the Ov shell receives input from the cup of RA [60]. The functional significance of projections from this song production region back to the auditory thalamus has yet to be determined. But it could contribute to a feedback system that allows the vocal motor and auditory systems to coordinate song learning and production.

A small number of studies of the songbird auditory thalamus indicate that many neurons in the region respond to tones with excitatory V-shaped tuning curves, some with inhibitory sidebands [119, 123]. The firing properties of single neurons appear to be slightly less reliable and linear than in the midbrain [123, 124]. Spontaneous firing rates are higher and responses to songs contain more bursts of action potentials than in the midbrain [124]. From this limited information, the frequency tuning in the thalamus and midbrain are thought to be similar; however, the stimulus–response relationship may be less precise in the thalamus than in the midbrain at the level of the single neuron. It is also possible that the increased level of spontaneous and sound-driven activity in thalamic neurons contributes to the increased heterogeneity of spectrotemporal tuning properties among single neurons in the primary cortex compared to the midbrain [87]. The general conclusion based on preliminary comparisons of midbrain, thalamus, and cortex tuning is that the auditory thalamus is not simply a relay nucleus but likely shapes the diversity of tuning features that emerges in the cortex [124].

Primary Cortex

The major thalamo-recipient layer in songbird auditory cortex, field L2, is homologous to layer IV of the primary auditory cortex in mammals [51]. Field L2 projects to adjacent layers L1 and L3 (Fig. 4.4). These two layers project reciprocally to higher auditory processing regions and receive weak inputs from the shell region

surrounding Ov [58, 59, 62]. Tonotopic gradients are most clear in L2 but span across L1, L2, and L3, with best frequencies increasing from dorsolateral to ventromedial regions of each layer [125, 126]. Mapping studies using STRFs calculated from responses to amplitude modulated noise have revealed spatial patterns of spectrotemporal tuning in the cortex that are helpful for understanding how each processing region may encode song features differently. Many L2 neurons are narrowly tuned in both spectral and temporal dimensions [126, 127] corresponding to the encoding of changes in energy over time in a specific frequency band, with the temporal precision needed to detect acoustic features of short duration song notes [87]. The single and multiunit STRFs of neurons in L1 and L3 are more complex and tend to show a trade-off in coding specificity between spectral and temporal tuning. For example, neurons with broad spectral tuning have narrow temporal tuning [126, 127], potentially specializing in rhythm encoding [87]. At the multiunit level, the organization of spectral and temporal tuning has been summarized as two mapped patterns; spectral tuning grows broader across field L subregions in the medial to lateral direction and temporal tuning is narrow in L2, intermediate in L1, and broad in L3 [126].

Vocalization coding in field L is also well studied by directly analyzing the spike trains evoked by songs. Field L responses to song are on average slightly less robust and reliable than midbrain responses [17, 87, 128]. Nonetheless, the temporal precision and reliability of song responses of some neurons is sufficient to discriminate among the songs produced by different birds [91, 92]. Moreover, when the responses to large sets of songs are used to estimate STRFs, field L neurons cluster into spectrotemporal tuning groups that are similar to those in the midbrain.

While the major tuning groups in midbrain and primary cortex are shared, each field L group contains a larger range of spectral and temporal response characteristics among neurons, and some smaller tuning groups, to emerge [87]. This indicates an overall pattern of increasing heterogeneity and complexity in spectrotemporal tuning at higher processing stages of the ascending auditory system. Tuning for specific spectrotemporal modulations in the human primary and secondary auditory cortex also appears to map onto the modulations in natural sounds [129]. This suggests that modulation tuning serves as a mechanism for processing complex sounds such as vocalizations that is common to birds and mammals, including humans.

At the population level, primary forebrain neurons show response selectivity for song over other complex sounds as measured by average firing rate. Neurons generally fire more to song over other sounds such as tone complexes that are matched to song in frequency spectrum and power but do not have the same spectrotemporal statistics as song [130, 131]. This selectivity for song over synthetic sounds is weakly present in juveniles that are just beginning to develop song [34] and is present in non-singing females [10] suggesting that basic response selectivity for song over synthetic sounds such as tone complexes and noise does not require vocal learning. However, the excitability and strength of song selectivity in field L are lower in juvenile males than in adults, indicating that cortical auditory processing matures while developing birds are hearing and learning to produce song [34]. There is some debate over how response properties such as song selectivity differ between the primary thalamo-recipient region L2 and the secondary subregions L1

and L3. However, multiple studies examining responses to song syllables or synthetic sounds agree that L1 and L3 neurons tend to show more complex tuning than do L2 neurons [57, 125, 127, 132–134]. In summary, the auditory coding of song and tuning in general become progressively more complex and nonlinear from midbrain to cortex.

Higher Cortex

Responses to songs in higher auditory regions reveal a variety of complex coding properties that are not prevalent in field L. Field L sends inputs to two large auditory processing fields in the songbird brain, NCM and CM (Fig. 4.4). Field L subregions L2 and L3 project to NCM, which is posterior to L3. Subregions L2 and L1 project to CM, which is dorsal to field L. The response properties of many NCM and CM neurons are poorly characterized by linear-nonlinear models such as STRFs. Therefore, attempts to characterize the spectrotemporal tuning properties of neurons in these regions have proven more difficult than in lower regions. While some neurons in these regions respond strongly to simple stimuli such as tones [135], responses to songs indicate that higher cortex neurons are sensitive to recent stimulus history [136–140], the sound environment [141], and the behavioral salience of song stimuli [8, 142–144]. For example, NCM responses habituate to repeated presentation of the same song [136, 139] and show species-related response preferences [140, 145, 146]. Additionally, training adult birds to recognize songs alters the responses of NCM and CM to those songs. For example, CM neurons respond more to song motifs that are present in perceptually learned songs than to songs that have no particular behavioral salience [142]. Similar learning processes are associated with weaker responses in NCM neurons [144]. Song-specific habituation, species response preferences, and adult learning effects in NCM and CM make these regions promising candidates for testing the interactions between vocal learning and auditory processing of song. For example, NCM neurons in starlings deprived of hearing adult song during development are less selective for song features than are NCM neurons in wild-caught starlings [147]. Understanding interactions between experience and auditory processing in the songbird brain can contribute to developing hypotheses for how speech experience shapes auditory processing in the human brain.

Questions about perceptual and neural categorization of communication vocalizations are addressed in songbirds that use repertoires of song types. The starling, for example, is useful for studying how different regions of the cortex may be specialized to code categories of song information such as syllable types and motif types. In contrast to zebra finches in which one individual sings only one stereotyped motif, individual starlings sing many different song motifs, each with a spectrotemporally unique set of syllables. Each motif type is stereotyped and may be repeated or interleaved with other motif types during singing bouts [148, 149]. This song complexity provides an opportunity to ask whether neurons in specific

cortical areas are differentially sensitive to categories of vocal signals, potentially demonstrating hierarchical sensitivity to syllables and motifs as acoustic units that are perceptually categorized (see [18] for a comparison with the hierarchical organization of language). Comparisons of single neuron responses to multiple motif types in the field L subregions show that sensitivity to acoustic differences among song motif types differs significantly among the subregions [150]. Response selectivity for motif type is higher in L3 than in L2 and L1, suggesting that the song features that drive cortical neurons become more specific beyond the thalamo-recipient layer. In turn, the higher cortical region NCM, which receives strong projections from L3, also shows strong motif selectivity. The sensitivity of neurons to acoustic variations among repetitions of the same motif type is lower in CM than in the field L, suggesting that this region may be important for categorizing motif types, potentially as auditory objects. In agreement with studies using STRFs to characterize tuning, neural selectivity for acoustic features of song becomes increasingly complex between L2 and higher regions, and some regions may be specialized for coding song at particular levels of organization. Studies of cortical sound processing in mammals also support the general conclusion that neural coding of complex sounds such as song and speech becomes more specialized and complex in higher cortical processing stages [151–156].

NCM is a specific focus for investigating the neural basis of song memories that serve as templates for song learning. The hypothesis is that NCM houses circuitry for the formation and storage of tutor song memories [61]. In support of this idea, female zebra finches that do not sing or have intact song control systems also memorize songs they hear during development [39]. In zebra finches, the NCM expression of an immediate early gene (a gene that alters transcription of other genes) called *zenk* [157] becomes sensitive to song playback at the age that birds begin to memorize tutor songs but not before [145, 158]. The NCM *zenk* expression following song presentation is lower in juvenile males and females that have not experienced adult song compared to juveniles that have heard song from an adult. This suggests that gene transcription in NCM is influenced by early auditory experience [158, 159]. *Zenk* expression following presentation of a tutor's song and other songs does not differ in the adult NCM [160] but it is higher in response to tutor song playback in the CM and NCM of juvenile males that are learning to sing [161]. *Zenk* expression levels in NCM are also correlated with the accuracy of song learning [162, 163], and song learning is impaired by disruptions of *zenk* expression [164]. Electrophysiological data on the habituation rates of NCM responses to tutor songs also suggest that NCM neurons store information about tutor songs [165]. Whether and how the habituated responses of NCM neurons subserve tutor song memory has yet to be worked out. In summary, song responses in the higher auditory cortex are nonlinear compared to song responses in upstream auditory regions and influenced by developmental and adult experience with vocalizations. Studying the influence of experience on song coding properties in higher auditory cortex may be instrumental in defining mechanisms for applying social salience to vocal signals and learning vocal motor gestures by imitation.

Early Experience and Song Processing

The influence of speech experience—auditory, social, and motor—on human speech development is well studied [38, 166–168]. But the neural mechanisms whereby experience shapes speech development and perception can only be approached with low resolution and/or indirect assays of neural activity such as electroencephalography (EEG) or functional magnetic resonance imaging (fMRI). Use of high-resolution methods such as single-unit electrophysiology and measurement of gene expression has shown that early experience shapes songbird auditory coding in the midbrain and cortex. Deprivation of exposure to adult song and exposure to abnormal song during development significantly affect neural response strength (i.e., firing rates) and the selectivity of neurons' responses to song features (see [169] for review). For example, cortical responsivity and selectivity for specific song elements such as whistles is diminished in starlings that are completely deprived of hearing song as juveniles [9, 170]. In male zebra finches, midbrain and field L coding of song is compromised by abnormal developmental experience of song, largely due to lower than normal firing rates of single neurons in response to songs [88]. The auditory systems of non-singing females may not be similarly sensitive to early song exposure [130, 171]. It is therefore possible that the ability to learn vocalizations is correlated with significant experience-dependent plasticity in the auditory system. More studies are needed to establish a relationship between early song experience and auditory development. In any event, the current behavioral and neurophysiological evidence supports the hypothesis that songbird auditory processing, like human auditory processing, is influenced by early experience with vocal communication sounds. The songbird therefore provides an opportunity to study neural mechanisms whereby vocal experience shapes auditory system development and perception. Because the human auditory system is also slow to fully develop [172], the knowledge gained by studying songbird auditory development may lead to hypotheses about the influence of speech sounds on human auditory development that will impact the diagnoses and treatment of language processing disorders.

Conclusions

The songbird auditory pathway is a particularly useful system for examining vocalization processing because songbirds, like humans and unlike most other animals, are vocal learners. A rich history of behavioral and neurobiological studies on song learning has paved the way for using songbirds to study how auditory processing subserves vocal communication. The studies reviewed here identify principles of auditory coding in songbirds that may apply to human speech. One example of shared coding principles is the tuning for spectrotemporal modulations that characterize vocal communication sounds. There are bound to be more mechanisms of vocalization perception and learning that can be characterized in songbirds and tested for in humans. Laboratory approaches using songbirds that include

experience manipulations, behavioral assessments of perception, and high-resolution measures of auditory coding can help us develop an understanding of how the human brain acquires and understands speech.

Acknowledgments I thank Svetlana Rosis for helpful comments. This work was supported by grants from the NIDCD R01-DC009810 and the NSF IOS-0920081.

References

1. Brenowitz EA, Beecher MD (2005) Song learning in birds: diversity and plasticity, opportunities and challenges. Trends Neurosci 28(3):127–132
2. Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. Annu Rev Neurosci 22:567–631
3. Marler P (1997) Three models of song learning: evidence from behavior. J Neurobiol 33(5):501–516
4. Saar S, Mitra PP, Deregnacourt S, Tchernichovski O (2008) Developmental song learning in the zebra finch. In: Zeigler HP, Marler P (eds) Neuroscience of birdsong. Cambridge University Press, Cambridge, pp 217–227
5. Konishi M (2004) The role of auditory feedback in birdsong. Ann N Y Acad Sci 1016:463–475
6. Woolley SM (2008) Auditory feedback and singing in adult birds. In: Zeigler HP, Marler P (eds) Neuroscience of birdsong. Cambridge University Press, Cambridge, pp 228–239
7. Catchpole CK, Slater PJB (2008) Bird song: biological themes and variations, 2nd edn. Cambridge University Press, Cambridge, UK
8. Gentner TQ (2004) Neural systems for individual song recognition in adult birds. Ann N Y Acad Sci 1016:282–302
9. George I, Cousillas H, Vernier B et al (2004) Sound processing in the auditory-cortex homologue of songbirds: functional organization and developmental issues. J Physiol Paris 98(4–6):385–394
10. Hauber ME, Cassey P, Woolley SM, Theunissen FE (2007) Neurophysiological response selectivity for conspecific songs over synthetic sounds in the auditory forebrain of non-singing female songbirds. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 193(7):765–774
11. Poirier C, Boumans T, Verhoye M, Balthazart J, Van der Linden A (2009) Own-song recognition in the songbird auditory pathway: selectivity and lateralization. J Neurosci 29(7):2252–2258
12. Williams H, Crane LA, Hale TK, Esposito MA, Nottebohm F (1992) Right-side dominance for song control in the zebra finch. J Neurobiol 23(8):1006–1020
13. Reiner A, Perkel D, Mello CV, Jarvis ED (2008) Songbirds and the revised avian brain nomenclature. In: Zeigler HP, Marler P (eds) Neuroscience of birdsong. Cambridge University Press, Cambridge, pp 58–62
14. 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
15. Chi T, Gao Y, Guyton MC, Ru P, Shamma S (1999) Spectro-temporal modulation transfer functions and speech intelligibility. J Acoust Soc Am 106(5):2719–2732
16. Singh NC, Theunissen FE (2003) Modulation spectra of natural sounds and ethological theories of auditory processing. J Acoust Soc Am 114(6 pt 1):3394–3411
17. Woolley SM, Fremouw TE, Hsu A, Theunissen FE (2005) Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. Nat Neurosci 8(10):1371–1379
18. Kiggins JT, Comins JA, Gentner TQ (2012) Targets for a comparative neurobiology of language. Front Evol Neurosci 4:6

19. Knudsen DP, Gentner TQ (2010) Mechanisms of song perception in oscine birds. Brain Lang 115(1):59–68
20. Allard HA (1939) Vocal mimicry of the starling and the mockingbird. Science 90:370–371
21. Dooling RJ (1982) Ontogeny of song recognition in birds. Am Zool 22:571–580
22. Konishi M (1969) Hearing, single-unit analysis, and vocalizations in songbirds. Science 166(3909):1178–1181
23. Okanoya K, Dooling RJ (1987) Hearing in passerine and psittacine birds: a comparative study of absolute and masked auditory thresholds. J Comp Psychol 101(1):7–15
24. Heffner E, Heffner RS (2008) Audition. In: Davis SF (ed) Handbook of research methods in experimental psychology. Wiley-Blackwell, Hoboken, pp 413–440
25. Gleich O, Langemann U (2011) Auditory capabilities of birds in relation to the structural diversity of the basilar papilla. Hear Res 273(1–2):80–88
26. Woolley SM, Rubel EW (1999) High-frequency auditory feedback is not required for adult song maintenance in Bengalese finches. J Neurosci 19(1):358–371
27. Burda H, Ballast L, Bruns V (1988) Cochlea in old world mice and rats (Muridae). J Morphol 198(3):269–285
28. Koppl C (2011) Birds—same thing, but different? Convergent evolution in the avian and mammalian auditory systems provides informative comparative models. Hear Res 273(1–2): 65–71
29. Pickles JO (1992) An introduction to the physiology of hearing. Academic, London
30. Woolley SM, Rubel EW (2002) Vocal memory and learning in adult Bengalese Finches with regenerated hair cells. J Neurosci 22(17):7774–7787
31. Woolley SM, Wissman AM, Rubel EW (2001) Hair cell regeneration and recovery of auditory thresholds following aminoglycoside ototoxicity in Bengalese finches. Hear Res 153(1–2):181–195
32. Gray L, Rubel EW (1985) Development of absolute thresholds in chickens. J Acoust Soc Am 77(3):1162–1172
33. Lickliter R, Stoumbos J (1991) Enhanced prenatal auditory experience facilitates species-specific visual responsiveness in bobwhite quail chicks (Colinus virginianus). J Comp Psychol 105(1):89–94
34. Amin N, Doupe A, Theunissen FE (2007) Development of selectivity for natural sounds in the songbird auditory forebrain. J Neurophysiol 97(5):3517–3531
35. Johnson KL, Nicol T, Zecker SG, Kraus N (2008) Developmental plasticity in the human auditory brainstem. J Neurosci 28(15):4000–4007
36. Eales LA (1985) Song learning in zebra finches: some effects of song model availability on what is learnt and when. Anim Behav 33:1293–1300
37. Immelmann K (1969) Song development in the zebra finch and other estrildid finches. In: Hinde RA (ed) Bird vocalizations. Cambridge University Press, Cambridge, pp 61–77
38. Kuhl PK (2004) Early language acquisition: cracking the speech code. Nat Rev Neurosci 5(11):831–843
39. Miller DB (1979) Long-term recognition of father's song by female zebra finches. Nature 280:389–391
40. Riebel K (2009) Song and female mate choices in zebra finches—a review. Adv Stud Behav 40:197–238
41. Campbell DLM, Hauber ME (2009) Cross-fostering diminishes song discrimination in zebra finches (Taeniopygia guttata). Anim Cogn 12(3):481–490
42. Clayton NS (1988) Song learning and mate choice in estrildid finches raised by 2 species. Anim Behav 36:1589–1600
43. Clayton NS, Prove L (1989) Song discrimination in female zebra finches and Bengalese finches. Anim Behav 38:352–353
44. 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(6):1249–1255
45. Nagle L, Kreutzer MI (1997) Song tutoring influences female song preferences in domesticated canaries. Behaviour 134:89–104

46. Peters S, Searcy WA, Marler P (1980) Species song discrimination in choice experiments with territorial male swamp and song sparrows. Anim Behav 28:393–404
47. Gleich O (1994) Excitation patterns in the starling cochlea: a population study of primary auditory afferents. J Acoust Soc Am 95(1):401–409
48. Young ED (2008) Neural representation of spectral and temporal information in speech. Philos Trans R Soc Lond B Biol Sci 363(1493):923–945
49. Shamma SA (1985) Speech processing in the auditory system. I: The representation of speech sounds in the responses of the auditory nerve. J Acoust Soc Am 78(5):1612–1621
50. Boord RL (1968) Ascending projections of the primary cochlear nuclei and nucleus laminaris in the pigeon. J Comp Neurol 133:523–542
51. Butler AB, Reiner A, Karten HJ (2011) Evolution of the amniote pallium and the origins of mammalian neocortex. Ann N Y Acad Sci 1225:14–27
52. Jarvis ED, Gunturkun O, Bruce L et al (2005) Avian brains and a new understanding of vertebrate brain evolution. Nat Rev Neurosci 6(2):151–159
53. Krutzfeldt NO, Logerot P, Kubke MF, Wild JM (2010) Connections of the auditory brainstem in a songbird, Taeniopygia guttata. I. Projections of nucleus angularis and nucleus laminaris to the auditory torus. J Comp Neurol 518(11):2109–2134
54. Krutzfeldt NO, Logerot P, Kubke MF, Wild JM (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
55. Wild JM, Krutzfeldt NO, Kubke MF (2010) Connections of the auditory brainstem in a songbird, Taeniopygia guttata. III. Projections of the superior olive and lateral lemniscal nuclei. J Comp Neurol 518(11):2149–2167
56. Fortune ES, Margoliash D (1992) Cytoarchitectonic organization and morphology of cells of the field L complex in male zebra finches (Taenopygia guttata). J Comp Neurol 325(3):388–404
57. Bonke BA, Bonke D, Scheich H (1979) Connectivity of the auditory forebrain nuclei in the guinea fowl (Numida meleagris). Cell Tissue Res 200(1):101–121
58. 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
59. Zeng S, Zhang X, Peng W, Zuo M (2004) Immunohistochemistry and neural connectivity of the Ov shell in the songbird and their evolutionary implications. J Comp Neurol 470(2):192–209
60. Mello CV, Vates GE, Okuhata S, Nottebohm F (1998) Descending auditory pathways in the adult male zebra finch (Taeniopygia guttata). J Comp Neurol 395(2):137–160
61. Bolhuis JJ, Gahr M (2006) Neural mechanisms of birdsong memory. Nat Rev Neurosci 7(5):347–357
62. Kelley DB, Nottebohm F (1979) Projections of a telencephalic auditory nucleus-field L-in the canary. J Comp Neurol 183(3):455–469
63. 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
64. Pinaud R, Terleph TA, Tremere LA et al (2008) Inhibitory network interactions shape the auditory processing of natural communication signals in the songbird auditory forebrain. J Neurophysiol 100(1):441–455
65. Carr CE, Code RA (2000) The central auditory systems of reptiles and birds. In: Fay RR, Popper AN (eds) Comparative hearing: birds and reptiles. Springer, New York
66. Konishi M (1970) Comparative neurophysiological studies of hearing and vocalizations in songbirds. Z Vergl Physiol 66:257–272
67. Sachs MB, Sinnott JM (1978) Responses to tones of single cells in nucleus magnocellularis and nucleus angularis of the redwing blackbird (Agelaius phoeniceus). J Comp Physiol 126:347–361
68. Sachs MB, Woolf NK, Sinnott JM (1980) Response properties of neurons in the avian auditory system: comparisons with mammalian homologues and consideration of the neural encoding

of complex stimuli. In: Popper AN, Fay RR (eds) Comparative studies of hearing in vertebrates. Springer, Berlin, pp 223–253
69. Oertel D (1999) The role of timing in the brain stem auditory nuclei of vertebrates. Annu Rev Physiol 61:497–519
70. Schneider DM, Woolley SM (2011) Extra-classical tuning predicts stimulus-dependent receptive fields in auditory neurons. J Neurosci 31(33):11867–11878
71. Roos MJ, May BJ (2012) Classification of unit types in the anteroventral cochlear nucleus of laboratory mice. Hear Res 289(1–2):13–26
72. Young ED, Brownell WE (1976) Responses to tones and noise of single cells in dorsal cochlear nucleus of unanesthetized cats. J Neurophysiol 39(2):282–300
73. Pfeiffer RR (1966) Classification of response patterns of spike discharges for units in the cochlear nucleus: tone-burst stimulation. Exp Brain Res 1(3):220–235
74. Kiang NY, Pfeiffer RR, Warr WB, Backus AS (1965) Stimulus coding in the cochlear nucleus. Ann Otol Rhinol Laryngol 74:463–485
75. Braun K, Scheich H, Heizmann CW, Hunziker W (1991) Parvalbumin and calbindin-D28K immunoreactivity as developmental markers of auditory and vocal motor nuclei of the zebra finch. Neuroscience 40(3):853–869
76. Logerot P, Krutzfeldt NO, Wild JM, Kubke MF (2011) Subdivisions of the auditory midbrain (N. Mesencephalicus lateralis, pars dorsalis) in zebra finches using calcium-binding protein immunocytochemistry. PLoS One 6(6):e20686
77. Zeng S, Li J, Zhang X, Zuo M (2007) Distinction of neurochemistry between the cores and their shells of auditory nuclei in tetrapod species. Brain Behav Evol 70(1):1–20
78. Cant NB (2005) Projections from the cochlear nuclear complex to the inferior colliculus. In: Winer JA, Schreiner CE (eds) The inferior colliculus. Springer, New York, pp 115–131
79. Conlee JW, Parks TN (1986) Origin of ascending auditory projections to the nucleus mesencephalicus lateralis pars dorsalis in the chicken. Brain Res 367(1–2):96–113
80. Takahashi TT, Konishi M (1988) Projections of the cochlear nuclei and nucleus laminaris to the inferior colliculus of the barn owl. J Comp Neurol 274(2):190–211
81. Wild JM (1995) Convergence of somatosensory and auditory projections in the avian torus semicircularis, including the central auditory nucleus. J Comp Neurol 358(4):465–486
82. Schumacher JW, Schneider DM, Woolley SM (2011) Anesthetic state modulates excitability but not spectral tuning or neural discrimination in single auditory midbrain neurons. J Neurophysiol 106(2):500–514
83. Woolley SM, Casseday JH (2004) Response properties of single neurons in the zebra finch auditory midbrain: response patterns, frequency coding, intensity coding, and spike latencies. J Neurophysiol 91(1):136–151
84. Woolley SM, Casseday JH (2005) Processing of modulated sounds in the zebra finch auditory midbrain: responses to noise, frequency sweeps, and sinusoidal amplitude modulations. J Neurophysiol 94(2):1143–1157
85. Woolley SM, Gill PR, Theunissen FE (2006) Stimulus-dependent auditory tuning results in synchronous population coding of vocalizations in the songbird midbrain. J Neurosci 26(9):2499–2512
86. Schneider DM, Woolley SM (2010) Discrimination of communication vocalizations by single neurons and groups of neurons in the auditory midbrain. J Neurophysiol 103(6):3248–3265
87. Woolley SM, Gill PR, Fremouw T, Theunissen FE (2009) Functional groups in the avian auditory system. J Neurosci 29(9):2780–2793
88. Woolley SM, Hauber ME, Theunissen FE (2010) Developmental experience alters information coding in auditory midbrain and forebrain neurons. Dev Neurobiol 70(4):235–252
89. van Rossum MC (2001) A novel spike distance. Neural Comput 13(4):751–763
90. Victor JD (2005) Spike train metrics. Curr Opin Neurobiol 15(5):585–592
91. Wang L, Narayan R, Grana G, Shamir M, Sen K (2007) Cortical discrimination of complex natural stimuli: can single neurons match behavior? J Neurosci 27(3):582–589

92. Narayan R, Best V, Ozmeral E et al (2007) Cortical interference effects in the cocktail party problem. Nat Neurosci 10(12):1601–1607
93. Ramirez AD, Ahmadian Y, Schumacher J, Schneider D, Woolley SM, Paninski L (2011) Incorporating naturalistic correlation structure improves spectrogram reconstruction from neuronal activity in the songbird auditory midbrain. J Neurosci 31(10):3828–3842
94. Pasley BN, David SV, Mesgarani N et al (2012) Reconstructing speech from human auditory cortex. PLoS Biol 10(1):e1001251
95. Calabrese A, Schumacher JW, Schneider DM, Paninski L, Woolley SM (2011) A generalized linear model for estimating spectrotemporal receptive fields from responses to natural sounds. PLoS One 6(1):e16104
96. Theunissen FE, Sen K, Doupe AJ (2000) Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. J Neurosci 20(6):2315–2331
97. 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(11):3374–3386
98. Andoni S, Li N, Pollak GD (2007) Spectrotemporal receptive fields in the inferior colliculus revealing selectivity for spectral motion in conspecific vocalizations. J Neurosci 27(18):4882–4893
99. Andoni S, Pollak GD (2011) Selectivity for spectral motion as a neural computation for encoding natural communication signals in bat inferior colliculus. J Neurosci 31(46):16529–16540
100. Casseday JH, Covey E (1996) A neuroethological theory of the operation of the inferior colliculus. Brain Behav Evol 47(6):311–336
101. Edwards CJ, Leary CJ, Rose GJ (2007) Counting on inhibition and rate-dependent excitation in the auditory system. J Neurosci 27(49):13384–13392
102. Elliott TM, Christensen-Dalsgaard J, Kelley DB (2011) Temporally selective processing of communication signals by auditory midbrain neurons. J Neurophysiol 105(4):1620–1632
103. Elliott TM, Theunissen FE (2009) The modulation transfer function for speech intelligibility. PLoS Comput Biol 5(3):e1000302
104. Sabin AT, Eddins DA, Wright BA (2012) Perceptual learning evidence for tuning to spectrotemporal modulation in the human auditory system. J Neurosci 32(19):6542–6549
105. Blake DT, Merzenich MM (2002) Changes of AI receptive fields with sound density. J Neurophysiol 88(6):3409–3420
106. Nagel KI, Doupe AJ (2006) Temporal processing and adaptation in the songbird auditory forebrain. Neuron 51(6):845–859
107. Lesica NA, Grothe B (2008) Dynamic spectrotemporal feature selectivity in the auditory midbrain. J Neurosci 28(21):5412–5421
108. Escabi MA, Miller LM, Read HL, Schreiner CE (2003) Naturalistic auditory contrast improves spectrotemporal coding in the cat inferior colliculus. J Neurosci 23(37):11489–11504
109. David SV, Vinje WE, Gallant JL (2004) Natural stimulus statistics alter the receptive field structure of v1 neurons. J Neurosci 24(31):6991–7006
110. Sharpee TO, Sugihara H, Kurgansky AV, Rebrik SP, Stryker MP, Miller KD (2006) Adaptive filtering enhances information transmission in visual cortex. Nature 439(7079):936–942
111. Fairhall AL, Lewen GD, Bialek W, de Ruyter Van Steveninck RR (2001) Efficiency and ambiguity in an adaptive neural code. Nature 412(6849):787–792
112. Maravall M, Petersen RS, Fairhall AL, Arabzadeh E, Diamond ME (2007) Shifts in coding properties and maintenance of information transmission during adaptation in barrel cortex. PLoS Biol 5(2):e19
113. Lesica NA, Grothe B (2008) Efficient temporal processing of naturalistic sounds. PLoS One 3(2):e1655
114. Webster MA, Georgeson MA, Webster SM (2002) Neural adjustments to image blur. Nat Neurosci 5(9):839–840

115. Dahmen JC, Keating P, Nodal FR, Schulz AL, King AJ (2010) Adaptation to stimulus statistics in the perception and neural representation of auditory space. Neuron 66(6):937–948
116. Johnson KL, Nicol TG, Kraus N (2005) Brain stem response to speech: a biological marker of auditory processing. Ear Hear 26(5):424–434
117. Kraus N, Skoe E, Parbery-Clark A, Ashley R (2009) Experience-induced malleability in neural encoding of pitch, timbre, and timing. Ann N Y Acad Sci 1169:543–557
118. Moore BC (2008) Basic auditory processes involved in the analysis of speech sounds. Philos Trans R Soc Lond B Biol Sci 363(1493):947–963
119. Bigalke-Kunz B, Rubsamen R, Dorrscheidt GJ (1987) Tonotopic organization and functional characterization of the auditory thalamus in a songbird, the European starling. J Comp Physiol A 161(2):255–265
120. Proctor L, Konishi M (1997) Representation of sound localization cues in the auditory thalamus of the barn owl. Proc Natl Acad Sci U S A 94(19):10421–10425
121. Karten HJ (1967) The organization of the ascending auditory pathway in the pigeon (Columba livia). I. Diencephalic projections of the inferior colliculus (nucleus mesencephali lateralis, pars dorsalis). Brain Res 6(3):409–427
122. Karten HJ (1968) The ascending auditory pathway in the pigeon (Columba livia). II. Telencephalic projections of the nucleus ovoidalis thalami. Brain Res 11(1):134–153
123. Margoliash D, Bankes SC (1993) Computations in the ascending auditory pathway in songbirds related to song learning. Am Zool 33:94–103
124. 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
125. Capsius B, Leppelsack H (1999) Response patterns and their relationship to frequency analysis in auditory forebrain centers of a songbird. Hear Res 136(1–2):91–99
126. Kim G, Doupe A (2011) Organized representation of spectrotemporal features in songbird auditory forebrain. J Neurosci 31(47):16977–16990
127. Nagel KI, Doupe AJ (2008) Organizing principles of spectro-temporal encoding in the avian primary auditory area field L. Neuron 58(6):938–955
128. Schneider DM, Woolley SM (2012) Sparse coding neurons encode individual vocalizations in complex auditory scenes. Computation and Systems Neuroscience, Salt Lake City
129. Schonwiesner M, Zatorre RJ (2009) Spectro-temporal modulation transfer function of single voxels in the human auditory cortex measured with high-resolution fMRI. Proc Natl Acad Sci U S A 106(34):14611–14616
130. Hauber ME, Woolley SMN, Theunissen FE (2007) Experience-dependence of neural responses to social vs. isolate conspecific songs in the forebrain of female zebra finches. J Ornithol 148(2):231–239
131. Grace JA, Amin N, Singh NC, Theunissen FE (2003) Selectivity for conspecific song in the zebra finch auditory forebrain. J Neurophysiol 89(1):472–487
132. Cousillas H, Leppelsack HJ, Leppelsack E, Richard JP, Mathelier M, Hausberger M (2005) Functional organization of the forebrain auditory centres of the European starling: a study based on natural sounds. Hear Res 207(1–2):10–21
133. Lim D, Kim C (1997) Emerging auditory response interactions to harmonic complexes in field L of the zebra finch. Auris Nasus Larynx 24(3):227–232
134. Sen K, Theunissen FE, Doupe AJ (2001) Feature analysis of natural sounds in the songbird auditory forebrain. J Neurophysiol 86(3):1445–1458
135. Terleph TA, Mello CV, Vicario DS (2006) Auditory topography and temporal response dynamics of canary caudal telencephalon. J Neurobiol 66(3):281–292
136. 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 U S A 92(8):3406–3410
137. Chew SJ, Vicario DS, Nottebohm F (1996) A large-capacity memory system that recognizes the calls and songs of individual birds. Proc Natl Acad Sci U S A 93(5):1950–1955
138. Gill P, Woolley SM, Fremouw T, Theunissen FE (2008) What's that sound? Auditory area CLM encodes stimulus surprise, not intensity or intensity changes. J Neurophysiol 99(6):2809–2820

139. 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(10):6919–6925
140. Mello CV, Clayton DF (1994) Song-induced ZENK gene expression in auditory pathways of songbird brain and its relation to the song control system. J Neurosci 14(11 pt 1): 6652–6666
141. Terleph TA, Lu K, Vicario DS (2008) Response properties of the auditory telencephalon in songbirds change with recent experience and season. PLoS One 3(8):e2854
142. Gentner TQ, Margoliash D (2003) Neuronal populations and single cells representing learned auditory objects. Nature 424(6949):669–674
143. 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
144. 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
145. Stripling R, Kruse AA, Clayton DF (2001) Development of song responses in the zebra finch caudomedial neostriatum: role of genomic and electrophysiological activities. J Neurobiol 48(3):163–180
146. Stripling R, Volman SF, Clayton DF (1997) Response modulation in the zebra finch neostriatum: relationship to nuclear gene regulation. J Neurosci 17(10):3883–3893
147. George I, Alcaix S, Henry L, Richard JP, Cousillas H, Hausberger M (2010) Neural correlates of experience-induced deficits in learned vocal communication. PLoS One 5(12):e14347
148. Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC (2006) Recursive syntactic pattern learning by songbirds. Nature 440(7088):1204–1207
149. Gentner TQ, Hulse SH (1998) Perceptual mechanisms for individual vocal recognition in European starlings, Sturnus vulgaris. Anim Behav 56(3):579–594
150. Meliza CD, Margoliash D (2012) Hierarchical emergence of selectivity and tolerance in the auditory cortex of European Starlings. In: Auditory System Gordon Research Conference, Lewiston, ME
151. Altmann CF, Doehrmann O, Kaiser J (2007) Selectivity for animal vocalizations in the human auditory cortex. Cereb Cortex 17(11):2601–2608
152. Leaver AM, Rauschecker JP (2010) Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. J Neurosci 30(22):7604–7612
153. Linden JF, Liu RC, Sahani M, Schreiner CE, Merzenich MM (2003) Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. J Neurophysiol 90(4):2660–2675
154. Nelken I (2008) Processing of complex sounds in the auditory system. Curr Opin Neurobiol 18(4):413–417
155. Petkov CI, Kayser C, Steudel T, Whittingstall K, Augath M, Logothetis NK (2008) A voice region in the monkey brain. Nat Neurosci 11(3):367–374
156. Rauschecker JP, Scott SK (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat Neurosci 12(6):718–724
157. Mello CV, Velho TA, Pinaud R (2004) Song-induced gene expression: a window on song auditory processing and perception. Ann N Y Acad Sci 1016:263–281
158. 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
159. Tomaszycki ML, Sluzas EM, Sundberg KA, Newman SW, DeVoogd TJ (2006) Immediate early gene (ZENK) responses to song in juvenile female and male zebra finches: effects of rearing environment. J Neurobiol 66(11):1175–1182
160. Terpstra NJ, Bolhuis JJ, den Boer-Visser AM (2004) An analysis of the neural representation of birdsong memory. J Neurosci 24(21):4971–4977
161. Gobes SM, Zandbergen MA, Bolhuis JJ (2010) Memory in the making: localized brain activation related to song learning in young songbirds. Proc Biol Sci 277(1698):3343–3351
162. Bolhuis JJ, Hetebrij E, Den Boer-Visser AM, De Groot JH, Zijlstra GG (2001) Localized immediate early gene expression related to the strength of song learning in socially reared zebra finches. Eur J Neurosci 13(11):2165–2170

163. Bolhuis JJ, Zijlstra GG, den Boer-Visser AM, Van Der Zee EA (2000) Localized neuronal activation in the zebra finch brain is related to the strength of song learning. Proc Natl Acad Sci U S A 97(5):2282–2285
164. London SE, Clayton DF (2008) Functional identification of sensory mechanisms required for developmental song learning. Nat Neurosci 11(5):579–586
165. 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 U S A 103(4):1088–1093
166. Dietrich C, Swingley D, Werker JF (2007) Native language governs interpretation of salient speech sound differences at 18 months. Proc Natl Acad Sci U S A 104(41):16027–16031
167. Kuhl PK (2007) Is speech learning 'gated' by the social brain? Dev Sci 10(1):110–120
168. Werker JF, Tees RC (1999) Influences on infant speech processing: toward a new synthesis. Annu Rev Psychol 50:509–535
169. Woolley SM (2012) Early experience shapes vocal neural coding and perception in songbirds. Dev Psychobiol 54(6):612–631
170. Cousillas H, Richard JP, Mathelier M, Henry L, George I, Hausberger M (2004) Experience-dependent neuronal specialization and functional organization in the central auditory area of a songbird. Eur J Neurosci 19(12):3343–3352
171. Maul KK, Voss HU, Parra LC et al (2010) The development of stimulus-specific auditory responses requires song exposure in male but not female zebra finches. Dev Neurobiol 70(1):28–40
172. Saffran JR, Werker JF, Werner LA (2006) The infant's auditory world: hearing, speech and the beginnings of language. In: Seigler R, Kuhn D (eds) Handbook for child development. Wiley, New York, pp 58–108
173. Heffner RS, Koay G, Heffner HE (2001) Audiograms of five species of rodents: implications for the evolution of hearing and the perception of pitch. Hear Res 157(1–2):138–152

Chapter 5
Prospective: How the Zebra Finch Genome Strengthens Brain-Behavior Connections in Songbird Models of Learned Vocalization

Sarah E. London

Abstract Songbirds are a premier model for speech and language; brain and behavior studies of song learning, perceptual processing, and production have uncovered processes fundamental to vocal communication. Recently, the first songbird genome has been sequenced and assembled. The genome represents great opportunity to advance discovery of neural mechanisms of song. Notably, it enables simultaneous measurements of thousands of genes, prediction and testing of the connections between those genes, examination of nonprotein coding RNAs, and functional structural elements. Researchers have just begun to explore these aspects. This review will therefore provide an overview of how genomic elements may interrelate with song.

Keywords Zebra finch • Songbird • Genome • Neural circuit • Plasticity • Experience • Song • Evolution

Introduction

We are at the start of an exciting era in the investigation of learned vocalizations using songbird model systems. Creative researchers are employing rapidly advancing technology to tackle various dimensions of complex behaviors such as song. Building on a solid 50-year foundation of behavioral work that established the songbird as a premier animal model for speech and language, recent studies have added new insights [1–5]. The genome is an essential resource for future discoveries into the fundamental neural mechanisms of song.

S.E. London, Ph.D. (✉)
Department of Psychology, University of Chicago, 940 East 57th Street, Chicago, IL 60637, USA
e-mail: london@uchicago.edu

S.A. Helekar (ed.), *Animal Models of Speech and Language Disorders*,
DOI 10.1007/978-1-4614-8400-4_5,

The genome tracks with behavior and brain on multiple timescales, from evolutionary time, across an animal's lifetime, and in acute moment-to-moment experiences. Species emerge on an evolutionary timescale. Close to 5,000 extant species of songbirds (order Passerine, suborder Oscine) live all over the world, each with their own unique neurobiology and behavior. A genetic program to direct development and maintenance of the brain circuitry that controls singing, and how it connects with the auditory forebrain area responsible for auditory processing and learning, must have evolved in songbirds.

Presumably, precise regulation of this genomic program differs between species whose strategy for song learning and singing differs, though all songbirds share some common features. Further, the process of vocal learning has strong parallels with speech acquisition. In both songbirds and humans, an individual forms an auditory memory of the vocal element to be copied, then uses a process of sensorimotor error correction to shape his own vocal output to resemble that sound element. In addition, the basic cortico-thalamic-basal ganglia neuroanatomical circuitry that directs learned vocalizations in humans is necessary for song learning in songbirds. Few animals are known to have the ability to learn vocalizations; songbirds are among the most tractable for deep investigation into the neurogenomic underpinnings of brain function and behavior that surround vocal learning.

Brain development in songbirds occurs primarily after hatching. Thus neural development is affected not only by genetic programming but also by the bird's experience. Especially in species that can only learn to sing once during development, the interaction of age- and experience-dependent genomic changes that alter brain function can have lifelong consequences for song. The genome must therefore respond to acute experience. This includes the experience interacting with other animals socially (including hearing singing) and the experience producing vocalizations. Dynamic regulation of the genome during acute experience reflects previous experience and is a mechanism by which a trace of the current experience can persist.

This chapter will examine genomic elements that reflect each of these timescales and how they might relate to brain function and behaviors like song. The zebra finch (*Taeniopygia guttata*) will be the primary species discussed, as it is the zebra finch songbird genome that is currently sequenced and assembled, and it has been the model for many studies investigating genome-brain-behavior interrelationships. Information will be presented at a high level because the zebra finch genome assembly was released recently, April 2010 [6].

The studies described here therefore represent how new approaches and access to information beyond individual protein-coding sequences will advance the songbird as a model for speech and language. Research is just beginning in this post-genomic era, thus almost all published work describes discoveries that require additional validation. The goal here is not to draw strong conclusions about how these early results directly relate to song. Rather the goal is to familiarize the reader with genomic characteristics that could be mechanisms of song and to review the recent studies as proof of concept for how the genome advances study from individual protein-coding gene-based examination to broader scale genomic dynamics. Ultimately, these types of studies will only deepen our appreciation for the parallels between vocal learning in songbirds and humans.

DNA Sequence and Song

Discovery of Coordinated Gene Sets

The study of individual genes has provided essential insight into the neurobiology of vocal learning in the last ~30 years [7–13]. However, no one gene can tell the whole story of how the brain interacts with behavior and experience; the Ensembl gene prediction (build version 71) contains 19,334 known and predicted gene transcripts, including pseudogenes, in the zebra finch genome. The post-genome perspective emphasizes simultaneous measurement of multiple genes, and consideration of the interactions between these gene sets, to gain insight into broader mechanisms that are involved in behavior. There are currently two major methods for measuring expression levels of thousands of genes simultaneously: microarray and "next-generation" whole-genome direct RNA sequencing.

Gene Sets: Microarray

A genome assembly is not necessary to produce and use a DNA microarray, although the genome aids greatly in annotation and permits identification of nonprotein-coding transcribed elements (discussed below). Microarrays do allow for gene discovery; no knowledge is necessary a priori to test if a gene is relevant to a particular brain area or process. Microarrays were therefore an important first step in the transition from single-gene examination to multigene set exploration. Four spotted DNA microarrays were created between 2004 and 2008 [14–17]. Each project generated DNA libraries with thousands of expressed sequence tags (ESTs) cloned from the zebra finch brain. All four arrays were used to identify genes that might be involved in zebra finch song.

Discovery of Static "Marker" Gene Sets

It is still unknown how the song system is masculinized but both steroid signaling and sex chromosome gene expression likely contribute [10, 11, 13–22]. To discover a gene set that differs between males and females, "brain sex markers," microarray experiments compared gene expression profiles in the telencephalon of developing birds. In one of the first songbird microarray experiments, ~300 ESTs showed a significant sex difference [17, 23]. The authors confirmed a sex difference in mRNA levels within major singing control nuclei for eight previously unstudied genes. Although small in number, this experiment demonstrated that meaningful gene discovery was possible with microarrays.

Two larger-scale experiments were done to answer the question of whether or not there were genes that had expression levels unique to the adult male song control nucleus HVC [14, 24]. Together, these experiments found ~1,200 ESTs (800 from Li; 400 from Lovell) that met the statistical criteria set to identify neuroanatomical

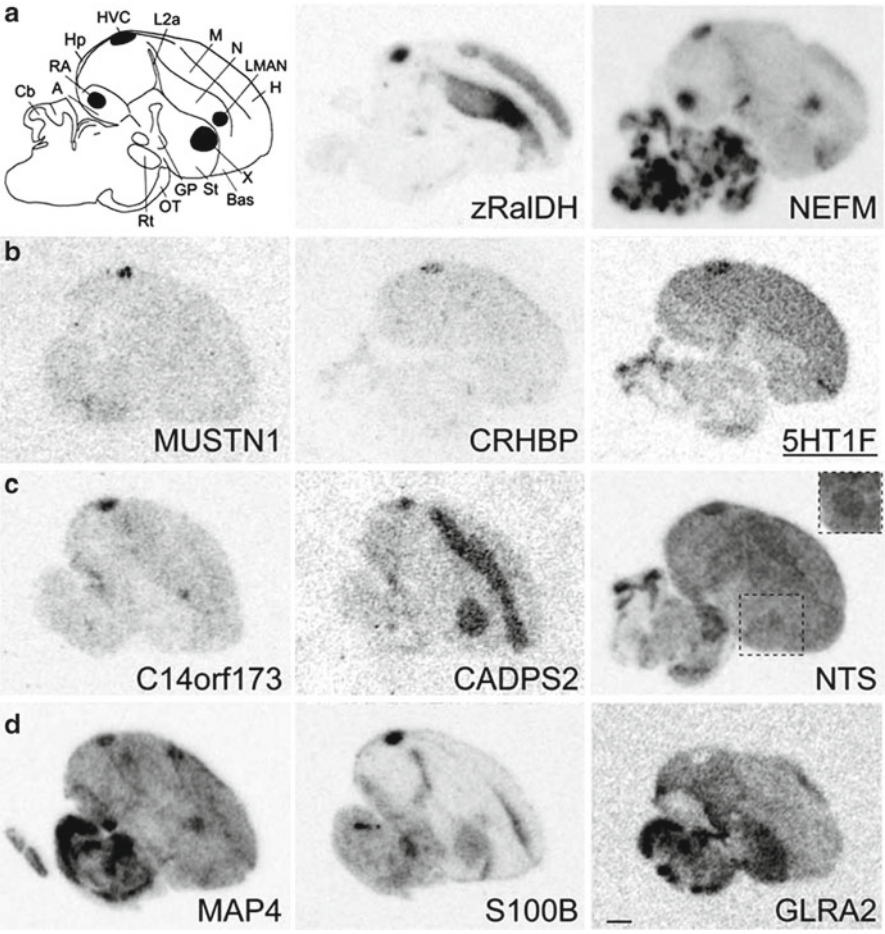


Fig. 5.1 Patterns of gene expression reveal specializations in song control areas. Expression of known HVC markers (**a**–**d**). Representative in situ hybridization autoradiograms of parasagittal sections of adult male zebra finches at the level of HVC (~1.4–2.4 mm from midline). (**a**, *left*) Schematic depicting neuroanatomy including the major telencephalic song control nucleus HVC (in *black*). *Reused with permission from Lovell, P. V., Clayton, D. F., Replogle, K. L. & Mello, C. V. Birdsong "transcriptomics": neurochemical specializations of the oscine song system. PloS One 3, e3440, doi:10.1371/journal.pone.0003440 (2008). Open Source Article* [24]

"marker" genes. With this number of genes, it is possible to analyze the functional categories over- and underrepresented in the gene sets to indicate what processes may be specialized to HVC. Genes in the Li study were disproportionately related to gene expression and protein translation, as analyzed by pathway analysis [14]. Gene ontology (GO) functional annotation analysis used in the Lovell study revealed that signal transduction, ion transport, and synaptic transmission had an overrepresentation of genes (Fig. 5.1) [24]. Based on GO analysis, HVC marker genes in the Lovell study were statistically more likely to encode proteins localized within the plasma membrane than the nucleus.

Discovery of Dynamic Gene Sets: Acute Timescale

Song processing and production requires fast modulation of cellular activity. Rapid and robust immediate early gene expression after hearing song playbacks and singing demonstrated that the genome also responded dynamically to these behaviors [9, 12, 25–29]. Microarrays were used to test whether or not immediate early genes were the only genes invoked by acute experience, and if not, to characterize other experience-dependent genes.

In the song control circuit, several microarray experiments identified genes regulated by singing and have organized these gene sets into functional categories and proposed transcriptional networks [6, 16, 30, 31]. In one experiment, 33 out of a predicted set of 150 genes were validated to show significant mRNA changes in several motor control nuclei after the bird sang; two of these were the well-studied immediate early genes ZENK (an acronym for zif286, egr-1, ngfi-a, krox24) and c-fos [16]. Thirty-one of the 33 genes showed increased expression levels after singing compared to non-singing controls. More than a third of all of the genes were predicted to function within the cell nucleus, and 19 % of them were categorized as transcription factors, consistent with the idea that song initiates several transcriptional cascades that could have cellular effects after the singing ceases [32]. Additional experiments targeted the basal ganglia component of the motor circuit and specifically analyzed thousands of singing-regulated genes for transcriptional control mechanisms. These studies reported the regulation of several co-expression networks and potential regulatory transcription factors including CREB, NFKB, NTRK2, and FOXP2 [6, 30, 31].

Hearing novel song playbacks also invokes the genome in adult male zebra finches in the auditory forebrain [33]. Approximately 600 ESTs showed significant changes in birds that heard a novel song playback compared to those who were not presented with auditory stimulus, i.e., left in silence. More than half of the song-responsive genes have lower mRNA levels after hearing song compared to their baseline level in the silence condition. Further, approximately 65 % of the song-responsive genes do not show a significant change from silence baseline when the song the bird heard was familiar to him. This process of genomic habituation, the attenuation of a response after repeated exposure to a stimulus, may relate to behavioral habituation, a form of nonassociative learning [33, 34].

Discovery of Dynamic Gene Sets: Developmental Timescale

The function of song areas shifts across development as birds enter, then exit, the sensitive period for song copying. Microarray experiments demonstrated that this functional change is associated with large-scale changes in gene expression. Approximately 900 genes show altered expression levels in the auditory forebrain between birds too young to learn song and adults, who have completed song learning [35]. These genes show the same expression levels regardless of whether or not the

birds had been exposed to acute song playback experience. In fact, the young birds do not show a song response at all; genes regulated by song playbacks in adults are expressed at a constitutively high level in the young auditory forebrain. Thus this 900-gene set likely represents "maturation" genes that reflect a combination of advanced age and accumulated experience. Functional GO categories of cell proliferation and differentiation are overrepresented in genes that showed higher P20 than adult expression levels. Genes that showed higher expression levels in adult than P20 auditory forebrain are overrepresented in cell death and transcriptional regulation.

Meta-Gene Sets: Microarray

With extreme care, statistical comparisons can be made across experiments. One algorithm employed for this purpose is called weighted gene co-expression network analysis (WGCNA) [36]. Cross-experiment analysis is aided when many of the technical aspects have been standardized. This was part of the design of the Songbird NeuroGenomics (SoNG) Initiative's microarray project [15]. WGCNA has been used to identify gene networks associated with particular areas and cell types, disease, and evolution in humans [37–39] and was applied to data from 15 experiments that utilized the SoNG microarray [15]. These experiments represent 488 tissue samples from six bird species assigned to 80 treatment groups. This comprehensive analysis revealed that brain area is the major determinant of overall patterns of gene expression [40]. This result is fascinating as previous use of the WGCNA algorithm on human brain expression data distinguished cell types; the genes associated with different components of the song circuitry may therefore provide essential information to discover what cellular characteristics permit them to function in song [38].

Gene Sets: Whole-Genome Direct RNA Sequencing

Whole-genome "next-generation" direct RNA sequencing technology has at least major advantages over microarrays: (1) no prior cloning needs to be performed; (2) all RNAs expressed in the experimental tissue can be sampled, i.e., measurement of a brain RNA is not dependent upon it also being represented in the sample used to make the array; (3) splice variants can potentially be identified; (4) unannotated sequences can be mapped to genomic locations; and (5) all transcribed genomic elements are sampled, including nonprotein-coding RNAs. Successful use of RNA sequencing technology still requires careful consideration of biological confounds and is most effective with full genome assemblies as read lengths are still relatively short (75–100 bp). Several labs have already taken advantage of this method and are actively analyzing whole-genome direct RNA sequencing data in relation to song. No detailed next-generation sequencing studies on songbirds in the context of song production or processing have been published to date, but data from zebra finch cell lines and auditory forebrain demonstrate that direct RNA sequencing will be a

highly sensitive strategy to detect dynamically regulated sets of genes, multiple types of transcribed genomic elements (see below), and alternatively spliced mRNAs [41]. RNA-sequencing technology is rapidly advancing; with continued improvements in analysis algorithms and falling costs, "next-next-generation" sequencing looks to become a standard method to investigate neurogenomic profiles associated with song in the near future.

Prediction and Testing of Transcriptional Cascades

Transcription factors are proteins that regulate the transcription of other genes. In this way, production of a few genes can set off a cascade of transcriptional activation—or repression—that alters the levels of multiple gene products [32]. Transcription factors often work in combination, but their potential to regulate a particular gene can be predicted by the presence of small stretches of DNA called binding sites/motifs or recognition sequences.

High Conservation of Transcription Factor Genes and Binding Motifs

Transcription factors are subject to evolutionary pressures [42–47]. However, evidence supports the high evolutionary conservation of key transcription factors. For example, ZENK, FOXP2, and steroid receptors are all transcription factors central to song neurobiology and behavior. Each of these transcription factors shows high sequence (78–99 % homology to human) and functional conservation. They show neuroanatomical expression distributions across age and experience that are similar to other species, are regulated by evolutionarily conserved mechanisms, and are implicated in neural processes such as learning and behavior as in other species [26–28, 48–58].

Rapid and Specific Regulation of Transcription

Individual transcription factors such as the immediate early gene ZENK were instrumental in mapping the biological relevance of a variety of specific behaviors and experiences [12, 25, 29, 50, 59]. The rapid and specific expression of transcription factors after hearing or producing song suggests larger alterations in patterns of gene expression function in song processing and production, consistent with microarray findings described above that show hundreds to thousands of transcriptional changes after song.

Transcriptional Cascade Prediction

With the genome assembly, it is possible to predict transcriptional cascades using conserved transcription factor binding motifs. In the zebra finch genome, the majority of transcription factor binding sites are clustered within 10 kb of the 5′ most predicted exon for protein-coding genes [6]. The position of binding motifs can therefore identify genes potentially regulated by a set of transcription factors. It is, however, nontrivial to assign specific transcription factor target genes; it is not universally true that binding sites are within 10 kb upstream of gene models and functional combinations of factors are not easily predicted [6]. Still, with the genome assembly, researchers can start with one transcription factor and computationally construct a multigene transcriptional cascade that can be empirically tested in ways that would have been nearly impossible before.

As proof of principle, predicted transcriptional networks that rely on transcription factor binding site information have been reported [6]. In this case, the starting data were expression data. Genes were grouped based on their temporal pattern of transcriptional changes after a bird sang. Transcription factor binding sites in the genomic regions of these gene sets were identified, and analysis predicted that particular transcription factors may coordinate transcription levels of these gene sets.

Functional Promoter Characteristics

Regions of the genome that are enriched for transcription factor and RNA polymerase binding sites are called promoters. Even small sequence changes within a promoter can have robust functional consequences. Binding site changes can be more cryptic than changes to the protein-coding portions of the gene because the promoter's location in relation to the protein-coding portion of the gene is variable and promoter sequences are not transcribed in RNAs. To date, there are no reports of promoter alterations directly affecting song behavior, but access to regulatory sequences via the genome assembly has great potential to discover gene-brain-behavior connections.

Alternative Promoters

Some genes have more than one promoter. The mechanisms that control usage of one promoter sequence over another are poorly understood but differential binding results in age-, sex-, and tissue-specific expression [60–63]. Evolutionarily conserved alternative promoters for two steroid-related genes, aromatase and androgen receptor, exist in songbirds, too [64, 65]. Alternative promoter usage for the aromatase gene regulates brain transcription independently from peripheral aromatase expression [64]. This may be particularly relevant in the songbird since aromatase

is the enzyme that produces estradiol. Estradiol is the most potent masculinizing factor known for the song control circuitry and can be rapidly synthesized in the brain [66–69]. Androgen receptors are abundant in the song control circuitry and their function is required for estradiol to have its masculinizing effects [70–74]. Thus, alternative promoters may be one mechanism by which signaling and transcriptional cascades are controlled to optimize song.

Noncoding RNA Regulation of mRNA Translation

In contrast to transcription factors that coordinate expression of multiple genes, noncoding RNAs (ncRNAs) typically regulate the availability of mRNAs that are already transcribed. ncRNAs are implicated in brain development, sexual differentiation, and neural plasticity in mammals [75–82]. ncRNAs therefore could affect behavior by acting within song areas at any stage of vocal learning and production. In addition to the more familiar ribosomal RNA (rRNA) and transfer RNA (tRNA) noncoding RNAs that facilitate translation, several other major types have been discovered including microRNA (miRNA), piwi-interacting RNA (piRNA), small nucleolar RNA (snoRNA), and long ncRNA, and the list of functional transcribed genomic elements continues to grow. The different classes of ncRNAs are structurally distinct and act on mRNA in unique ways. All major ncRNA types have been identified in the zebra finch brain, but to date, only miRNAs have been described in relation to song [6, 14, 83, 84].

miRNAs

Mature miRNAs are ~21–25 nucleotides long. They are processed by endogenous cellular machinery that cleaves a larger hairpin structure transcribed from the genome [85]. The guidelines for miRNA regulation of mRNAs are still being refined, but generally, miRNAs bind to similar sequences in the reverse complement orientation in mRNAs [86–89]. Often, only one strand of the miRNA functions and the other strand is degraded, but in some cases both strands act as functional miRNAs. Prior to miRNA-mRNA binding, the miRNA is loaded into a protein complex that either disrupts translation or marks the mRNA for degradation [85–91].

miRNA Is Transcribed in Songbird Brain

In the zebra finch, early reports identified five highly conserved miRNAs expressed in the whole brain throughout development [14]. Later next-generation RNA-sequencing experiments designed to preferentially sequence small RNAs in

the auditory forebrain reported over 100 ncRNAs that share high sequence homology with known miRNAs in mammals and chicken [83, 84]. Additional bioinformatic prediction identified ~35 more sequences that have many characteristics of miRNAs and are putative novel zebra finch miRNAs.

miRNA Is Dynamically Regulated After Specific Stimulus

Song playback experiments demonstrated that miRNAs have the potential to dynamically regulate mRNAs after specific experiences. Approximately 50 miRNAs are expressed at different levels after adult male birds heard song playbacks than when they were left in silence [83]. As with the mRNAs in the same paradigm, some miRNAs are present in higher levels after hearing song and others are reduced. Again, as with the mRNA, the miRNA response is selective for novel conspecific song; hearing playbacks of auditory stimulus with identical acoustic properties but with disordered temporal properties does not invoke the miRNA response in the auditory forebrain [83]. Notably, both strands of one song-responsive miRNA, tgu-mir-2954, appear to be functional, though males preferentially process one strand and females the other (Fig. 5.2) [83].

miRNA Target Prediction

To understand the functional implications of miRNA biosynthesis, it is necessary to identify the mRNAs that are targeted for degradation. Sequence-based predictions identified about 30 target mRNAs for the five miRNAs described in development, and eight mRNAs were predicted as targets of the song-responsive miRNA tgu-miR-2954-3p [14, 83]. Interestingly, all eight mRNAs code for proteins with related function. While intriguing, predictions of mRNA targets need to be experimentally validated. Song-responsive miRNA populations suggest that ncRNAs are sensitive to the broad context of the experiment [83]. Further, prediction algorithms can produce errors given the mismatch nature of the miRNA-mRNA binding. In addition, one gene can have multiple binding sites for a miRNA and can have binding sites for multiple types of miRNAs [6]. This raises the possibility of combinatorial and dosage-based miRNA regulation of a transcript that is difficult to bioinformatically predict at this time.

Repetitive Elements

Repetitive genomic elements serve many essential functions, including chromosomal stability and transcriptional regulation [92–97]. The composition of repetitive elements is different in the songbird genome compared to the human genome:

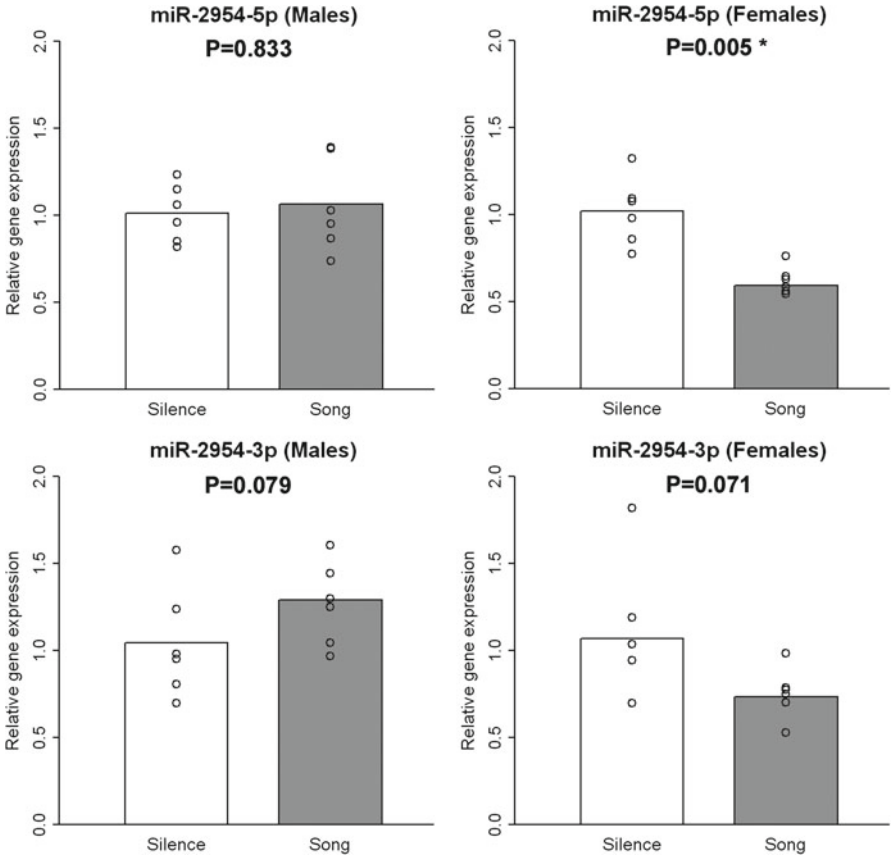


Fig. 5.2 Hearing song results in genome dynamics in the auditory forebrain that include ncRNAs. Each strand of song-responsive miRNA miR-2954 is differentially regulated in male and female auditory forebrain. Graphs show RT-PCR data of the relative expression of each strand (3p or 5p) after either silence (*white bar*) or song playback (*gray bar*); *open circles* mark individual measures. *Reused with permission from Gunaratne, P. H. et al. Song exposure regulates known and novel microRNAs in the zebra finch auditory forebrain. BMC Genomics 12, 277, doi:10.1186/1471-2164-12-277 (2011) Open Source Article* [83]

~50 % of the human genome is characterized as repetitive elements but in the zebra finch genome, this estimation is ~8 % [6]. Major classes of repetitive elements, long interspersed elements (LINES), short interspersed elements (SINES), long terminal repeats (LTRs), and microsatellites are, however, present in the zebra finch genome [6]. Interestingly, SINES are essentially absent in the chicken genome, suggesting one way that the songbird genome may differ functionally from the non-songbird genome [6, 98]. The role of repetitive elements in genome function is under active research in all models, and there are likely implications for song: at least one repetitive element is expressed in the brain and is enriched in song control areas [6].

DNA Structure and Song

Epigenetic DNA Changes

Epigenetic modifications are small molecular changes that regulate transcription without alterations in genomic sequence. DNA methylation in promoters prevents gene transcription, and posttranslational modifications to histones, the structural proteins of chromatin, modify the probability of gene transcription. Epigenetics is especially intriguing when looking for genome-behavior interconnections because experience alters the epigenomic landscape, and the epigenetic "marks" are reversible [99–101].

To date, there are no published papers that tie specific epigenetic modifications to song. But there is an unpublished report that expression of epigenetic-related genes such as histone H3 are expressed at a higher level in song control area when birds sing plastic song compared to stable song (Kobayashi et al., Society for Neuroscience Abstracts No. 413.09, 2011, personal communication). In zebra finches, tutor experience and song rehearsal during the sensitive period for song learning may promote epigenetic marks that prevent song acquisition later in life. In other species that can learn song more than one time, mechanisms that erase previous marks may also be crucial for reopening behavioral plasticity.

Large-Scale Chromosome Rearrangements

Relatively large structural genomic changes also occur. Insertions, deletions, inversions, and duplications or expansions can alter single genes and swathes of genes in more than one chromosome. These rearrangements are positively correlated with the position of repetitive elements discussed above, and several instances of chromosomal rearrangements were described in the zebra finch genome [6].

Deletion

Comparison of the zebra finch and human genomes revealed that synapsin I (SYN1) is not present in the zebra finch assembly [6]. Further, the adjacent genes, covering at least 100 Mb, are also apparently absent from the zebra finch genome. The functional effect of this deletion remains to be discovered. SYN1 is one of the most abundant proteins found in the presynaptic terminal and is instrumental in modulating release of synaptic vesicles, thus could have profound effects on neural signaling [102]. However, the chicken genome also lacks this stretch of DNA, so it may be that birds have evolved alternative or paralogous genes to perform the same roles as the deleted genes and/or that these genes are not necessary for any aspect of song.

Duplication and Expansion

Three genes that may have relevance to the development and function of song brain areas may be duplicated in the zebra finch genome based on comparisons to chicken and mammalian genomes: growth hormone, caspase-3, and β-secretase [6, 103, 104]. Growth hormone could affect the size and connectivity of neural circuits; other passerines species have a similar duplication [104, 105]. Zebra finches appear to have a unique duplication in the caspase-3 gene. Caspase-3 is activated in the auditory forebrain of adult male zebra finches within 30 min of hearing novel conspecific song playbacks [106]. Caspase-3 activation in mammals initiates apoptosis but there was no evidence of cell death after song playbacks. Perhaps the different gene forms permit the uncoupling of caspase-3 and programmed cell death; the caspase family of genes differs across species with duplicated and deleted genes in zebra finches, chicken, and mammals, and it has been suggested that zebra finches may therefore have a different caspase-mediated apoptosis cascade than other animals [103]. β-secretase is best understood for its role in cleaving the proteins that aggregate in Alzheimer's disease [107]. There are no reports that songbirds suffer from brain abnormalities similar to the Alzheimer's plaques, and how either version of the β-secretase gene might influence songbird brain function is not immediately clear [108]. All of these in silico predictions need to be carefully validated as having biological relevance.

Two protein kinase genes, p21-activated serine/threonine kinase 3 (PAK3) and proviral integration site 1 (PIM1), were predicted to have undergone expansions [109, 110]. PAK3 was predicted to have 31 separate zebra finch genes. At least two genes are expressed in the brain, including in major song control areas. Each gene has a distinct neuroanatomical expression distribution, consistent with their separate identities and potential functions. PIM1 has ten additional family members in the zebra finch genome. At least one of the PIM1 expanded genes is expressed in the brain. Both PAK3 and PIM1 appear as only one gene in the chicken genome; thus, it is possible that these are Passerine-specific expansions that have effects within and outside of the brain; as with other predicted genomic traits, these gene expansions need to be experimentally validated.

Inversion

The zebra finch Z sex chromosome (male birds are ZZ and female birds are ZW) is polymorphic due to a large-scale inversion [111] (Fig. 5.3). A similar chromosomal rearrangement is not found on the chicken sex chromosomes. Further, the zebra finch polymorphism is present in varying ratios in distinct populations, including those in the species' native Australia. The zebra finch inversion visibly alters the position of the centromere. It affects at least three protein-coding genes: histidine triad nucleotide-binding protein 1 (HINT1), doublesex- and mab-3-related

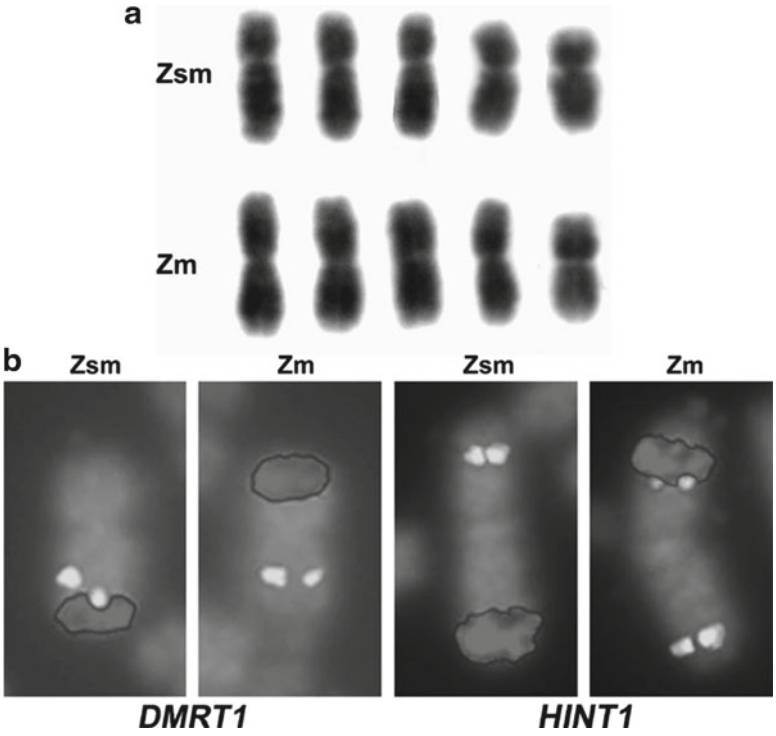


Fig. 5.3 Sex chromosome genes may contribute to masculinization of song circuitry; inversion events could therefore affect song by altering chromosomal structure. (**a**) Two types of Z chromosomes found in wild and domesticated zebra finches show alterations in centromere position and (**b**) gene locations such as for DMRT1 and HINT1. In (**b**) *medium-gray hazy label* is DAPI nonselective DNA stain, outlined *dark-gray* label shows the late replication site (here, provides spatial reference), and bright dots are fluorescent in situ hybridization (FISH) signal for specific genes. *Reused with permission from Itoh, Y., Kampf, K., Balakrishnan, C. N. & Arnold, A. P. Karyotypic polymorphism of the zebra finch Z chromosome. Chromosoma 120, 255–264 (2011)* [110]

transcription factor 1 (DMRT1), and PAK3. In fact, 11 of the 31 reported PAK3 gene replications were mapped to the Z chromosome; one gene has been validated [109, 111]. Like PAK3, HINT1 is associated with protein kinase function and is implicated in human psychiatric disorders. DMRT1 is a transcription factor necessary for sex determination. Although the functional implications of this Z inversion chromosome are unknown, chromosomal inversions in other birds have a profound effect on behavioral phenotype [112]. The position of this inversion on a sex chromosome may be particularly relevant to song because Z-linked genes are not subject to dosage compensation silencing [113, 114].

Comparative Approaches

Finally, as a model for speech and language, songbirds benefit from a rich natural phylogeny [8, 115, 116]. The focus of this chapter was the zebra finch, but zebra finches are only one of almost 5,000 extant songbird species that can be used as comparisons. Songbirds are also closely related to a group of sub-oscine birds that do not learn their vocalizations. Comparisons of the chicken and zebra finch genomes have already uncovered ion channels as a potential set of song-responsive genes that may be under different evolutionary pressures in songbirds [6, 117]. Other bird genome sequences, assemblies, and transcriptomes are in the pipeline and represent more power for songbirds to contribute to our understanding of complex behaviors such as learned vocalizations.

Current Challenges and Opportunities

We are just beginning to uncover how genomic dynamics track with the dynamic brain during the learning, perceptual processing, and production of song. There are challenges and opportunities for using genome assemblies to advance investigations into learned vocal communication [118]. Most of these issues are not specific to the songbird, but represent the scientific revolution that all systems must undergo. The songbird community of researchers has an exemplary track record of using this natural system to identify fundamental mechanisms of brain-behavior interconnections. The genome assembly is now one more resource to be employed in these studies. To make best use of the genome assemblies, it is necessary to merge genomic investigations with existing strengths in behavioral, molecular, and physiological methodologies that have proven so successful in making the songbird a powerful model in which to discover neural mechanisms that underlie complex behaviors such as speech and language.

References

1. Margoliash D (2010) Sleep, learning, and birdsong. ILAR J 51:378–386
2. Riters LV (2011) Pleasure seeking and birdsong. Neurosci Biobehav Rev 35:1837–1845. doi:10.1016/j.neubiorev.2010.12.017
3. Pinaud R, Terleph TA, Wynne RD, Tremere LA (2008) Neural correlates of auditory processing, learning and memory formation in songbirds. Prog Theor Phys Supp 173:270–282
4. Sizemore M, Perkel DJ (2011) Premotor synaptic plasticity limited to the critical period for song learning. Proc Natl Acad Sci U S A 108:17492–17497. doi:10.1073/pnas.1104255108
5. Prather JF, Peters S, Nowicki S, Mooney R (2008) Precise auditory-vocal mirroring in neurons for learned vocal communication. Nature 451:305–310. doi:10.1038/nature06492
6. Warren WC, Clayton DF, Ellegren H, Arnold AP, Hillier LW, Künstner A et al (2010) The genome of a songbird. Nature 464:757–762. doi:10.1038/nature08819
7. Clayton DF (1997) Role of gene regulation in song circuit development and song learning. J Neurobiol 33:549–571

8. Clayton DF, Balakrishnan CN, London SE (2009) Integrating genomes, brain and behavior in the study of songbirds. Curr Biol 19:R865–R873. doi:10.1016/j.cub.2009.07.006
9. Jarvis ED, Nottebohm F (1997) Motor-driven gene expression. Proc Natl Acad Sci U S A 94:4097–4102
10. London SE, Clayton DF (2010) The neurobiology of Zebra Finch song: insights from gene expression studies. Emu-Austral Ornithology 110:219–232. doi:10.1071/Mu09079
11. London SE, Remage-Healey L, Schlinger BA (2009) Neurosteroid production in the songbird brain: a re-evaluation of core principles. Front Neuroendocrinol 30:302–314. doi:10.1016/j.yfrne.2009.05.001
12. Mello CV, Velho TA, Pinaud R (2004) Song-induced gene expression: a window on song auditory processing and perception. Ann N Y Acad Sci 1016:263–281. doi:10.1196/annals.1298.021
13. Wade J, Arnold AP (2004) Sexual differentiation of the zebra finch song system. Ann N Y Acad Sci 1016:540–559. doi:10.1196/annals.1298.015
14. Li X, Wang XJ, Tannenhauser J, Podell S, Mukherjee P, Hertel M et al (2007) Genomic resources for songbird research and their use in characterizing gene expression during brain development. Proc Natl Acad Sci U S A 104:6834–6839. doi:10.1073/pnas.0701619104
15. Replogle K, Arnold AP, Ball GF, Band M, Bensch S, Brenowitz EA et al (2008) The songbird neurogenomics (SoNG) initiative: community-based tools and strategies for study of brain gene function and evolution. BMC Genomics 9:131. doi:10.1186/1471-2164-9-131
16. Wada K, Howard JT, McConnell P, Whitney O, Lints T, Rivas MV et al (2006) A molecular neuroethological approach for identifying and characterizing a cascade of behaviorally regulated genes. Proc Natl Acad Sci U S A 103:15212–15217. doi:10.1073/pnas.0607098103
17. Wade J, Peabody C, Coussens P, Tempelman RJ, Clayton DF, Liu L et al (2004) A cDNA microarray from the telencephalon of juvenile male and female zebra finches. J Neurosci Methods 138:199–206. doi:10.1016/j.jneumeth.2004.04.007
18. Arnold AP (1997) Sexual differentiation of the zebra finch song system: positive evidence, negative evidence, null hypotheses, and a paradigm shift. J Neurobiol 33:572–584
19. Arnold AP, Itoh Y (2011) Factors causing sex differences in birds. Avian Biol Res 4:44–51. doi:10.3184/175815511x13070045977959
20. Remage-Healey L, London SE, Schlinger BA (2010) Birdsong and the neural production of steroids. J Chem Neuroanat 39:72–81. doi:10.1016/j.jchemneu.2009.06.009
21. Schlinger BA, London SE (2006) Neurosteroids and the songbird model system. J Exp Zool 305:743–748. doi:10.1002/jez.a.303
22. Schlinger BA, Soma KK, London SE (2001) Neurosteroids and brain sexual differentiation. Trends Neurosci 24:429–431
23. Wade J, Tang YP, Peabody C, Tempelman RJ (2005) Enhanced gene expression in the forebrain of hatchling and juvenile male zebra finches. J Neurobiol 64:224–238. doi:10.1002/neu.20141
24. Lovell PV, Clayton DF, Replogle KL, Mello CV (2008) Birdsong "transcriptomics": neurochemical specializations of the oscine song system. PLoS One 3:e3440. doi:10.1371/journal.pone.0003440
25. Jarvis ED, Scharff C, Grossman MR, Ramos JA, Nottebohm F (1998) For whom the bird sings: context-dependent gene expression. Neuron 21:775–788
26. Mello C, Nottebohm F, Clayton D (1995) Repeated exposure to One song leads to a rapid and persistent decline in an immediate-early genes response to that song in Zebra finch telencephalon. J Neurosci 15:6919–6925
27. Mello CV, Clayton DF (1994) Song-induced ZENK gene expression in auditory pathways of songbird brain and its relation to the song control system. J Neurosci 14:6652–6666
28. Mello CV, Vicario DS, Clayton DF (1992) Song presentation induces gene expression in the songbird forebrain. Proc Natl Acad Sci U S A 89:6818–6822
29. Woolley SC, Doupe AJ (2008) Social context-induced song variation affects female behavior and gene expression. PLoS Biol 6:e62. doi:10.1371/journal.pbio.0060062
30. Hilliard A, Miller J, Fraley ER, Horvath S, White S (2012) Molecular microcircuitry underlies functional specification in a basal ganglia circuit dedicated to vocal learning. Neuron 73:537–552

31. Hilliard AT, Miller JE, Horvath S, White SA (2012) Distinct neurogenomic states in basal ganglia subregions relate differently to singing behavior in songbirds. PLoS Comput Biol 8(11):e1002773. doi:10.1371/journal.pcbi.1002773
32. Clayton DF (2000) The genomic action potential. Neurobiol Learn Mem 74:185–216
33. Dong S, Replogle KL, Hasadsri L, Imai BS, Yau PM, Rodriguez-Zas S et al (2009) Discrete molecular states in the brain accompany changing responses to a vocal signal. Proc Natl Acad Sci U S A 106:11364–11369. doi:10.1073/pnas.0812998106
34. Dong S, Clayton DF (2009) Habituation in songbirds. Neurobiol Learn Mem 92:183–188. doi:10.1016/j.nlm.2008.09.009
35. London SE, Dong S, Replogle K, Clayton DF (2009) Developmental shifts in gene expression in the auditory forebrain during the sensitive period for song learning. Dev Neurobiol 69:437–450. doi:10.1002/dneu.20719
36. Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9:559. doi:10.1186/1471-2105-9-559
37. Oldham MC, Horvath S, Geschwind DH (2006) Conservation and evolution of gene coexpression networks in human and chimpanzee brains. Proc Natl Acad Sci U S A 103:17973–17978. doi:10.1073/pnas.0605938103
38. Oldham MC, Konopka G, Iwamoto K, Langfelder P, Kato T, Horvath S et al (2008) Functional organization of the transcriptome in human brain. Nat Neurosci 11:1271–1282. doi:10.1038/nn.2207
39. Miller JA, Oldham MC, Geschwind DH (2008) A systems level analysis of transcriptional changes in Alzheimer's disease and normal aging. J Neurosci 28:1410–1420. doi:10.1523/JNEUROSCI.4098-07.2008
40. Drnevich J, Replogle KL, Lovell P, Hahn TP, Johnson F, Mast TG et al (2012) Impact of experience-dependent and -independent factors on gene expression in songbird brain. Proc Natl Acad Sci U S A 109(Suppl):17245–17252. doi:10.1073/pnas.1200655109
41. Balakrishnan CN, Lin YC, London SE, Clayton DF (2012) RNA-seq transcriptome analysis of male and female zebra finch cell lines. Genomics 100:363–369. doi:10.1016/j.ygeno.2012.08.002
42. Gilad Y, Oshlack A, Smyth GK, Speed TP, White KP (2006) Expression profiling in primates reveals a rapid evolution of human transcription factors. Nature 440:242–245. doi:10.1038/nature04559
43. Prendergast JG, Semple CA (2011) Widespread signatures of recent selection linked to nucleosome positioning in the human lineage. Genome Res 21:1777–1787. doi:10.1101/gr.122275.111
44. Zheng W, Gianoulis TA, Karczewski KJ, Zhao H, Snyder M (2011) Regulatory variation within and between species. Annu Rev Genomics Hum Genet 12:327–346. doi:10.1146/annurev-genom-082908-150139
45. Sholtis SJ, Noonan JP (2010) Gene regulation and the origins of human biological uniqueness. Trends Genet 26:110–118. doi:10.1016/j.tig.2009.12.009
46. Hamilton AT, Huntley S, Kim J, Branscomb E, Stubbs L (2003) Lineage-specific expansion of KRAB zinc-finger transcription factor genes: implications for the evolution of vertebrate regulatory networks. Cold Spring Harb Symp Quant Biol 68:131–140
47. Nowick K, Hamilton AT, Zhang H, Stubbs L (2010) Rapid sequence and expression divergence suggest selection for novel function in primate-specific KRAB-ZNF genes. Mol Biol Evol 27:2606–2617. doi:10.1093/molbev/msq157
48. Cheng HY, Clayton DF (2004) Activation and habituation of extracellular signal-regulated kinase phosphorylation in zebra finch auditory forebrain during song presentation. J Neurosci 24:7503–7513. doi:10.1523/JNEUROSCI.1405-04.2004
49. Dong S, Clayton DF (2008) Partial dissociation of molecular and behavioral measures of song habituation in adult zebra finches. Genes Brain Behav 7:802–809. doi:10.1111/j.1601-183X.2008.00423.x
50. Jin H, Clayton DF (1997) Localized changes in immediate-early gene regulation during sensory and motor learning in zebra finches. Neuron 19:1049–1059

51. London SE, Clayton DF (2008) Functional identification of sensory mechanisms required for developmental song learning. Nat Neurosci 11:579–586. doi:10.1038/nn.2103
52. 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:e321. doi:10.1371/journal.pbio.0050321
53. Haesler S, Wada K, Nshdejan A, Morrisey EE, Lints T, Jarvis ED et al (2004) FoxP2 expression in avian vocal learners and non-learners. J Neurosci 24:3164–3175. doi:10.1523/JNEUROSCI.4369-03.2004
54. Miller JE, Spiteri E, Condro MC, Dosumu-Johnson RT, Geschwind DH, White SA (2008) Birdsong decreases protein levels of FoxP2, a molecule required for human speech. J Neurophysiol 100:2015–2025. doi:10.1152/jn.90415.2008
55. Schulz SB, Haesler S, Scharff C, Rochefort C (2010) Knockdown of FoxP2 alters spine density in Area X of the zebra finch. Genes Brain Behav 9:732–740. doi:10.1111/j.1601-183X.2010.00607.x
56. 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 24:3152–3163. doi:10.1523/JNEUROSCI.5589-03.2004
57. Teramitsu I, Poopatanapong A, Torrisi S, White SA (2010) Striatal FoxP2 is actively regulated during songbird sensorimotor learning. PLoS One 5:e8548. doi:10.1371/journal.pone.0008548
58. Teramitsu I, White SA (2006) FoxP2 regulation during undirected singing in adult songbirds. J Neurosci 26:7390–7394. doi:10.1523/JNEUROSCI.1662-06.2006
59. Bailey DJ, Wade J (2003) Differential expression of the immediate early genes FOS and ZENK following auditory stimulation in the juvenile male and female zebra finch. Brain Res 116:147–154
60. Stocco C (2008) Aromatase expression in the ovary: hormonal and molecular regulation. Steroids 73:473–487. doi:10.1016/j.steroids.2008.01.017
61. Cao-Lei L, Leija SC, Kumsta R, Wüst S, Meyer J, Turner JD et al (2011) Transcriptional control of the human glucocorticoid receptor: identification and analysis of alternative promoter regions. Hum Genet 129:533–543
62. Lenartowski R, Goc A (2011) Epigenetic, transcriptional and posttranscriptional regulation of the tyrosine hydroxylase gene. Int J Dev Neurosci 29:873–883
63. Davuluri RV, Suzuki Y, Sugano S, Plass C, Huang THM (2008) The functional consequences of alternative promoter use in mammalian genomes. Trends Genet 24:167–177
64. Ramachandran B, Schlinger BA, Arnold AP, Campagnoni AT (1999) Zebra finch aromatase gene expression is regulated in the brain through an alternate promoter. Gene 240:209–216
65. Luo M, Yu Y, Kim H, Kudrna D, Itoh Y, Agate RJ et al (2006) Utilization of a zebra finch BAC library to determine the structure of an avian androgen receptor genomic region. Genomics 87:181–190. doi:10.1016/j.ygeno.2005.09.005
66. 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 25:865–877. doi:10.1002/neu.480250710
67. Remage-Healey L, Maidment NT, Schlinger BA (2008) Forebrain steroid levels fluctuate rapidly during social interactions. Nat Neurosci 11:1327–1334. doi:10.1038/nn.2200
68. Schlinger BA, Arnold AP (1992) Circulating estrogens in a male songbird originate in the brain. Proc Natl Acad Sci U S A 89:7650–7653
69. Grisham W, Arnold AP (1995) A direct comparison of the masculinizing effects of testosterone, androstenedione, estrogen, and progesterone on the development of the zebra finch song system. J Neurobiol 26:163–170. doi:10.1002/neu.480260202
70. 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
71. 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 19:2628–2636
72. Grisham W, Lee J, McCormick ME, Yang-Stayner K, Arnold AP (2002) Antiandrogen blocks estrogen-induced masculinization of the song system in female zebra finches. J Neurobiol 51:1–8

73. Harding CF, Walters MJ, Parsons B (1984) Androgen receptor levels in hypothalamic and vocal control nuclei in the male zebra finch. Brain Res 306:333–339
74. Kim YH, Perlman WR, Arnold AP (2004) Expression of androgen receptor mRNA in zebra finch song system: developmental regulation by estrogen. J Comp Neurol 469:535–547. doi:10.1002/cne.11033
75. Bian S, Sun T (2011) Functions of noncoding RNAs in neural development and neurological diseases. Mol Neurobiol 44:1–15
76. Fineberg SK, Kosik KS, Davidson BL (2009) MicroRNAs potentiate neural development. Neuron 64:303–309
77. Schratt G (2009) Fine-tuning neural gene expression with microRNAs. Curr Opin Neurobiol 19:213–219
78. St Laurent Iii G, Wahlestedt C (2007) Noncoding RNAs: couplers of analog and digital information in nervous system function? Trends Neurosci 30:612–621
79. Konopka W, Kiryk A, Novak M, Herwerth M, Parkitna JR, Wawrzyniak M et al (2010) MicroRNA loss enhances learning and memory in mice. J Neurosci 30:14835–14842
80. Li Q, Bian S, Hong J, Kawase-Koga Y, Zhu E, Zheng Y et al (2011) Timing specific requirement of microRNA function is essential for embryonic and postnatal hippocampal development. PLoS One 6(10):e26000
81. Mercer TR, Dinger ME, Mariani J, Kosik KS, Mehler MF, Mattick JS et al (2008) Noncoding RNAs in long-term memory formation. Neuroscientist 14:434–445
82. Takada S, Berezikov E, Young LC, Yamashita Y, Mano H (2009) Potential role of miR-29b in modulation of Dnmt3a and Dnmt3b expression in primordial germ cells of female mouse embryos. RNA 15:1507–1514
83. Gunaratne PH, Lin YC, Benham AL, Drnevich J, Coarfa C, Tennakoon JB et al (2011) Song exposure regulates known and novel microRNAs in the zebra finch auditory forebrain. BMC Genomics 12:277. doi:10.1186/1471-2164-12-277
84. Luo GZ, Hafner M, Shi Z, Brown M, Feng GH, Tuschl T et al (2012) Genome-wide annotation and analysis of zebra finch microRNA repertoire reveal sex-biased expression. BMC Genomics 13:727
85. He L, Hannon GJ (2004) MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet 5:522–531. doi:10.1038/nrg1379
86. Thomson DW, Bracken CP, Goodall GJ (2011) Experimental strategies for microRNA target identification. Nucleic Acids Res 39:845–6853
87. Nelson PT, Kiriakidou M, Mourelatos Z, Tan GS, Jennings MH, Xie K et al (2010) High-throughput experimental studies to identify miRNA targets directly, with special focus on the mammalian brain. Brain Res 1338:122–130
88. Smalheiser NR, Torvik VI (2006) Complications in mammalian microRNA target prediction. Methods Mol Biol 342:115–127
89. Szabó PM (2010) MicroRNA target prediction: problems and possible solutions. Curr Bioinformatics 5:81–88
90. Eulalio A, Huntzinger E, Izaurralde E (2008) Getting to the root of miRNA-mediated gene silencing. Cell 132:9–14
91. Huntzinger E, Izaurralde E (2011) Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat Rev Genet 12:99–110
92. Cohen CJ, Lock WM, Mager DL (2009) Endogenous retroviral LTRs as promoters for human genes: a critical assessment. Gene 448:105–114. doi:10.1016/j.gene.2009.06.020
93. Abeysinghe SS, Chuzhanova N, Cooper DN (2006) Gross deletions and translocations in human genetic disease. Genome Dyn 1:17–34
94. Jensen-Seaman MI, Furey TS, Payseur BA, Lu Y, Roskin KM, Chen CF et al (2004) Comparative recombination rates in the rat, mouse, and human genomes. Genome Res 14: 528–538
95. Moltó E, Fernández A, Montoliu L (2008) Boundaries in vertebrate genomes: different solutions to adequately insulate gene expression domains. Brief Funct Genomic Proteomic 8:283–296

96. Waldman AS (2008) Ensuring the fidelity of recombination in mammalian chromosomes. Bioessays 30:1163–1171
97. Xing J, Witherspoon DJ, Ray DA, Batzer MA, Jorde LB (2007) Mobile DNA elements in primate and human evolution. Am J Phys Anthropol 50:2–9
98. Sela N, Kim E, Ast G (2010) The role of transposable elements in the evolution of non-mammalian vertebrates and invertebrates. Genome Biol 11:R59. doi:10.1186/gb-2010-11-6-r59
99. Gudsnuk KMA, Champagne FA (2011) Epigenetic effects of early developmental experiences. Clin Perinatol 38:703–717
100. Gräff J, Mansuy IM (2009) Epigenetic dysregulation in cognitive disorders. Eur J Neurosci 30:1–8
101. McGowan PO, Szyf M (2010) The epigenetics of social adversity in early life: implications for mental health outcomes. Neurobiol Dis 39:66–72
102. Shupliakov O, Haucke V, Pechstein A (2011) How synapsin I may cluster synaptic vesicles. Semin Cell Dev Biol 22:393–399
103. Quesada V, Velasco G, Puente XS, Warren WC, Lopez-Otin C (2010) Comparative genomic analysis of the zebra finch degradome provides new insights into evolution of proteases in birds and mammals. BMC Genomics 11:220. doi:10.1186/1471-2164-11-220
104. Yuri T, Kimball RT, Braun EL, Braun MJ (2008) Duplication of accelerated evolution and growth hormone gene in passerine birds. Mol Biol Evol 25:352–361. doi:10.1093/molbev/msm260
105. Arai N, Iigo M (2010) Duplicated growth hormone genes in a passerine bird, the jungle crow (Corvus macrorhynchos). Biochem Biophys Res Commun 397:553–558. doi:10.1016/j.bbrc.2010.05.156
106. Huesmann GR, Clayton DF (2006) Dynamic role of postsynaptic caspase-3 and BIRC4 in zebra finch song-response habituation. Neuron 52:1061–1072. doi:10.1016/j.neuron.2006.10.033
107. Citron M (2001) Human beta-secretase and Alzheimer's disease. Expert Opin Ther Targets 5:341–348. doi:10.1517/14728222.5.3.341
108. 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
109. Kong L, Lovell PV, Heger A, Mello CV, Ponting CP (2010) Accelerated evolution of PAK3- and PIM1-like kinase gene families in the zebra finch, Taeniopygia guttata. Mol Biol Evol 27:1923–1934. doi:10.1093/molbev/msq080
110. Itoh Y, Kampf K, Balakrishnan CN, Arnold AP (2011) Karyotypic polymorphism of the zebra finch Z chromosome. Chromosoma 120:255–264
111. Itoh Y, Kampf K, Balakrishnan CN, Arnold AP (2011) Karyotypic polymorphism of the zebra finch Z chromosome. Chromosoma 120:255–264. doi:10.1007/s00412-010-0308-3
112. Thomas JW, Cáceres M, Lowman JJ, Morehouse CB, Short ME, Baldwin EL et al (2008) The chromosomal polymorphism linked to variation in social behavior in the white-throated sparrow (Zonotrichia albicollis) is a complex rearrangement and suppressor of recombination. Genetics 179:1455–1468. doi:10.1534/genetics.108.088229
113. Itoh Y, Melamed E, Yang X, Kampf K, Wang S, Yehya N et al (2007) Dosage compensation is less effective in birds than in mammals. J Biol 6:2. doi:10.1186/jbiol53
114. Melamed E, Arnold AP (2007) Regional differences in dosage compensation on the chicken Z chromosome. Genome Biol 8:R202. doi:10.1186/gb-2007-8-9-r202
115. Barker FK, Cibois A, Schikler P, Feinstein J, Cracraft J (2004) Phylogeny and diversification of the largest avian radiation. Proc Natl Acad Sci U S A 101:11040–11045. doi:10.1073/pnas.0401892101
116. Hackett SJ, Kimball RT, Reddy S, Bowie RC, Braun EL, Braun MJ et al (2008) A phylogenomic study of birds reveals their evolutionary history. Science 320:1763–1768. doi:10.1126/science.1157704
117. Nam K, Mugal C, Nabholz B, Schielzeth H, Wolf JB, Backström N et al (2010) Molecular evolution of genes in avian genomes. Genome Biol 11:R68. doi:10.1186/gb-2010-11-6-r68
118. Clayton DF (2004) Songbird genomics: methods, mechanisms, opportunities, and pitfalls. Ann N Y Acad Sci 1016:45–60. doi:10.1196/annals.1298.028

Chapter 6
The Molecular Convergence of Birdsong and Speech

Mugdha Deshpande and Thierry J. Lints

Abstract Songbird vocal learning depends on the anterior forebrain pathway, the organization of which reflects a conserved vertebrate cortico-basal ganglia-thalamocortical loop architecture. We review the involvement of *FoxP2* in this circuit, as well as *FoxP1* and *Cntnap2*, both posited to participate alongside *FoxP2*. In the avian striatum, *FoxP2* expression is regulated by singing, highlighting the possibility that developmental verbal dyspraxia arising from human *FOXP2* mutation might primarily reflect a deficit in ongoing neural signaling, rather than developmental miswiring. We explore genes co-regulated with *FoxP2* during singing and propose that Wnt trafficking and p63 signaling pathways may be crucial to speech and language.

Keywords Birdsong • Speech • *FOXP2* • *WNT* • Exosomes • Multivesicular body • Apical ectodermal ridge • Autism • Angelman • Potocki–Lupski • Williams and Phelan–McDermid syndromes

Introduction

The Comparative Approach to the Molecular Biology of Speech and Language

Our imitativeness and capacity for vocal learning form the bedrock upon which modern civilization rests. Yet, we still know very little about how our brain and genome together accomplish the unique feat of speech and language. Tackling this

M. Deshpande, Ph.D. • T.J. Lints, M.Sc., Ph.D. (✉)
Department of Biology, Texas A & M University, College Station, TX, USA
e-mail: tlints@bio.tamu.edu

S.A. Helekar (ed.), *Animal Models of Speech and Language Disorders*,
DOI 10.1007/978-1-4614-8400-4_6,

problem is somewhat simplified by adopting an appropriate model system with which to examine core features of the vocal learning process. In keeping with the August Krogh Principle, that for many problems there is an animal on which it can be most conveniently studied [1, 2], one model has emerged as preeminent in the study of vocal learning. The zebra finch (*Taeniopygia guttata*), a cage domesticated songbird, has been intensively investigated by ethologists and neuroscientists, dating back to a seminal study by Desmond Morris [3]. Historically, the zebra finch model has made significant contributions to understanding adult neurogenesis [4–6], sensorimotor learning [7–10], the role of sleep in learning [11, 12], brain sexual dimorphism [13, 14], mechanisms of sexual selection [15, 16], group affiliative behavior [17], and vocal behavior [18–21]. Thus, a wealth of electrophysiological, neuroanatomical, ecological, and behavioral data exists for this songbird.

The zebra finch is also the most facile animal model of the few species, beside *Homo sapiens*, both capable of vocal learning and for which whole genome sequence data exists [22]. This bird might therefore seem an ideal model for understanding how genes affect vocal learning, speech, and language. Crucially, there remain some significant obstacles to progress on the molecular and genetic analysis of vocal learning in songbirds. With very few exceptions (and setting aside focal virus-mediated manipulations of songbird gene expression [23]), there have been no forward or reverse genetic experiments conducted to interrogate aspects of songbird biology. In an exciting development, zebra finch transgenesis has been demonstrated, representing a major technical advance [24]. There is also a paucity of natural songbird mutants affected in vocal learning, and comprehensive forward genetic screening for such mutants does not seem a particularly viable approach. Currently, for the zebra finch, there are some moderately inbred laboratory populations [25], but no isogenic strains. Indeed, the "best"-characterized mutation affecting songbird vocal communication is in Belgian Waterslager canaries, which have a hereditary degenerative sensorineural hearing loss at higher frequencies and, compensatorily, a loud, low-pitched song—for which it was bred [26–28]. The gene, or genes, responsible remains at large. Finally, strong inbreeding depression exhibited by the zebra finch [29, 30], the presence of a large number of poorly characterized microchromosomes, a genomic landscape featuring extremely heterogeneous recombination rates [31], and the very few laboratories engaged in such work conspire to present a significant challenge in applying traditional genetic approaches to the functional analysis of genes involved in birdsong.

Not surprisingly, the molecular and genetic dissection of birdsong is still in its infancy, as indeed is the study of genes contributing to speech and language function. Correspondingly, at the level of the gene, there are at present only a few tantalizing points of intersection between song learning and human speech and language acquisition. Over the next decade there will be a dramatic increase in the number of vocal learner species for which whole genome sequence data exists. In addition to the human and zebra finch, substantial genome sequence data is already available for other vocal learner groups, including parrots [32], bats [33], and dolphins [34, 35]. In time, it may be possible to identify some of the crucial genetic changes that underlay the convergent evolution of this trait. The emerging technological

armamentarium, reviewed in Chap. 5 by London, will pave the way for a bioinformatic revolution in our understanding of the genetic basis for vocal learning. For the moment, the present chapter provides a timely opportunity to take stock of the first decade of investigations into the contribution of "speech and language genes" to vocal learning in songbirds.

Inevitably, the function of the avian *FoxP2* gene in songbird vocal learning has been central to many of these studies and, as is hoped for speech and language, may provide a fulcrum for prizing apart the genetic control of learned vocalization. The many reviews in the literature on the function of *FoxP2* in birdsong, speech, and language are a testament to the excitement in this nascent field [36, 37]. The first part of this chapter therefore covers some well-worn ground to summarize the findings on avian *FoxP2* function, elaborating on songbird studies briefly touched on in Chap. 2 by Vernes and Fisher. Before doing so, however, we first describe the contribution to song learning of the anterior forebrain pathway (AFP) of the song system, in order to place the songbird *FoxP2* data in context. In this respect, the present chapter complements Chap. 4 by Woolley on the songbird auditory system and Chap. 3 by Tchernichovski and Margoliash, which focuses primarily on timing mechanisms mediated through the vocal motor pathway.

In the second half of this chapter, we venture into less certain territory, speculating on the involvement of other genes in birdsong and speech production. Although some of the connections drawn must be regarded as tentative, by bringing together songbird data and disparate human genetics findings, we advance a hypothesis that may provide a path through the seemingly impenetrable genetic complexity of disorders affecting speech production, including childhood apraxia of speech (CAS) (developmental verbal dyspraxia) and Angelman, Potocki–Lupski, Williams and Phelan–McDermid syndromes. Thus, the framework developed here illustrates a tremendous synergy that exists between examining gene expression during song and what is, effectively, a genetic screen of many millions of human genomes for deficits in speech and language.

AFP Architecture Is Analogous to Cortico-Basal Ganglia-Thalamocortical Loops

The primary circuit in the songbird brain controlling the production of song is known as the vocal motor pathway (see Fig. 3.1a). Pioneering work by Nottebohm and colleagues showed that bilateral lesions of the premotor cortical nucleus HVC in male canaries completely abolished song production. Similarly, lesions in the motor nucleus RA caused severe deficits in song [38]. Since these initial observations, a detailed understanding of the circuits underlying neuronal control of song has emerged through recording techniques that allowed chronic recordings of identified neurons in awake, singing birds. These experiments indicate that the timing of bursts within HVC microcircuitry is responsible for keeping the clock for song timing [9, 39] although, as discussed in Chap. 3, this view is challenged by recent findings.

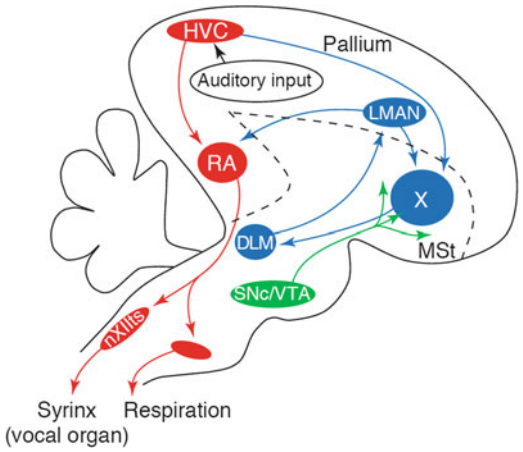


Fig. 6.1 Simplified layout of the song system of oscine songbirds. Vocal control is mediated by two major pathways. Shown in *red*, the motor pathway for song production descends from nucleus HVC (which is also the main station of auditory input into the song system) via $HVC_{(RA)}$ projection neurons to the robust nucleus of the arcopallium (RA). RA projects to the tracheosyringeal portion of the hypoglossal nucleus (nXIIts), which innervates the syrinx, and to brainstem respiratory centers. Shown in *blue*, the anterior forebrain pathway (AFP) conforms to a cortico-basal ganglia-thalamocortical (CBGTC) architecture, which has been specialized for song learning. A second class of pallial (i.e., equivalent of cortical) neurons in HVC, $HVC_{(x)}$, project to the basal ganglia nucleus Area X, within the medial striatum (MSt). The *dashed line* represents the approximate pallial–striatal divide. Area X projects to the dorsolateral thalamic nucleus (DLM), and the CBGTC path is completed by DLM neurons projecting back to a pallial nucleus, the lateral magnocellular nucleus of the nidopallium (LMAN). Neural circuit loops are closed with projections from LMAN back to motor circuitry at the RA, with collateral axons projecting to Area X. Shown in *green*, dopaminergic neurons from the ventral tegmental area (VTA) and the substantia nigra pars compacta (SNc) project strongly to Area X and elsewhere in MSt. Reprinted with permission from Doupe AJ, Perkel DJ, Reiner A, Stern EA. Birdbrains could teach basal ganglia research a new song. Trends Neurosci. 2005 Jul;28(7):353–63 [44]

Neurons in RA that receive projections from HVC are responsible for generating a motor code for actual song production. RA neurons are therefore thought to be equivalent to layer five neurons of the primary motor cortex in mammals [40]. Unique ensembles of RA neurons are sequentially activated by the firing of neurons that project from HVC to RA (i.e., $HVC_{(RA)}$ neurons) during production of syllables constituting the song [41, 42]. The current model proposed for song motor control by RA postulates that convergent activity of an RA ensemble is translated into specific syllable features [42].

The internal circuit dynamics of the vocal motor nucleus RA are modified by another input, arising from the AFP. A crucial advance in relating avian vocal learning to studies on mammals came with the reappraisal of the structural organization of the avian brain [40, 43]. With that, it became clear that the overarching design of the song system AFP reflects a highly conserved vertebrate cortico-basal ganglia-thalamocortical (CBGTC) loop architecture (Fig. 6.1) [44]. In humans and other

mammals, it is known that CBGTC loops play an important role in motor sequence learning, performance, motivation, and neurological disease [45–48]. Thus, rather than being an oddity of the passerine brain, the song system represents a neuroanatomically discrete and vocally dedicated CBGTC subset that may be paralleled by circuits mediating other behaviors [49]. The song system additionally has the advantage of producing a richly quantifiable behavior that is not crucial to the survival of the individual, allowing experimental manipulations of underlying neural circuits to be sensitively assayed, without significantly impairing physiological well-being.

The integration of the AFP with the vocal motor pathway, to form a song control CBGTC loop, can be considered as beginning at the cortical-like nucleus HVC. In addition to local interneurons and $HVC_{(RA)}$ neurons, HVC also contains a second group of projection neurons, those innervating the basal ganglia structure Area X ($HVC_{(X)}$ neurons). Molecular events occurring in Area X (and surrounding striatum) during song production are central to a hypothesis developed in later sections of this chapter, hence our focus on the AFP here. Area X is the largest of song system nuclei and is positioned in the medial striatum [50], although the development of the nucleus suggests it is of ventral striatal origin [51]. During song production, $HVC_{(X)}$ neurons fire in a sparsely bursting mode, somewhat similar to the firing of $HVC_{(RA)}$ neurons. Neural activity is propagated through subsequent AFP connections as follows (Fig. 6.1): Area X sends projections to the dorsolateral nucleus of medial thalamus (DLM) which then projects to the cortical-like nucleus, the lateral magnocellular nucleus of anterior nidopallium (LMAN), thus completing the CBGTC loop. LMAN forms the output nucleus of the AFP, sending a projection to RA. Like HVC, Area X and LMAN neurons are active during singing, as revealed by chronic electrophysiological recordings in the awake bird. These regions of the AFP are also genomically activated by the act of singing, as demonstrated by the induction of the immediate-early gene *zenk* (an acronym for *ZIF-268*, *EGR-1*, *NGFI-A*, *KROX-24*) [52].

Analogy between avian Area X and the mammalian striatum is propelled by the following similarities. Similar to the cortical projections into mammalian striatum, Area X receives glutamatergic cortical inputs from HVC and LMAN. Analysis of firing activity of Area X neurons in singing birds has identified four classes of neurons with firing properties comparable to mammalian striatal neurons [53, 54]. Area X neurons express conserved markers of mammalian striatal spiny neurons such as enkephalin and Substance P [53, 55]. Area X is also similar to the mammalian striatum in its strong dopaminergic innervation arising from the substantia nigra pars compacta (SNc) and ventral tegmental area (VTA), and dopaminergic modulation of neuronal responses [56].

There are some differences, however. Unlike the distinct separation of striatum and globus pallidus (GP) in the mammalian basal ganglia, Area X contains pallidal-like neurons, intermixed with striatal neurons [53, 54]. Two classes of pallidal-like neurons have been described in songbird Area X. In the case of the first of these two classes, GABAergic pallidal-like neurons of Area X send an efferent projection to thalamic nucleus DLM. In some respects, but not all, this pallidal output from Area X

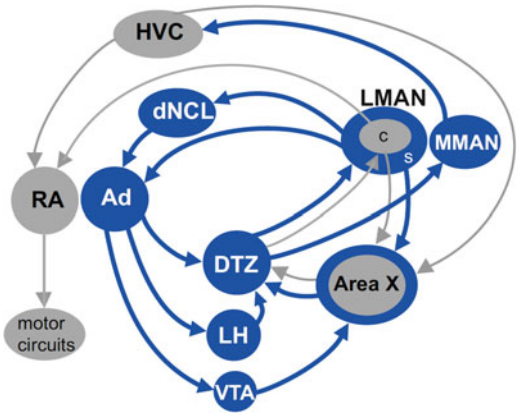


Fig. 6.2 A second pathway traverses the songbird forebrain in parallel with the classical AFP circuit. Shown in *gray*, the song system motor and AFP connections as also depicted in Fig. 6.1. Shown in *blue*, a parallel pathway that prominently includes the shell region of LMAN (s, c = core) and a pallial polymodal association region, the dorsal region of the caudolateral nidopallium (dNCL). Neurons in dNCL display the interesting property of being most active in juvenile birds (as judged by *zenk* expression) when juvenile birds hear the tutor song while engaged in vocal practice, working toward their imitation. Beginning at LMAN shell, neurons project to Ad, a region of the dorsal arcopallium (motor cortex-like) adjacent to RA. $LMAN_{shell}$ also indirectly connects to Ad, through dNCL. Ad sends projections to the VTA, to the lateral hypothalamus (LH) and to the dorsomedial nucleus of the posterior thalamus (DMP, not shown) within the dorsal thalamic zone (DTZ). DMP neurons connect to the medial magnocellular nucleus of the anterior nidopallium (MMAN), which in turn projects to HVC. The DTZ also includes DLM and projections of the *gray* and *blue* pathways from Area X and the surrounding medial striatum remain segregated, respectively, in dorsolateral and ventromedial portions of DLM. These two regions of DLM project to $LMAN_{core}$ and $LMAN_{shell}$, respectively, thereby producing two parallel closed circuits. Overall, the schematic reinforces the reiterated and nested loop design of forebrain circuits mediating motor learning. Reprinted with permission from Bottjer SW, Alderete TL, Chang D. Conjunction of vocal production and perception regulates expression of the immediate early gene ZENK in a novel cortical region of songbirds. J Neurophysiol. 2010 Apr;103(4):1833–42 [59]

to the thalamus appears to be functionally similar to the mammalian basal ganglia-to-thalamus direct pathway [57]. It is not yet known whether the internal circuitry of Area X also contains an equivalent of the mammalian indirect pathway. Suggestively, however, the second class of pallidal-like neurons of Area X is internally projecting and possesses firing properties similar to internal globus pallidus neurons (GPi) [54].

Completing the final leg of this CBGTC loop, LMAN activity is inhibited by Area X, through DLM. LMAN activity thus represents the summation of the AFP and sends the output of the CBGTC loop to RA via a glutamatergic projection modulating the activity of RA neurons. In addition to the loop beginning at HVC, a recurrent CBGTC loop is formed by connections from LMAN to Area X, onto DLM, and back to LMAN again (Fig. 6.2). The two above-mentioned loops engage

the core region of LMAN and the dorsolateral region of DLM. Bearing further resemblance to parallel recurrent loops involving the mammalian basal ganglia [58], an additional recurrent CBGTC loop connects the shell of LMAN to the medial striatum surround of Area X, onto the ventromedial portion of DLM, then back to the shell of LMAN (Fig. 6.2) [59]. This loop also spawns an indirect feed-forward connection to HVC and further recurrent loops to the polymodal association cortex-like area, dorsal caudolateral nidopallium (dNCL), and to the dorsal arcopallial region (Ad) adjacent to RA. These additional regions of the songbird brain appear to participate in song learning alongside the canonical CBGTC loop encompassing HVC–Area X–DLM–LMAN [59, 60]. How these pathways interact requires further investigation and might reveal whether there are instructive analogies to be drawn from the interplay between associative (dorsomedial) and sensorimotor (dorsolateral) basal ganglia loops in the mammalian striatum during learning [58] and, respectively, the LMAN shell and LMAN core encompassing CBGTC circuitries.

The AFP Drives Vocal Exploration and Plasticity

Vocal production in zebra finches begins with the production of soft unstructured babbling-like subsong around 30 days post-hatch (dph), marking the beginning of the sensorimotor phase in vocal learning (see Chap. 3). The first step in the vocal learning process involves memorizing an adult song. During the sensorimotor phase juvenile males develop their own song to be a replica of the memorized template through a series of trial and error learning steps. Similar to mammalian basal ganglia circuits involved in motor sequence learning, the AFP plays an important role in driving the vocal exploration central to this trial and error learning process [61].

The requirement for LMAN and Area X function in the process of vocal development was first revealed by ablation studies. As the source of cortical output from the AFP onto the vocal motor pathway, lesioning of LMAN was naturally an experimental priority. Comparison of the effect of LMAN lesions on the song of young birds treated at the onset of vocal learning with those treated later in development or in adulthood demonstrated that LMAN is required for normal song development, but, once learned, song is resistant to LMAN ablation [50, 62]. Surprisingly, songs produced by juvenile birds immediately following LMAN lesions prematurely crystallized, having syllables that were more stereotyped and oversimplified in structure and sequence. Conversely, lesions to Area X of juvenile males led to a failure to stabilize the song even in adulthood [62, 63]. Thus, these AFP perturbations have diametrically opposing outcomes, but, in either case, adult song was aberrant and there was impoverished imitation of the tutor song. However, when LMAN or Area X is lesioned in adults that have established a crystallized song, there is very little change in the adult song, indicating the critical role for these AFP nuclei in vocal learning yet their apparent dispensability for adult song production.

Recent findings indicate that, in juveniles, premotor nucleus HVC is dispensable for the production of subsong. In adults, removing HVC inputs to RA converts the stereotyped song into a more subsong-like form, implying the presence of two independent circuits for driving these two modes of song [64]. During the earliest stages of song production, LMAN was found to play a premotor role. Inactivation of LMAN at later stages of vocal development (after 45 dph) leads to abnormally stereotyped song, indicating the premature loss of LMAN-driven variability [65]. Under normal circumstances, as vocal ontogeny progresses, inputs from HVC get progressively stronger and become sufficient to drive the maturing song, supplanting LMAN drive and leading to greater stereotypy [66, 67].

Importantly—although contributing more subtly—the AFP does continue to exert an influence on song production by adult males. Increased vocal plasticity due to inputs from LMAN is observed in the adult, depending on the social context of song production. In the presence of females, zebra finch males sing "directed" song, which has a faster tempo and less spectral variability as compared to their "undirected" song that is not aimed at a listener [68, 69]. This modulation of singing by social context can be observed from around 55 dph in juvenile males [70]. In adult males, both Area X and LMAN neurons exhibit lowered firing rates in the context of directed singing [69], and LMAN neurons show spikes precisely time-locked to the song [71]. Opposingly, increased variability in the mean frequency and syllable timing of undirected songs is accompanied by variable burst firing of LMAN neurons [71, 72].

Context-dependent differences in the firing activity of AFP neurons are also mirrored in the transcriptional control of gene expression during directed and undirected singing. The precise mechanism(s) by which these phenomena are linked is still obscure. The act of singing leads to induction of *zenk* in vocal motor pathway nuclei, which is positively correlated with the duration of singing [52]. However, singing-driven *zenk* expression in the AFP and RA varies based on the social context of song production. Induction of immediate-early genes sensitive to neuronal depolarization (*zenk* and *c-fos*) in HVC, Area X, LMAN, and RA is observed when the birds engage in undirected singing [49, 68]. This induction does not occur in AFP nuclei of birds counter-singing or singing female-directed song [68, 73]. Moreover, activity in the basal ganglia is required for *zenk* induction in RA, such that a lesion of Area X in adult birds abolishes the *zenk* induction in RA following undirected singing [74].

In contrast to the substantial advances made in understanding the neurophysiological involvement of the AFP in vocal learning and in modulation of song under different social contexts, the functional contributions of *zenk* and *c-fos* transcription factor activity under these conditions are still unknown. Indeed, the identification of underlying molecular players supporting birdsong across different regions of the avian brain is still in its early stages. Nevertheless, some progress has been made in uncovering the contribution of avian Forkhead box P2 (*FoxP2*), a transcription factor implicated in human speech and language disorders (reviewed in Chap. 2), to which we now turn.

Involvement of *FOXP2* in Birdsong

Expression of FOXP2 *in the Avian Brain*

The amino acid sequence of the FoxP2 protein is highly conserved in vertebrates, and zebra finch FoxP2 shares 98 % amino acid identity with the human FOXP2 sequence [75, 76]. Consistent with the *FOXP2* gene expression patterns observed in the mammalian brain [77], songbird *FoxP2* is expressed strongly in avian striatum (Table 6.1) [75, 76]. In Area X and surrounding striatum, FoxP2 protein is localized in projection neurons co-expressing Darpp-32, a marker for dopaminergic signaling in the adult striatum. *FoxP2* expression in the pallium is lower than in striatum throughout development and in adulthood. Expression in HVC is similar to the surrounding nidopallium, and LMAN shows lower *FoxP2* signal intensity as compared to the nidopallium adjacent to it. Arcopallium, along with the premotor song nucleus RA, does not express *FoxP2*.

With respect to *FoxP2* expression in subtelencephalic regions, the dorsal thalamic zone (DTZ), which includes DLM, the thalamic recipient of Area X projections, strongly expresses *FoxP2*. Sensory relay stations in thalamus that express *FoxP2* include nucleus ovoidalis (Ov, through which auditory input passes to the forebrain and song system) and nucleus rotundus (Rt, a relay for visual input to the forebrain). *FoxP2* is also expressed in midbrain structures providing widespread dopaminergic projections to the brain, namely, the substantia nigra (SN) and VTA. In the cerebellum, *FoxP2* expression is found in Purkinje cells and in the inferior olive, a brainstem nucleus which gives rise to climbing fibers that project onto the Purkinje cells. As in humans, the red nucleus in songbirds also expresses *FoxP2* [75, 76]. Thus, *FoxP2* is expressed in key components of a cerebello-rubro-olivocerebellar loop. In conjunction with the inferior olive and cerebellum, the RN might therefore play an important role in cerebellar-based motor learning or timing control [77–79]. This circuit is extensively modulated by cortical and subcortical afferents, and moreover, the cerebellum is activated during the process of learning to produce covert articulations of novel phoneme combinations [80]. Also perhaps suggestive of a role in articulatory control, the RN is activated during stuttering [81]. The red nucleus has often undergone extensive remodeling in mammalian species, and it has been proposed that, in humans, changes to the red nucleus may have contributed the emergence of hominin bipedality and language [82]. The possible contribution of cerebellar circuits to vocal learning in songbirds [83] and humans requires further exploration.

Even though the neuronal substrates underlying song production are highly sexually dimorphic in zebra finches, *FoxP2* expression patterns are not sexually dimorphic during development and in adulthood. Also, this expression pattern is similar between avian vocal learners and non-learners, as well as in the crocodilian brain [75]. Together, the absence of sexually dimorphic expression as well as conservation of expression patterns in reptiles, birds, and mammals, irrespective of their vocal learning abilities, indicates a more widespread role for *FoxP2*, which may be

Table 6.1 Expression pattern of FoxP2 and FoxP1 in selected regions of the songbird brain

Songbird brain regions	Mammalian counterpart	FoxP2 expression	FoxP1 expression
HVC		+	+++
RA (robust nucleus of arcopallium)		–	+++
Ad (dorsal arcopallium)		–	+
HD (hyperpallium densocellulare)		+	+
GP (globus pallidus)	GP	+ (+)	–
*Area X**		+	+++
StM (medial striatum)	Striatum	+ (+)	+
StL (lateral striatum)	Striatum	+ (+)	+
LMAN (lateral magnocellular nucleus of anterior nidopallium)		–	–
Field L	A1	–	–
DTZ (dorsal thalamic zone)	IMMC	+ (+)	+
DLM (dorsolateral nucleus of medial thalamus)		+	+
VIA (ventral intermediate area)	Ventral tier of thalamus	+ (+)	
Ov (nucleus ovoidalis)		+	–
Rt (nucleus rotundus)		+	–
MLd (dorsal lateral mesencephalic nucleus)		+	–
TeO (optic tectum)		++	+
VTA (ventral tegmental area)	VTA	++	+
Purkinje cells (cerebellum)	Purkinje cells	+ (+)	–
Inferior olive (brainstem)	Inferior olive	+ (+)	–

Qualitative level of gene expression is marked by the "+" and "–" signs. Where possible for FoxP2, the corresponding expression level in the mammalian brain is indicated in parenthesis. * Area X represents a specialized region in the medial striatum in the songbird brain that shows both striatal and pallidal properties. The data are distilled from Ferland et al. [102] for the mammalian brain and Teramitsu et al. [76] and Haesler et al. [75] for the songbird. IMMC: intralaminar, midline, and mediodorsal thalamic nuclear complex

necessary for, but not limited to, learned vocalizations. Beginning in the section entitled "Network Analyses of Molecular Processes Supporting Birdsong," we attempt to address what some of the core molecular signaling processes surrounding *FOXP2* function might be. Although our focus is on Area X and an adjacent region of ventral striato-pallidum (VSP), we would not rule out that similar processes might be at play in the red nucleus, inferior olive, and cerebellar Purkinje cells.

FOXP2 *Expression During Vocal Plasticity*

In juvenile zebra finch males, the level of *FoxP2* mRNA expression increases in Area X relative to surrounding striatum from 15 to 50 dph [75]. There is some disparity in the literature as to whether this heightened expression of *FoxP2* in Area X

then returns to baseline expression levels or continues at a stable higher level at 75 dph [75, 84]. The sensory phase when memory of the tutor song is formed begins around 20 dph, before the young males start to sing [85]. The ability to memorize a song model declines by 65 dph, and the song is crystallized in its mature form by 90 dph. Elevated expression of *FoxP2* in the striatum during juvenile development therefore coincides with the early phase of song learning, when syllable structure and sequence is highly variable. Consistent with these observations of elevated zebra finch *FoxP2* during the period of vocal plasticity, a similar increase in *FoxP2* expression in Area X relative to surrounding striatum is found in adult canaries during the months when song is plastic and new syllables are added to the song repertoire [75]. During the breeding season, when the song is relatively stable, this elevated *FoxP2* expression is not seen.

During the critical period of zebra finch song learning, the volume of Area X increases from 25 to 75 dph [86]. The addition of new neurons during the first 2 months post-hatch contributes largely to the increase in the size of Area X [87]. The majority of the newborn neurons recruited to Area X express FoxP2 protein and develop into Darpp-32$^+$ medium spiny neurons. The highest number of FoxP2$^+$ cells is recruited at 25 dph, possibly due to elevated levels of neurogenesis around 4 dph, as cell migration from the proliferative ventricular zone and incorporation into Area X takes several weeks. The increase in FoxP2 expression during juvenile song learning is attributed to the increased recruitment of FoxP2-expressing newborn neurons in Area X from 35 to 75 dph [88]. Interestingly, although expressed in the avian adult ventricular zone, Rochefort et al. [88] did not find parallel FoxP2 expression in mammalian adult neurogenic niches, such as the subventricular zone and the subgranular zone of the hippocampus. As a possible cellular correlate of the association between FoxP2 expression and plasticity, the intensity of FoxP2 expression within newly born neurons of the avian brain is highest within the first 3 weeks of their birth date [89]. Once firmly established in Area X, the expression level of FoxP2 in neurons goes down. In adult animals, as the recruitment of newborn neurons in Area X declines, the number of neurons showing weak intensity of *FoxP2* staining increases. Moreover, across individual adult males, the density of strongly *FoxP2*-positive cells in Area X negatively correlates with song stereotypy [89].

In addition to this age-dependent regulation, *FoxP2* mRNA expression in Area X is regulated by the motor act of singing and social context, but not by hearing song [84]. In adult males, *FoxP2* mRNA and protein are downregulated by undirected singing [90, 91]. This decrease in *FoxP2* levels is not observed following directed singing. Similar downregulation is seen in juvenile males at 75 dph following 2 h of vocal practice [84, 92]. Juvenile vocal practice may be likened to adult undirected singing, with more variability in song structure as compared to the stereotyped delivery of directed adult song. Decreased *FoxP2* expression in Area X in both these contexts implicates this downregulation in increasing vocal plasticity. Unlike the weakly stained neurons found in adult Area X, FoxP2 protein levels in the intensely stained neurons, which are predominantly younger than 3 weeks of age, are not regulated by the amount of singing [89], possibly reflecting the incomplete incorporation of these new neurons into singing-driven circuits. Taken together, *FoxP2*

levels of expression in Area X correlate with plasticity in two distinct ways. On the one hand, high *FoxP2* expression in juvenile Area X reflects the flux of plastic new neurons into this nucleus. On the other hand, expression of *FoxP2* in low-expressing Area X cells may provide a brake (or gate) on plasticity, needing to be shut off for song variability and vocal exploration.

Perturbation to FOXP2 *Function Inhibits Normal Vocal Development*

A potentially important advantage of the song system as a model for speech and language is the (relative) ease with which it is possible to interrogate temporally and spatially discrete molecular processes underlying vocal learning. In order to confirm the involvement of *FoxP2* in vocal development, expression of the gene in Area X was perturbed using an RNA-interference approach [93]. Persistent downregulation of FoxP2 protein levels in Area X was achieved using a virally encoded short hairpin RNA against *FoxP2*, driven by a constitutive promoter. By utilizing this approach, starting at 23 dph, Haesler and colleagues were able to reduce *FoxP2* levels in Area X throughout the critical period of vocal learning and into adulthood. RNAi-mediated downregulation of FoxP2 protein did not lead to cell death or reduced cell density in Area X.

Bilaterally injected juvenile males were tutored individually by adults in sound isolation boxes. This form of tutoring leads to near-perfect imitation of tutor songs by the majority of pupils [94]. Lowered *FoxP2* levels in Area X during the sensorimotor learning period impaired the ability of juvenile males to accurately imitate the tutor song. Along with the reduced accuracy of imitation of syllable features, omission or repetition of syllables in a song motif was observed in *FoxP2* knockdown animals. This decline in the ability to imitate the tutor song was evident even when only 20 % of the Area X volume was affected by the targeting injections. Songs of the birds with lowered FoxP2 levels showed increased variability between syllable renditions as adults. The increased variability of the final songs and deficient tutor song imitation by *FoxP2* knockdown birds is thought to reflect the outcome of enhanced vocal plasticity with attendant defects in auditory-guided motor learning [93].

What cellular processes might mediate these effects of *FoxP2* knockdown? Although *FoxP2*-expressing newborn neurons are recruited in Area X during vocal development, lowered levels of FoxP2 at 30 dph in Area X do not affect the integration of new neurons [95]. However, FoxP2^{+} medium spiny neurons do show a reduced density of dendritic spines following *FoxP2* downregulation. Connections between HVC neurons projecting to Area X are known to be formed by 23–25 dph [96], before *FoxP2* knockdown was carried out in Area X. Thus, this downregulation does not affect the initial organization of pallial inputs, but perhaps impacts further refining of those connections and their ongoing activity.

Similarly, *FoxP2* knockdown—in the ventricular zone—at 23 dph does not alter precursor cell proliferation or subsequent recruitment and differentiation of newborn

neurons. However, the density of spines in the newly generated neurons is reduced, mimicking the effects seen as a consequence of FoxP2 knockdown in Area X [95]. This indicates a role for *FoxP2* in spine formation in both immature and mature spiny neurons. As these FoxP2^{+} spiny neurons receive their inputs from pallial sensorimotor nucleus HVC as well as the dopaminergic signal from SNc, they represent prime candidates for feedback-dependent tuning of motor output.

Canonical and Other Targets of *FOXP2* Function in Birdsong

Participation of the FOXP1 *and* CNTNAP2 *Genes in the Song System*

The aforementioned research on avian *FoxP2* highlights the value of applying to songbirds a candidate-gene approach informed by progress in human genetics. Adopting the same rationale, a member of the neurexin gene family, contactin-associated protein-like 2 (*Cntnap2*), has also attracted the attention of researchers working in the song system. Human *CNTNAP2* (discussed in Chap. 2 by Vernes and Fisher) is enriched in developing human frontal cortex, striatum, and dorsal thalamus and, through genetic linkage analysis, has been shown to be associated with autistic diagnosis and language impairments [97]. In parallel, chromatin immunoprecipitation experiments demonstrated that *CNTNAP2* is a direct target of FOXP2 transcriptional regulation [98]. Zebra finch Cntnap2 shares 84 % amino acid identity with the human CNTNAP2 protein. In situ hybridization with an mRNA probe for zebra finch *Cntnap2* revealed elevated expression in key song system nuclei in the adult male brain [99].

Expression of *Cntnap2* in LMAN and MMAN (medial MAN) is enriched as compared to the surrounding nidopallium. Similarly, expression in RA is elevated compared to surrounding arcopallium. In contrast, Area X shows a reverse trend of lowered expression in comparison with the surrounding striatum. Strong signals are also found in cerebellum, specifically in the Purkinje cell layer, in optic tectum and in the habenula, as well as in midbrain structures involved in auditory processing such as MLd (dorsal lateral mesencephalic nucleus). The enhancement of *Cntnap2* expression in song system nuclei is sexually dimorphic and not seen in females [99]. Studying the developmental pattern of *Cntnap2* expression revealed that the onset of sexual dimorphism in *Cntnap2* expression is concurrent with male-specific development of LMAN and RA. In the case of RA, both males and females show enriched expression of *Cntnap2* until 30 dph. By 35 dph, male $HVC_{(RA)}$ neurons reach their target field, coinciding with the onset of singing. From this point onwards, RA continues to enlarge in males until 50 dph, while in females it shrinks in size drastically. Paralleling this developmental trajectory, *Cntnap2* expression remains elevated in male RA relative to its surrounds, whereas in females the entire arcopallium shows a uniform intensity of expression.

Regulation of downstream targets by FOXP2 is modulated by its interactions with other Forkhead Box domain proteins [100, 101]. *FOXP1*, a close homolog of *FOXP2*, shows overlapping expression in many brain regions [76, 102]. *FOXP1* mutations have been implicated in autism spectrum disorders, gross motor delay, general cognitive impairment, and language deficit (Chap. 2) [103–105]. However, orofacial dyspraxia—the major phenotype found in *FOXP2* human variant studies—has not been associated with *FOXP1* mutations. Considering these broad cognitive deficits, it is unclear whether the expressive language impairments that are common between *FOXP1* and *FOXP2* mutations arise from a specific shared mechanism or are a consequence of autism spectrum disorder and general cognitive disability resulting from *FOXP1* mutation [106].

FOXP2 and *FOXP1* expression patterns overlap one another in the developing and mature brain. In the mouse, as well as in the developing human embryo, both mRNAs are found in the striatum [76, 102]. In zebra finches, Area X shows higher levels of *FoxP1* expression than the surrounding striatum. This expression is sexually dimorphic from as early as 35 dph [75, 76]. As both FoxP2 and FoxP1 are prominently co-expressed in the striatum, beginning in embryonic development, they might function in concert to regulate the development and function of striatal neurons. These proteins can potentially function as a heterodimer [101], but, regardless, it is unlikely that such a complex is obligatory for function. In the mammalian brain, FOXP2 is expressed in deeper layers of cortex, whereas FOXP1 is found in superficial layers. In songbirds, *FoxP1* is expressed strongly in sexually dimorphic song system nuclei—particularly HVC, RA, and Area X. However, unlike *FoxP2*, *FoxP1* is highly expressed in the nidopallilum surrounding LMAN, but is not expressed in LMAN itself. Similarly, only partial overlap is observed in the expression of these two genes in subtelencephalic brain nuclei (Table 6.1).

In summary, there is still much work to be done before the song system functions of the three genes discussed above are clearly delineated. Nevertheless, important insights have come from the studies already conducted on songbird *FoxP2*, providing an archetype for the molecular dissection of vocal learning. Increasingly sophisticated experimental manipulations of *FoxP1*, *FoxP2*, and *Cntnap2* in the songbird brain will illuminate some of the molecular contributions these genes make to human speech, language, and cognitive function.

Network Analyses of Molecular Processes Supporting Birdsong

Are there recurrent themes in the functional roles of target genes regulated by FOXP2, revealing the core componentry necessary for the production of speech? Because of obvious ethical and technical limitations, efforts to identify relevant FOXP2 targets have employed a variety of strategies to tackle this problem, predominantly utilizing in vitro methods on human cell lines and in vivo studies in the mouse (reviewed in Chap. 2). Several studies have examined the potential set of target genes that FOXP2 (and variant forms) might regulate in mammalian neurons,

using chromatin immunoprecipitation techniques (ChIP-chip) and whole genome microarray analysis on a human neuroblastoma-derived cell line, SH-SY5Y [98, 107]. FOXP2 ChIP-chip experiments have also been run on mid-gestation human fetal basal ganglia and frontal cortical tissue [108] and on embryonic brain from wild-type mice and from mice lacking functional Foxp2 protein [109]. In addition, a comparison of gene expression in mice carrying only one functional copy of wild-type *Foxp2*, or two copies of a "humanized" *Foxp2*, revealed significant differences in striatal gene expression between mutant and wild-type mouse embryos [110]. In each of the above studies, up to several hundred genes/FOXP2 targets were identified. Thus, a major challenge for researchers interested in the molecular genetic basis of speech and language is how to select the salient genes among an embarrassment of riches. There was some overlap between the sets of genes identified in these studies, but also significant differences due to the different biological starting material used (i.e., cell lines versus embryonic tissues).

The contextual framework for most of the above studies was the identification of FOXP2 targets that might, *during embryonic development*, contribute to the emergence of circuits that support speech and language. As discussed in earlier sections of this chapter songbird *FoxP2* may contribute to the vocal imitation process during juvenile development, and *FoxP2* expression is dynamically modulated by vocal activity in adult birds [90, 111]. Arguing by analogy, these observations raise the possibility that, rather than genes expressed during embryogenesis, targets of FOXP2 transcriptional regulatory function *during speech* may be more relevant to speech dyspraxia.

Advancing upon their prior studies revealing vocal activity-dependent downregulation of FoxP2 expression [84, 90], Hilliard and colleagues conducted a gene expression microarray analysis of Area X of the adult male zebra finch anterior forebrain after singing [111]. It was possible, using an unsupervised hierarchical clustering approach, termed weighted gene co-expression network analysis (WGCNA) [112], to cluster these genes based solely on gene expression data into distinct gene modules. Three such modules were specific to Area X and not present in an adjacent ventral striato-pallidal region. In total, Hilliard et al. identified a large number of genes (~2,000) whose expression was modulated in Area X by singing activity [111]. However, we are again faced with a surfeit of genes that may function in the production of learned vocalization and the problem of how to identify those that are most critical.

Following up on their initial analysis, the same group used WGCNA to directly compare the genetic microcircuitry of Area X and VSP [113]. Here it is important to note that both regions express *FoxP2* and both regions are active during singing, VSP recruitment being thought to be associated primarily with other body movements, such as the learned dance [114], that the bird produces during singing activity (it is pertinent to recall here that nonspeech motoric deficits are also seen in KE family members bearing a FOXP2 mutation). Curiously, the expression of genes within a subset of VSP gene modules correlated with acoustic features of song (such as frequency modulation, pitch, and goodness of pitch), in contrast to genes in several

Area X modules, whose expression best correlated with the amount of song (or motif number) produced.

For the current purpose of providing an overview of the parallels between genes implicated in birdsong and those potentially contributing to speech and language, we here make two simplifying assumptions to winnow the field. First, we selectively focus on only a subset of the data obtained by Hilliard and colleagues, considering only the modules in Area X and VSP that include the *FoxP2* gene. In Area X, the *FoxP2*-including "brown" module includes 829 genes, and in VSP the *FoxP2-including* "blue" module includes 817 genes. Second, we filter these gene lists further by focusing on just the 158 genes that are common to both (Table 6.2). The rationale for this approach is that, as the genetic microcircuitry within the two modules differs [113], the overlap between them may be enriched for striato-pallidal processes that are especially crucial to the operation of *FoxP2* within the basal ganglia. Even across songbird species there are significant differences in the transcriptome of the same region of the brain [115]. Therefore, filtering candidate genes down to a potentially conserved sub-module seems a useful and perhaps necessary place to start in the search for common molecular underpinnings of birdsong and speech.

The genes highlighted in bold in Table 6.2 are among those mentioned (albeit some of them briefly) in the hypothesis laid out in the following sections. Undoubtedly, many of the other genes in Table 6.2 may be of interest in their own right, but do not (yet) obviously fit into our model. For example, among those *not* highlighted, Slit1, Eya1, Dcn, and PCDH17 have the distinction of being identified as differentially regulated by FOXP2 and $FOXP2^{chimp}$ in the study by Konopka et al. [107]. These genes may therefore be of interest to investigate further in the singing bird, as might others cited in Table 6.2 [116–139].

A Triangulation on Speech Through Birdsong and Human Genetics

Involvement of the WNT Pathway in Birdsong and Speech

We hypothesize a fundamental role for synaptic Wnt pathway signaling in the basal ganglia during the production of learned vocalizations. The framework we advance unifies a wide range of human genetics data on speech and language deficits, focusing on speech apraxia/dysarthria. The members of the Wnt signaling pathway are named as a portmanteau for the *Drosophila wingless* (*wls*) gene and the mammalian *Int-1* gene, the founding members of this pathway identified by Nusse and Varmus, as a preferred mammary tumor virus integration site in the murine genome [140]. Wnts are cysteine-rich secreted proteins that function as signaling ligands in a wide variety of biological contexts, spanning development and disease [140, 141]. There is now known to be considerable complexity to Wnt pathway signaling, including a canonical Wnt pathway that, via β-catenin, activates transcription through the LEF1

Table 6.2 The intersection of singing-regulated FoxP2-bearing modules in songbird Area X and ventral striato-pallidum

ADCY2	ADRA2A	AGPAT4	AP2A2	ARRDC3	ATP5S	C11ORF58	C1QTNF3
C21ORF63	C22ORF36	C3ORF39	C7ORF30	CAMK1	CCBL2	CDC23	CDH13 [116]
CLDN12	**CPEB3**	CPNE2 [117]	**CRK**	**CX3CL1**	CYB5R3	DCN	EFNA5
EFNB2	ELOVL6	ENOPH1 [118–120]	**EPB41L3**	EVL	EYA1	FABP5	**FAT1**
FBXL20	**FGF18**	FNDC4	**FOXP2**	FRMD3	FRMPD3	GABBR2 [121]	GBAS
GBX1	GPR162	GRIA1	GRID2	GSTO2	HAGHL	HGSNAT	HM13
HMGCR	IL7	ING1 [122, 123]	**INHBB**	**IPO5**	**IQGAP1**	KCTD4 [117]	KIAA0182 [124]
KIAA1045	KLHDC2	LHX6 [125]	LHX8 [125, 126]	LOC420454	LOC421792	LOC770209	LRTM2 [127]
MAP1B [128, 129]	MAPK13	MFSD6	MPST	NDUFAF2	NEGR1 [130]	NKTR	NOL4
NPY2R	NR2F1 [117]	**NRXN1**	**ODZ2**	**OPRD1**	**OPRM1**	PALM2	PCDH17 [131]
PDDC1	PDXP [132]	PET112L	PFKFB4	**PHLDB2**	PHYHIPL	PIK3IP1	PLA2G12A
PLCXD3	PLD6	PLXNA1	**PLXNC1**	POLE4	**PRKX**	**PRKY**	PSMD9
PTDSS2	PTPRG	PTPRR [133]	PYCR2	**REEP5**	**RGS17**	RNF169	RPE
RTN1	RUFY2	RUFY3	SCFD2 [134]	SCN3A	SEMA3A	SESN2 [135]	**SH3GL3**
SH3GLB2	**SHISA4**	**SHISA6**	**SIX3**	SLC4A11	SLIT1	**SMAD6**	SNORD12
SNX17	SNX24	SPG7	STK25 [136]	**STMN2**	SY(…)004485	SY(…)010312	SY(…)017925
TAGLN3	TALDO1	**TBC1D15**	TCF12 [137]	**TESK1**	TM9SF3	TMCC3 [138]	TMEM171
TMEM184C	TMEM45B	TMEM93	TRIM36	**TSGA14**	TTC28 [139]	TUBB2A	TYSND1
UBE2B	UBFD1	UCHL1	VAT1L	**WNT5A**	**WNT5B**	WSCD1	XP_2187084.1
XP_2188032.1	XP_2198838.1	ZAR1L	ZC3H7B	ZFP462	ZRANB2		

In total, 158 genes are common between the FoxP2-containing Area X and VSP modules. Those genes highlighted in bold text are mentioned elsewhere in this chapter. Several general themes emerge from surveying the literature regarding the genes in this list. First, as mentioned in the text, many have a direct or indirect involvement with Wnt pathway signaling. Second, more than a few of the listed genes may relate to aspects of p53 function (e.g., ING1, SCFD2, SESN2, TMCC3, TTC28). Third, several of the genes may be regulated by Bcl11b (also termed CTIP2), a master regulator of striatal medium spiny neuron identity. Bcl11b transcriptional regulatory function can occur in conjunction with COUP-TF1 (listed in the table as NR2F1) at particular target genes, perhaps including FOXP2, CPNE2, and KCTD4. Bracketed numbers provide a few highly selective pointers to genes or loci that might contribute to speech and language development, or to processes focused on in this review, but that have not been mentioned elsewhere in the text. The reference key to first authors (only) is as follows: Lodewyckx [116], Desplats [117], Bonnet [118], Rosenstierne [119], White [120], Kantamneni [121], Soliman [122], Thalappilly [123], Quéméner-Redon [124], Paschou [125], Zhao [126], Abdelmoity [127], Tymanskyj [128], Shao [129], Veerappa [130], Hoshina [131], Zhou [132], Hendriks [133], Krieg [134], Maiuri [135], Matsuki [136], Buonincontri [137], Neilsen [138], Brady [139]

transcription factor (Fig. 6.3) and at least two noncanonical pathways that in certain contexts operate antagonistically to the canonical pathway [142].

Although dynamic gene expression changes during birdsong have prompted our reinterpretation of human genetics data in light of potential deficits in synaptic Wnt signaling, in practice it is difficult to disambiguate such deficits from those arising from developmental errors in Wnt-mediated neural patterning or wiring. Indeed, other researchers have sought to identify recurrent themes underlaying speech and language disorders, usually invoking errors in neuronal connectivity, and the reader is encouraged to consider these views alongside the current discussion [143–145].

Surveying the genes in Table 6.2, one might ask whether song is carried by the Wnt. The *Wnt5A* and *Wnt5b* genes encode Wnt-family ligands that, via the so-called noncanonical Wnt signaling pathway, often—but not always—inhibit canonical Wnt signaling [146, 147]. In the human genome, *WNT5B* resides at 12p13.33, almost adjacent to *ERC1*. Peculiarly, the *Drosophila* homolog of *ERC1*, known as *bruchpilot* (*brp*), is next to *Drosophila wnt2*, which is most similar in humans to *WNT7A*/*WNT7B*. Despite the permutation, this might indicate some functional connection between ERC1/brp and WNTs, something we return to in the concluding discussion. Remarkably, microdeletion of *ERC1* and/or nearby genes results in CAS [148]. In addition to *ERC1*, *WNT5B* was also lost for all but one of the ten families having 12p13.33 microdeletion, but even in the case of that family (3 patients) retaining both copies of *WNT5B*, a substantial portion of a large intergenic region between *ERC1* and *WNT5B* was lost. Marrying the songbird data of Hilliard et al. [111, 113] with human genetics and speech evaluations conducted on 12p13.33 microdeletion patients by Thevenon et al. [148], we are tempted to speculate that a key contribution of *FOXP2* function (or that of p63, as will be discussed later) in speech production might be in the regulation of upstream or downstream components of *WNT5B* signaling or in regulating the expression of the ligand itself.

The two Shisa family members in Table 6.2, *Shisa4* and *Shisa6*, might also be of interest as Shisa protein has been found to antagonize Wnt signaling by increasing the retention of the Wnt frizzled receptor in the endoplasmic reticulum [149]. The *Smad6* gene has also been implicated in Wnt pathway inhibition via Smad6 formation of a repressive complex with CtBP and direct binding to the β-catenin/TCF complex [150]. *Prky* and *Prkx*, which also appear in the table, are significantly upregulated Wnt pathway components in lung squamous cell carcinoma [151] and PRKX phosphorylates SMAD6 [152]. The *Six3* gene also encodes a Wnt pathway repressor [153, 154]. Conversely, the Ube2b ubiquitin ligase Rad6b, which adds a polypeptide moiety (ubiquitin) onto target proteins, regulating their proteolytic degradation, may stabilize β-catenin function and is itself also a target of β-catenin/Lef1 signaling [155, 156]. The *Fgf18* gene, present in Table 6.2, is a target of canonical Wnt signaling [157, 158]. In addition, a smörgåsbord of other genes such as *Fat1* [159], *Iqgap1* [160–162], *Inhbb* [163, 164], *Stmn2* [165], the cytoskeletal regulator *Tesk1* [166], and *Plxnc1* [167] also impact aspects of Wnt signaling, to name but a few.

The potential for extensive interactions between Wnt signaling and *FOXP* gene function in neural circuits is still emerging, but may constitute a major unifying

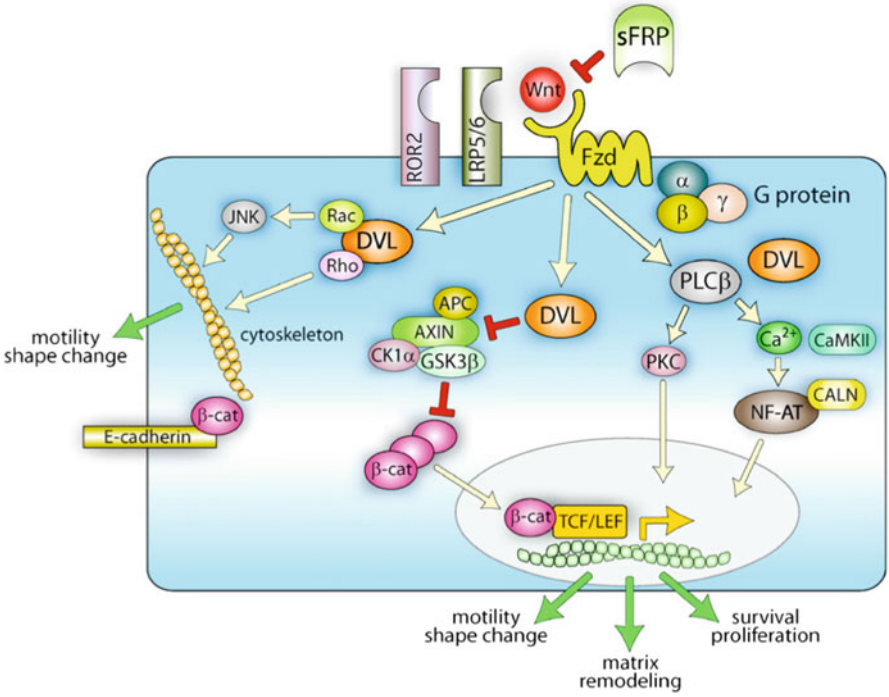


Fig. 6.3 Core Wnt canonical and noncanonical signaling pathways. The canonical Wnt/β-catenin pathway initiates with Wnt binding to Fzd and either LRP5 or LRP6. This Wnt/receptor interaction, mediated by heterotrimeric G proteins, triggers the dissociation of a multiprotein complex that normally functions to facilitate the turnover of cytosolic β-catenin. The complex is composed of an Axin scaffold, which binds APC, β-catenin, and two enzymes that phosphorylate β-catenin and target it for degradation (CK1α and GSK3β). Wnt signaling activates Dvl, prompting destabilization of the Axin complex. Consequently, β-catenin accumulates in the nucleus and associates with TCF/LEF family transcriptional regulators, inducing the expression of specific genes. Signaling via the noncanonical pathway is favored by particular Wnt-family ligands, most notably Wnt5a, and involves Wnt binding to Fzd3, 5, or 7, the receptor tyrosine kinase ROR2, or a complex of both receptor types. This leads to activation of phospholipase Cβ, PKC, and release of calcium from intracellular stores, enhancing the activity of calcium-dependent enzymes such as CaMKII and CALN. These, in turn, modulate the activity of a distinct set of transcription factors (e.g., NF-AT), thereby inducing expression of another set of genes. Dvl is also a component of the Wnt/PCP (planar cell polarity) pathway, transducing Wnt signals to Rac/JNK and Rho-dependent changes in the actin cytoskeleton that alter cell shape and enhance motility (*left side* of figure). Noncanonical signaling antagonizes canonical signaling at several points in the pathway. In contrast, the sFRP family likely inhibits all of these mechanisms of Wnt activity. *APC* adenomatous polyposis coli protein; *CALN* calcineurin; *CaMKII* calmodulin-dependent kinase II; *β-cat* β-catenin; *CK1α* casein kinase 1α; *DVL* dishevelled; *Fzd* frizzled; *GSK3β* glycogen synthase kinase 3β; *JNK* c-Jun-amino-terminal-kinase; *LRP5/6* LDL-receptor-related protein 5 or 6; *NF-AT* nuclear factor in activated T cells; *PKC* protein kinase C; *PLCβ* phospholipase Cβ; *sFRP* secreted Fzd-related protein; *TCF/LEF* T cell factor/lymphoid enhancer factor. Reprinted with permission and minor modification from Rubin JS, Bottaro DP. Loss of secreted frizzled-related protein-1 expression in renal cell carcinoma reveals a critical tumor suppressor function. Clin Cancer Res. 2007 Aug 15;13(16):4660–3 [141]

theme in understanding the genetic underpinning of speech and language. Indeed, it has been noted that Wnt pathway genes may be preferential targets of *FOXP2* transcriptional regulation, suggesting an intimate involvement of *FOXP2* in positive and/or negative feedback regulation of the Wnt pathway signaling [109, 168]. Speaking to the evolutionary conservation of this interaction, in teleosts it has been found that *FoxP2* expression is regulated by the Wnt pathway transcription factor Lef1 [169]. If such a Wnt-*FoxP* link contributes to birdsong, as suggested from the subset of genes shown in Table 6.2, as well as to speech and language, then valuable support for the involvement of this pathway in the production of learned vocalizations might come from mutations in additional components of this pathway that appear to impact speech and language production. Several observations are suggestive of such a link. As just mentioned, Hemizygous deletion of a 1.39 Mb region of chromosome 12p13.33 results in developmental delay, including delayed and/or slurred speech, possibly as a consequence of haploinsufficiency of the *WNT5B* gene [127, 148]. The orphan nuclear receptor ROR2 is a key component of Wnt5A signaling [170], and familial deletion at 9q22, encompassing *ROR2* as well as ~30 adjacent genes, results in severe dysarthria [171]. Polymorphisms in the *WNT2* gene (which resides at the 7q31.1–q31.2 region containing *FOXP2*) are associated with speech delay occurring in conjunction with autism [172]. A constellation of severe neurological structural (microcephaly) and functional deficits (Rett syndrome-like) occurs in patients carrying deletions in the *FOXG1* gene, including absent language development [173, 174]. FOXG1 has several important roles in patterning the developing telencephalon, in part by directly repressing the transcription of Wnt ligands [175]. *FOXG1* expression may also be required for the generation of FOXP1^{+}/FOXP2^{+} striatal medium spiny neurons [176]. It remains to be seen whether *FOXG1* inhibition also contributes to deficits in learned vocal production as a consequence of impaired synaptic WNT signaling during postnatal learning—something that could be tested in songbirds.

Protean WNT Contributions to Speech and Language Deficits in Autism

The posited association between Wnt signaling and speech may also provide insights on autism and schizophrenia. As discussed in the Chap. 2 by Vernes and Fisher, the same genes that constitute risk factors for relatively circumscribed disorders of speech and language may also contribute risk, in some individuals, to neurodevelopmental disorders that more pervasively affect social and cognitive function. Equally, autism and schizophrenia exhibit shared genetic susceptibility loci and provide a prime example of the degree to which variation in the phenotypic expressivity of neuropsychiatric disorders emerges from similar genetic mutations [177, 178]—even despite what has been depicted by some as diametrically opposing behavioral features characterizing these disorders [179]. For the sake of simplicity, here we focus the discussion on possible genetic overlaps between autism and disorders of speech and language.

Dysregulation of the Wnt pathway has been posited to play an important role in schizophrenia and autism spectrum disorders [180, 181]. Similarly, *FOXP* genes

and some of their known targets, including *CNTNAP2*, have been implicated in autism [106, 182, 183]. As a possible neuroanatomical correlate, developmental deficits in Wnt [184–186] and *FOXP2* activity [187, 188] both contribute to impaired cerebellar development and function, which has been commonly reported in autism [189]. Moreover, a significant number of genes identified as singing regulated and common to the *FoxP2*-containing modules in songbird VSP and Area X have been implicated in autism (Table 6.2). In humans, mutation of these genes has traditionally been thought to exact their toll on speech, language, or other aspects of social functioning as a consequence of perturbations to neural development.

Taking the above observations together, it is tempting to speculate that the frequent co-occurrence of speech and language disorders with autism [190–192] reflects the particulars of how this hypothesized FoxP–Wnt genetic circuitry plays out for each individual subject, given the specific constellation of risk-associated mutations and normal genetic variation they possess. In one scenario, the degree of Wnt pathway impairment in the cerebellum or basal ganglia might determine, respectively, the mix of classical autistic features and linguistic deficits, although it should be noted that the cerebellum—where *FOXP2* is prominently expressed in Purkinje cells—also likely participates in speech production [80, 193]. Alternatively, it is possible that a rather specific set of molecular processes are involved in speech production and autism, but these are built on a fundamental core machinery that is broadly required for cognitive function. It is only in the very terminal branches that this pathway diverges, such that few mutations occur where cognitive function is largely unimpaired but speech is highly affected (most notably in the KE family) and also only a subset of cases where social affiliative behavior is affected, but speech and some other complex cognitive tasks are not (e.g., in Asperger's syndrome and high-functioning autism).

In summary, a significant insight provided through the songbird work of Hilliard et al. [111, 113] is that it reveals the extent to which striato-pallidal genetic microcircuitry is actively engaged during the learned performance of a critical vocal communication signal. This permits some hope that a deeper understanding of the operation of this genetic microcircuitry in juveniles and adults might lead to pharmacological strategies ameliorating linguistic and other social deficits in autism spectrum disorders and other patient groups. As a precedent, recent studies demonstrate that in animal models of severe neurodevelopmental disorders, such as Rett syndrome, rectification of deficient neuronal (or glial) activity postnatally may profoundly benefit cognitive function [194–197].

Looking Beyond FOXP2: Other Contributions to Childhood Apraxia of Speech

Above, we have highlighted the prevalence of Wnt signaling components in FoxP2-containing gene expression modules activated in Area X and adjacent VSP during the production of birdsong. Continuing this line of reasoning, it is also of interest that deficits in expressive speech have been noted in cases of familial adenomatous

polyposis, caused by loss of the adenomatosis polyposis coli gene (*APC*) at 5q22 [198, 199]. APC is a critical component of a multiprotein destruction complex that degrades β-catenin when in the absence of active Wnt signaling [200]. It is noteworthy that genes in close proximity to APC may also impact Wnt signaling, such as *MCC* and *EPB41L4A* [201, 202]. The *Reep5* gene (also known as *TB2* or *Yop1*), residing between *MCC* and *APC*, appears among the songbird "core FoxP2" module genes listed in Table 6.2, as does Epb41l4a-related gene, Epb41l3. Both *Ctnnb1* (β-catenin) and *Reep5* mRNAs are transported into axons [203], and, given the potential role of the latter in regulating membrane curvature, endoplasmic reticulum function, and intracellular membrane trafficking [204, 205], Reep5 protein might participate in Wnt-stimulated axon remodeling, synaptic assembly, and perhaps also dendritogenesis [206–209].

Of special interest are the recent set of findings by Shriberg and colleagues, obtained in a microarray-based comparative genomic hybridization (aCGH) study of CAS. These authors were able to identify copy number variations in 12 of the 24 participants in their study and, in addition, checked for *FOXP2* mutations by sequencing all 17 *FOXP2* coding exons in each subject [210]. One subject was found to have a likely pathogenic mutation in *FOXP2* and another subject a deletion in *CNTNAP2*. Focusing our attention on the remaining 10 subjects carrying copy number variations, we find tantalizing hints of Wnt pathway involvement. A small 53 Kb deletion at 17q23.2 affects the copy number of a single gene, musashi RNA-binding protein 2 (*Msi2*). The *Msi2* gene is transcriptionally repressed by a TCF1 isoform containing a long c-terminal "E tail," whereas the same TCF1E isoform is required for LEF1 promoter activation [211, 212]. *Msi2* functions to repress translation of the p21 cyclin-dependent kinase inhibitor and likely many other targets in the CNS. Besides prominent expression in regions of adult neurogenesis, such as the subventricular zone, *Msi2* is also widely expressed in parvalbumin-positive GABAergic interneurons and also in some cholinergic interneurons in the striatum [213]. Either of these latter two cell populations might contribute to the production of learned vocalizations by the basal ganglia and/or other regions of the brain. It is an intriguing, but still unapproached, question as to whether the yin–yang nature of MSI2 and LEF1 responses to TCF1E represents a molecular gating mechanism, such that WNT/β-catenin targets in a poised state of chromatin configuration [214] are transcriptionally activated only under appropriate contexts during the production of learned vocalizations.

A WNT–Estrogen Nexus May Underlay Deficits Due to 16p11.2 Microdeletion

We now turn to the most interpretively complex of the verbal apraxia-associated CNVs identified by Shriberg and colleagues: two cases of CAS associated with 16p11.2 microdeletion syndrome [210, 215]. This microdeletion syndrome results in haploinsufficiency of 29 genes, thereby making the attribution of speech deficits to any one gene rather problematic. This is a relatively frequent pathogenic microdeletion and the second-most common chromosomal abnormality associated with autism [216], with developmental delay and speech and language deficits occurring

in the majority of patients [217, 218]. Duplications at 16p11.2 can contribute to autism and schizophrenia, as well as other cognitive or psychiatric impairment [178]. Several groups have set out to systematically dissect the possible syndromic contributions of genes in the deletion interval by testing the consequences of their reduced function in zebrafish [219, 220]. Assuredly, an argument could be made for deficits in many of these 16p11.2 genes adversely affecting neuronal circuits, including those involved in speech—for example, the *ALDOA* gene has been identified as a target of FOXP2 transcriptional regulation [168].

Most interestingly, for the current argument, a case can be made for 7 of the 29 genes potentially contributing to aspects of Wnt pathway activity. Working from the centromeric to telomeric end of the deletion interval, mitogen-activated protein kinase 3 (*MAPK3*, also known as *ERK1*) is involved in many signal transduction pathways, the Wnt pathway included [221, 222]. The product of the T-box transcription factor 6 gene, *TBX6*, can cooperate with Wnt pathway transcriptional LEF/TCF factors to regulate target gene expression [223]. Also at the transcriptional level, the INO80 chromatin remodeling complex subunit CCDC95 (INO80E) could impinge on Wnt signaling, as other subunits of the complex, INO80H and INO80J (pontin52 and reptin52, respectively), bind β-catenin to antagonistically control the β-catenin-TCF transactivation complex [224]. Located adjacent to *INO80E*, the *HIRIP3* gene product interacts with the chromatin remodeling repressor HIRA [225], and HIRA may be a target of Wnt pathway activity via phosphorylation by GSK3β [226]. Interestingly, HIRIP3 and HIRA proteins have human-specific phosphorylation sites [227], raising the question of whether the genes encoding these proteins have contributed to some human-specific evolutionary trait (given the current context, speech and language represent an interesting possibility).

The BTB/POZ domain-containing protein KCTD13 (potassium channel tetramerization domain-containing 13) has been favored as a candidate for contributing to the 16p11.2 microdeletion syndrome phenotype as gene dosage of *KCTD13* appears to be important [220]. Of particular note, KCTD13 interacts with the dishevelled 2 protein, DVL2 [228], which functions as a key regulator of the Wnt pathway [229, 230]. The *MAZ* gene encodes a zinc finger transcription factor that binds guanine quadruplex motifs. The promoters of WNT pathway genes are enriched for these motifs [231], and MAZ potentially regulates expression of the low-density lipoprotein receptor LRP5 [232], which functions as a co-receptor for the frizzled family of Wnt receptors. When expressed in cells in conjunction with frizzled-4, levels of LRP5 expression dictate whether WNT5A signaling occurs by the noncanonical or canonical pathway [233]. Finally, at the telomeric end of the 16p11.2 microdeletion resides the *SPN* gene, encoding sialophorin (also known as leukosialin or CD43). The cytoplasmic domain of CD43 translocates to the nucleus, physically interacts with β-catenin, and may be required for wnt/β-catenin signaling pathway function [234]. Taking the above observations together, it seems reasonable to conclude that there is significant potential for copy number variation at 16p11.2 to affect Wnt signaling. It is also worth noting that a large region of this locus is coordinately regulated by estrogen [235]. We speculate that 16p11.2 might be a critical locus for the convergence of the Wnt and estrogen signaling pathways [236–238], thereby mediating some of the effects of estrogen on vocalization.

Estrogenic Contributions to Perception, Vocal Production, and…Literacy

There is a substantial body of songbird literature concerned with the contribution of different mechanisms of estrogen signaling to the establishment and function of sexually dimorphic circuits for vocal perception and production [239, 240]. Estradiol administration to hatchling female zebra finch chicks potently masculinizes the developing songbird brain both structurally and functionally [241, 242]. Moreover, estrogen can exert very rapid effects at the synapse, acting on the timescale of seconds, to alter the excitability of song circuits [243], as has also been demonstrated in regions of the mammalian brain [244]. Possibly, the effect of variation in endogenous estrogen levels on Wnt pathway components (at 16p11.2 and perhaps elsewhere) could contribute to fluctuations in speech over the menstrual cycle and in postmenopausal declines to verbal fluency [245, 246]. Whether the equivalent 16p11.2 syntenic region of the songbird genome might, via Wnt pathway modulation, contribute to the development of sexual dimorphisms of the song system and whether sex steroid influenced seasonal alterations in adult neurogenesis and song structure, or alterations in song perception due to synaptic effects of estrogens, remain open questions.

The enzyme aromatase converts androgen into estrogen and is central to brain estrogenic signaling. In humans, mapping of a chromosomal translocation breakpoint in a dyslexic individual revealed that the translocation disrupted the promoter region of *CYP19A1*, which encodes aromatase [247]. CYP19A1 resides at 15q21, a region of the genome known to harbor the DYX1 locus—one of the nine genetic loci tied to developmental dyslexia [248]. In an earlier study, a translocation breakpoint associated with developmental dyslexia had been mapped to the *DYX1C1* gene in this same region [249]. *DYX1C1* also has demonstrated links to estrogen signaling. DYX1C1 interacts with estrogen receptor *a* and *b* [250], and expression of this gene appears to be regulated by 17β-estradiol and estrogen receptor *b* [251]. Together these data point toward estrogen/steroidal signaling as a contributing factor in the development of dyslexia [247, 252]. There is potentially an interesting convergence of this genetic data with the fact that several lines of evidence point toward cerebellar deficits in developmental dyslexia [253] and, moreover, the cerebellum is a major site of neurosteroid synthesis [254]. There is, however, some dissenting opinion on steriodal contributions to dyslexia risk [255, 256], and we suggest a way to reconcile these views in the conclusion.

Traffic Complexity Contributes to the Genetic Heterogeneity of Neurodevelopmental Disorders

One of the patients in the Laffin study [210] showed a deletion at 8q21.13. Beyond the diagnosis of CAS, phenotypic data on this particular patient is lacking. The deleted interval does not include any known gene and falls in an intergenic region. However, immediately adjacent to this region of haploinsufficiency is the

SNX16 (sorting nexin 16) gene. The protein encoded by this gene functions in endosomal trafficking [257] and is localized to a still poorly understood late endosomal subcompartment that is comprised of highly dynamic tubules that require association with microtubules for their biogenesis and movement [258]. Based on studies in *Drosophila*, SNX16 is required for synaptic growth signaling mediated by BMP and *wingless* (Wnt) pathways [259].

Wnt ligand and receptor trafficking events within cells are likely to be critical points at which the signaling process can be disrupted by genetic lesions. In general, the handling of endosomal trafficking by the cell is evolutionarily very highly conserved, with many genes involved in the process being homologous between yeast, invertebrate metazoans, and man [260]. However, the significant complexity to intracellular sorting processes also allows the possibility that Wnt pathway events may preferentially engage embellishments on core trafficking pathways, such that deficits in those particular aspects of trafficking could compromise neurons dependent on synaptic Wnt signaling (in the context of this hypothesis, those involved in speech and language), while comparatively sparing other neuronal populations. On the other hand, where mutations impact components that are central to core aspects of endosomal trafficking, we would expect speech and language deficits to be accompanied by more general cognitive impairment, as might also be expected from mutations that drastically impair all Wnt signaling during neurodevelopment.

The retromer complex [261, 262] functions in the endosome-to-Golgi retrieval of many important transmembrane proteins. Typically, the tubulation of endosomal membranes required for this retrieval process occurs via the oligomerization of sorting nexin dimers (of SNX1/SNX2 or SNX5/SNX6), and the dimer is thought to then recruit a ternary complex of Vps26-Vps29-Vps35 proteins, which constitute a cargo recognition complex. Atypically, however, sorting nexin SNX3 is required for retromer-mediated retrieval of the protein wntless, a critical component of the Wnt signaling pathway [263]. Disruption of SNX3 as a consequence of chromosomal translocation results in a constellation of severe developmental morphological defects, microcephaly, and profound intellectual impairment [264].

It also seems likely that mutations compromising the gene encoding VPS35, which interacts with SNX3, affect speech production. VPS35 is also known to be required for retrieval of Wntless (WLS) [265, 266], perhaps providing a basis for the fact that a functional disomy including *VPS35* results in severe speech apraxia, despite good nonverbal communication skills [267]. Other lesions at the same locus, for example, resulting in overexpression of *VPS35* due to chromosomal translocation, may contribute to severe cognitive impairment and an averbal phenotype [268]. It is worth noting that, in the latter case, the patient was initially suspected of having Angelman syndrome (see section “Syndromic Disorders of Speech, with Lessons from Papillomavirus”), and the hypothesis advanced in this chapter provides a basis for interpretation of this phenotypic similarity.

In the Laffin study [210], a CNV microdeletion at 9q32 compromises two genes encoding zinc finger proteins (*ZNF883* and *ZFP37*). Of *ZNF883*, nothing is known. *ZFP37* is expressed in neurons and encodes a nucleolar-localized protein that may

bind centromeric DNA and influence nucleolar/centromeric architecture, thereby perhaps regulating ribosomal RNA synthesis and ribosome assembly [269]. For our purposes, *ZFP37* is currently of little interest. Of more relevance to us are the genes that flank the 9q32 deletion interval, namely, *FKBP15* and *SNX30* (the latter, we consider in the following section).

FKBP15 (FK506-binding protein 15, also known as *WAFL*) is highly expressed throughout the brain [270] and during development may modulate neuronal growth cone behavior [271]. FKBP15 is able to interact with FAM21 and, via this interaction, participates in the interaction between the WASH (Wiskott–Aldrich syndrome homologue) complex and the retromer VPS35 subunit [265, 272]. Although the function of FKBP15 is still unclear, its association with endosomes appears to depend on vATPase proton pump activity [265], which contributes to acidification of the endosomal lumen during the early-to-late endosome maturation process. We will return to how aberrant acidification of intracellular organelles may contribute to speech and language deficits in due course. Parenthetically, in addition to known retromer components, Harbour and colleagues identified CCDC22 as binding FAM21 [265]. *CCDC22* has recently been put forward as a candidate gene for X-linked intellectual disability [273].

Preferential Compromise of Speech Might Reflect Deficits in Exosomal Signaling

SNX30, the other flanking gene at 9q32, may further point to a hypothesized role for Wnt signaling in childhood speech apraxia, albeit with the caveat that this supposition involves a connect-the-dots exercise. *SNX30* is highly conserved and in yeast corresponds to *SNX42*, the product of which acts in concert with two other yeast sorting nexins, SNX4 and SNX41. In yeast, the SNX4/41/42 ternary complex is involved in a retrieval pathway that is separable from that served by the retromer complex. That is, in yeast the SNX4/41/42 complex is required for retrieval of Snc1p to the Golgi from post-Golgi endosomes [274], whereas the retromer complex is involved in retrieving proteins from pre-vacuolar endosomes, sparing them from degradation in the vacuole (the yeast equivalent of the lysosome). As we will return to in a moment, the human homolog of Snc1p is the synaptobrevin, vesicle-associated membrane protein 2 (VAMP2). SNX4/41/42 also co-localize with Snf7p in the late endosome/multivesicular bodies (MVBs) [275] and Snf7p is involved in MVB formation. Significantly, Snf7p additionally appears to play a role in determining the protein composition of yeast extracellular vesicles, which bear numerous resemblances to mammalian exosomes [276].

In humans, the three paralogs of Snf7 are charged MVB proteins 4a, 4b, and 4c (CHMP4A/B/C). These Snf7 homologs in mammalian cells are components of the ESCRT-III complex, which is required for cytokinesis (Fig. 6.4), viral budding, and the formation of MVBs [277–279]. CHMP4 proteins interact with ALIX [280, 281], a homolog of yeast Bro1/Vps31, an interaction that may promote assembly of

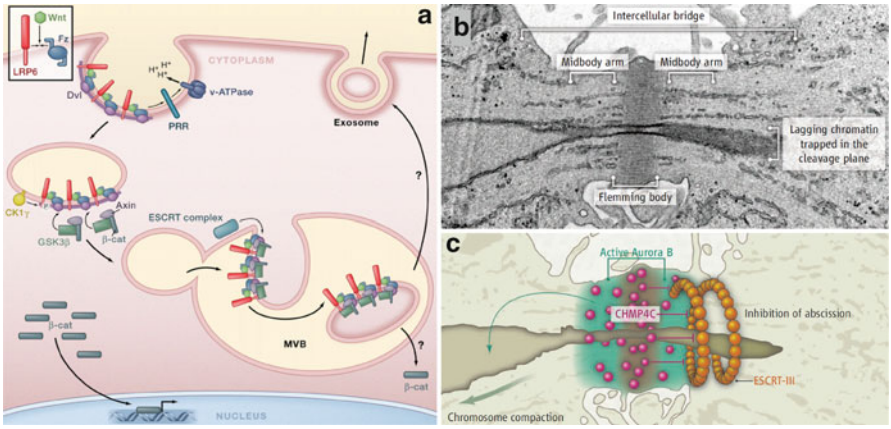


Fig. 6.4 Wnt signaling depends on multivesicular body formation. (**a**) Signaling by the canonical Wnt pathway requires sequestration of GSK3β in multivesicular bodies. Wnt ligand binding to a receptor comprised of a Fzd-family member and a LDL-receptor-related protein (LRP) results in assembly of a ligand–receptor–dishevelled (Dvl) complex. Internalization of this complex into the endocytic pathway involves engagement of the pro-renin receptor (PRR) adaptor protein and vesicle acidification through the action of the vacuolar ATPase (vATPase). Receptor clustering induces LRP phosphorylation by Casein kinase 1 gamma (CK1**γ**) initiating signalosome formation by the recruitment of GSK3β and the destruction complex, which includes Axin and adenomatous polyposis coli (APC). Signalosomes are recognized by component(s) of the ESCRT (endosomal sorting complex required for transport) complex and sorted to vesicles destined for intralumenal budding into multivesicular bodies. MVB formation sequesters GSK3β, dishevelled, Axin, and β-catenin from the cytoplasm. As a result of GSK3β isolation from the cytoplasm, remaining cytosolic β-catenin (and many other proteins) is protected from degradation and translocates to the nucleus. As emphasized in the text, MVBs also participate in the biogenesis of exosomes and through this mechanism could additionally contribute to intracellular Wnt signaling. (**b**, **c**) Multivesicular body formation utilizes cellular machinery that ancestrally evolved for the abscission step of cytokinesis. In (**b**) a thin-section electron micrograph of a cell cytokinetic furrow is shown revealing the relative position of the Flemming and midbodies in the intercellular bridge between two daughter cells. The corresponding schematic in (**c**) depicts the positioning of the Aurora B kinase at the midbody arms, where it is activated by lagging chromatin and phosphorylates the ESCRT-III protein CHMP4C, thereby causing it to localize at the Flemming body and delay abscission. Several proteins localized at the midbody or Flemming body are mentioned in this chapter with respect to speech and language production, most likely through their presumed contributions to MVB formation. Part (**a**) is reproduced with permission from Niehrs and Acebron [290]. Part (**b**, **c**) are reprinted with permission from Petronczki M, Uhlmann F. Cell biology. ESCRTing DNA at the cleavage site during cytokinesis. Science. 2012 Apr 13;336(6078):166–7 [297] (illustration by P. Huey/Science)

the ESCRT-III complex [279]. Importantly, the intralumenal vesicles of mammalian MVBs are directed toward two trafficking pathways: one to the lysosome where their contents are degraded and the other to the plasma membrane where intralumenal vesicles are released as exosomes. Notably, VAMP2, ALIX, CHMP4C, and components of the WNT pathway are all released into exosomes [282]. The release of WNT pathway components into exosomes may have great functional consequence, representing a mechanism of intercellular WNT signaling that overcomes the

hydrophobic nature of WNT ligands [283] enabling, for example, transsynaptic WNT communication between neurons [284, 285]. The *Tbc1d15* gene encodes a GTPase-activating protein involved in exosome secretion [235] and is represented in Table 6.2, alongside *Wnt5a* and *Wnt5b*, perhaps pointing to a possible functional basis for its inclusion in "core FoxP2" WGCNA modules alongside Wnts. Along similar lines, it may be worth noting that ALIX induces cytoplasmic vacuolization and binds endophilins in the process [286]. Two endophilin genes, endophilin A3 (*Sh3gl3*) and endophilin B2 (*Sh3glb2*), are shared by the two singing-induced basal ganglia *FoxP2* modules, shown in Table 6.2. Endophilin A3 exhibits a potent ability to drive the formation of small vesicles [287]. Interestingly, severe speech delay, with intellectual disability and behavioral features, has been reported in a number of individuals carrying an interstitial microdeletion at 15q25.2 [288]. The 1.7 Mb interval deleted in common across patients in this group involves 27 genes. Although Palumbo and colleagues speculate on the contribution some of these genes might make to the speech and other cognitive deficits apparent in 15q25.2 deletion patients [288], we suggest that *SH3GL3* should be considered a high-priority candidate (targets of a microRNA these authors discuss also fit well with the overarching themes here). This example is yet another illustration of the potential synergy that might come from interweaving birdsong and human genetics evidence in the search for molecular mechanisms controlling complex learned vocalization.

Unexpectedly, the inhibition of GSK3β function during canonical WNT signaling occurs by sequestration of GSK3β into MVBs (Fig. 6.4) [289, 290], and conversely, WNT signaling may be antagonized by release of β-catenin into exosomes [291, 292]. Possibly representing a point of crosstalk between the canonical and noncanonical Wnt signaling pathways, WNT5A signaling through the frizzled 2 (Fz2) WNT receptor regulates ESCRT-III complex function through a WNT5A-stimulated association between Fz2 and CHMP4B [293]. In summary, we propose that the 9q32 childhood speech apraxia microdeletion identified by Laffin et al. [210] compromises (synaptic) WNT signaling via endosomal trafficking through MVBs and exosomes. This could provide a unifying basis for interpreting speech production deficits associated with mutations in many of the genes mentioned above, including *SNX3*, *VPS35*, *SNX30*, *SH3GL3*, and perhaps also *FKBP15* and *SNX16*.

The gene adjacent to *SNX16* at 8q21.13 (see above) is *CHMP4C*, perhaps suggesting an abstruse connection between SNX16 and ESCRT-III function that is reflected in the genome. A similarly curious genomic co-localization can be found in the fact that *VAMP2* resides on human chromosome 17p13.1, just 42 kilobases from the Aurora B kinase gene (*AURKB*). This kinase phosphorylates a serine-/threonine-rich sequence found only in CHMP4C and not in the other mammalian Snf7 paralogs [294, 295], directing CHMP4C localization during cytokinesis to a structure termed the Flemming body (Fig. 6.3), where it may function to regulate the timing of daughter cell abscission [295, 296]. VAMP2 has also been implicated in cytokinesis [297]. These details serve to underscore the fact that the ESCRT-III complex and many associated components function in cytokinesis, the ancestral function—and topologically equivalent—of membrane scission events involved in

MVB formation, an essential step in exosome biogenesis. Indeed, CHMP4C, ALIX, and TSG101 are all detected in exosomes [282] and localize to the midbody or Flemming body during cytokinesis (see Fig. 6.3) [277, 296].

Following the logic of the hypothesis proposed here, one might predict that genetic disruption of the 17p13.1 region harboring *VAMP2* and *AURKB* would result in speech and language deficits (and perhaps significant other cognitive impairments). Absent speech or severe speech and language impairment has been reported by several groups to occur as a consequence of microdeletions that include the 17p13.1 region containing *VAMP2* and *AURKB* [298, 299]. Nevertheless, there are important—and informative—complications that arise in attempting to attribute speech deficits at 17p13.1 solely to these two genes. Firstly, as a caveat to the following selective discussion, it should be noted that the telomeric half of 17p13.1, where these deletions occur, is gene-rich (85 refseq genes fall in the interval between Hg18 coordinates 6.0 and 8.3 Mb). Secondly, there is no single region of overlap when considering the microdeletions identified in the eight patients combined from the studies by Krepischi-Santos et al. [298] and Shlien et al. [299]. This indicates that multiple genes at this locus may contribute to the phenotype in these patients, not just the closely linked *VAMP2* and *AURKB* genes. Indeed, 17p13.1 may be a genomic hub of WNT and exosomal control. The biomedical epicenter of this locus is *TP53*, encoding the tumor suppressor gene p53. Indeed, the papers cited above by Krepischi-Santos, Shlien, and their various colleagues focus on not only cognitive deficits associated with this region of 17p13.1 but also cancer risk due to haploinsufficiency of the gene encoding p53. Strikingly, recent evidence from the Levine laboratory suggests that, in mouse embryonic fibroblasts and in a human non-small cell lung cancer cell line, p53 may play an important role in the regulation of exosomal secretion, in part by regulating transcription of CHMP4C [300, 301]. We will return to p53 and its connection to human papillomavirus (HPV) infection shortly, but for the moment it is useful to mention just that inhibition of HPV E6/E7 gene function results in induction of p53 and some of its target genes that facilitate exosome production, namely, *CHMP4C* and *STEAP3* (six-transmembrane prostate protein 3, also known as *TSAP6*) [302]. Continuing this thread—but requiring a brief aside on 2q14.2—we conjecture that genomic alterations in proximity to *STEAP3* that alter its activity (i.e., exosome secretion) could contribute to speech and language delay or to autism [303]. Moreover, the very close proximity of *STEAP3* to the Engrailed gene, *EN1*, and the contribution of the latter to Wnt signaling at multiple regions throughout the body [304–306] and also perhaps to autism [307] may not be mere happenstance.

Several other genes at 17p13.1 reinforce our speculation that this region constitutes a nexus for WNT–exosome–p53 signaling. Most obviously, the dishevelled-encoding gene *DVL2*, an important WNT pathway component, resides within the region. The Scr3 phospholipid scramblase protein, encoded by the *PLSCR3* gene, has been identified as a component of exosomes that is subsequently taken up in a paracrine fashion by nearby cells [308–310]. PLSCR3 physically interacts with ALG-2, as does ALIX (which, as mentioned above, is generally detected in exosomes). Conceivably, this scramblase could contribute to some step in exosome

biogenesis by regulating the asymmetric distribution of specific lipids across inner and outer membrane leaflets [311, 312]. The 17p13.1 genes *CTDNEP1* (in *Drosophila*, provocatively known as *dullard*) and *C17orf81* (also termed *ELP5* or *DERP6*) may also be relevant. The adjacency of these genes is evolutionarily conserved in vertebrates. ELP5 is a component of the elongator complex [313] which has a nuclear function, regulating RNA polymerase II-mediated transcription, and a cytosolic function in which it acetylates alpha-tubulin and controls aspects of vesicular trafficking [314, 315]. ELP5 has also been demonstrated to regulate p53 expression [316]. *CTDNEP1* positively regulates WNT signaling, as decreased *CTDNEP1*/*dullard* function results in a reduction in *DVL2* levels [317]. *CTDNEP1* and the gene encoding its interacting protein, *CNEP1R1*, represent ancient, highly conserved, regulators of the lipin pathway and are able to functionally replace *Nem1p-Spo7p*, their orthologs in yeast [318, 319]. *Nem1p-Spo7p* regulate diacylglycerol synthesis and membrane biogenesis [320], thereby potentially impacting vacuolar/vesicular fission and fusion at multiple stages of intracellular trafficking [321, 322].

Syndromic Disorders of Speech, with Lessons from Papillomavirus

The intimate connection that may exist between the WNT pathway and p53 warrants further emphasis. The latter may powerfully regulate WNT signaling via transcriptional control of the miR-34 microRNA, which targets mRNAs encoding multiple components of the WNT pathway [323, 324]. Moreover, in terms of understanding the molecular underpinnings of speech and language, perhaps of equal interest is the idea that mutations in the Angelman syndrome gene, *UBE3A*, impact this critical nexus of WNT–p53 interaction. *UBE3A* has received considerable attention as a consequence of its involvement in cancer and equal interest because of its—still enigmatic—contributions to neurological features of Angelman syndrome, which include significant cognitive deficits and virtually absent speech and language [325]. One way of getting a foothold on understanding the function of UBE3A in the brain is to look at the cellular consequences of interference with *UBE3A* function in cancer. Human papillomaviruses (HPVs) are causative agents in the majority of cervical cancers and contribute significantly to oropharyngeal carcinoma. A critical step in HPV-mediated oncogenesis is the subjugation of p53 tumor suppressor activity and cellular stress response, via the interaction of the virion-encoded HPV16 E6 protein with the human host's E6-AP protein, encoded by *UBE3A* [326, 327]. Germane to the current argument, the efficacy of HPV in promoting carcinoma is not solely due to the effect of E6, via E6-AP, on p53: the WNT/β-catenin pathway may also participate via interlacing mechanisms [328–331].

The above discussion highlights the possible existence of parallel and complementary p53 and WNT signaling mechanisms in certain forms of carcinoma, and that HPV's hijacking of UBE3A/E6-AP might also inform us of E6-AP's involvement in processes pertinent to the cellular and molecular substrates of speech and language. In order to develop this argument, a short detour is necessary. Critical to

HPV infection is its subversion of the endocytic pathway, enabling the uptake and intracellular trafficking of the virus, with eventual release from a late endosome/lysosomal compartment, uncoating of the virus and transmittal of viral DNA to the nucleus. Exposure to low pH in late endosomes or lysosomes is believed to be necessary for viral uncoating and infection [332]. However, the protracted timecourse of HPV infection may also necessitate mechanisms to prevent premature exposure to low pH. The viral E5 protein, for example, may contribute to alkalinization of endosomes and/or prevention of early endosome fusion with acidified vesicles [333, 334]. Although there is some debate concerning the initial uptake mechanism(s) for HPV, recent evidence suggests that—at least for HPV16—it may occur through a clathrin- and caveolae-independent mechanism that instead involves tetraspanin microdomains and, in particular, the tetraspanin CD151 [335, 336]. Once inside the cell, passage of the virus through the correct subcellular compartments appears to depend on its interactions with the sorting nexin SNX17 [337, 338].

Given the foregoing and the many reports supporting the idea that altered WNT signaling augments HPV infection, perhaps HPV has co-opted endosomal routing mechanisms specifically utilized by the WNT pathway. We speculate that an important centerpiece of the speech and language deficit (and perhaps more general cognitive impairment) in Angelman syndrome is the role E6-AP plays at the conjunction of the p53 and WNT signaling pathways. We will get to the more tortuous details shortly, but for the moment how do we even get HPV16 into the ballpark? HPV infection occurs in a discrete population of squamocolumnar junction cells characterized by a distinct molecular signature that includes notable expression of the WNT/β-catenin target *MMP7*, encoding an interactor with CD151 [339]. HPV infection is thought to occur by entry into basal keratinocytes via the basolateral surface [340]. This fits well with the predominant basolateral surface localization of CD151, which in basal keratinocytes is a constituent of hemidesmosomes [341], and data indicating that SNX17 may function specifically in basolateral receptor recycling through basolateral sorting endosomes, BSEs [342, 343]. Recently, CD151 was identified as an important factor in trigeminal placode development, and the question was raised as to whether CD151 might be required for Wnt signaling [344]. Perhaps even from the get-go, HPV16 infection co-opts the WNT pathway.

In neurons, the somatodendritic portion of the neuron is, in terms of polarized trafficking, the equivalent of basolateral transport in epithelial cells [345, 346]. Consistent with this, SNX17 appears to be required for appropriate cell surface expression of proteins transported into the somatodendritic compartment in neurons [342]. One of the best-characterized targets of SNX17 trafficking is the low-density lipoprotein receptor-related protein 1 (LRP1) [343]. LRP1 has recently been found to mediate WNT5A canonical signaling [347, 348]. Thus, the existence of Snx17 in the songbird FOXP2 module genes listed in Table 6.2 (along with Wnt5a and Wnt5b) may further reflect participation of WNT signaling in the basal ganglia during learned vocal production.

The HPV16 E5 protein may have other strings to its bow, targeting the nucleocytoplasmic transport protein, karyopherin β-3 (KPNB3) [349], and reducing the expression of the cytoplasmic polyadenylation element-binding protein 3 (CPEB3) [350].

Normally, KPNB3 facilitates the nuclear import of CPEB3 [351], and this may be one of the ways that KPNB3 affects the secretory pathway [349]. As CPEB3 function appears to suppress STAT5B (signal transducer and activator of transcription 5B)-activated transcription of the epidermal growth factor (EGF) receptor gene [352], E5-mediated reduction of CPEB3 activity in HPV-positive cells might promote tumor cell transformation and migration by enhancing signaling pools of EGF receptor [353]. Again, the way in which HPV interacts with the cellular machinery in cervical cancer may highlight certain pathways that also cohere in normal striatal neurons, for example, linking *Cpeb3* and *Kpnb3* (see Table 6.2) with Wnt pathway function. If these facets do indeed converge, it may be necessary to consider the possibility that the harmonious intersection of *TP53* (*or the related TP63*; *see later*), *UBE3A*, *CPEB3*, and *WNT* pathways is essential for striatal support of speech, language, and various forms of learning. *CPEB3* has recently been found to have an important role in hippocampal plasticity and memory storage [354, 355], raising the question of whether mutations in *UBE3A* might, for example, impair similarly critical CPEB3-dependent functions in forms of learning and memory dependent on the striatum. Barely glimpsed through a glass, darkly, the outline of these interactions will only take shape with more research.

Angelman Syndrome: The Golgi, the WNT Pathway, and MVBs

But what of the Angelman syndrome gene, UBE3A? Earlier, we touched on the need for HPV to control the timing and location of encounter with an acidified endolysosomal compartment in order to uncoat efficiently and ultimately transfer virion DNA to the nucleus. As mentioned, the HPV protein E5 may play a part in modulating endosomal pH [334, 335], but so too may E6 via its interaction with E6-AP. Ehlers and colleagues have recently demonstrated that lack of *Ube3a* in cortical neurons of Angelman syndrome model mice (*Ube3a*$^{m-/p+}$) results in a dramatic distention of the Golgi apparatus, likely as a consequence of severe under-acidification and consequent osmotic swelling [356]. Alteration to pH in another intracellular compartment of the secretory pathway (namely, the endoplasmic reticulum) was also noted. The mechanistic details of this disturbance to organellar pH are still opaque, but Condon et al. [356] suggest the possibility this may arise as a consequence of impaired ubiquitination, and hence overabundance, of (Na+, K+)/H+ exchangers in the secretory pathway. Moreover, they reveal that protein sialylation is diminished in the brains of mice with insufficient *Ube3a* function. The authors emphasize some of the potential avenues by which impaired protein or lipid sialylation could disrupt neuronal morphogenesis or function, thereby potentially contributing to Angelman syndrome features, such as ataxia, severe developmental delay, and virtual absence of speech. As detailed below, we provide a slightly different interpretation and mesh the work of Ehlers, and other research groups, with our overarching theme.

Of what are probably many, we emphasize three points at which WNT signaling and *UBE3A* function may collide. The first of these is in the regulation of Golgi pH. Ehlers and colleagues mention the Na+/H+ exchangers encoded by the NHE6 and

NHE9 genes, which are both involved in regulation of lumenal pH in the sorting/recycling endosome system [357] and which have both been implicated in autism [358]. The mechanism by which E6-AP regulates Golgi pH has yet to be determined, so it is plausible that NHE6 and NHE9 mutations simply phenocopy Ube3a mutations, without being involved in the same mechanistic pathway. It is therefore a useful exercise to explore the possibilities further. We propose that p53/UBE3A/WNT signaling might modulate Golgi pH through indirect actions on one, or all, of three distinct (but perhaps physically associated) Golgi-localized chloride/anion channels. E6-AP physically interacts with the cystic fibrosis transmembrane regulator-associated ligand (CAL, also termed Golgi-associated PDZ and coiled-coil motif containing protein, GOPC). Interestingly, E6-AP binding to CAL is augmented when HPV E6 protein is associated with E6-AP, resulting in more efficient ubiquitination of CAL, targeting it for proteasomal degradation [359]. CAL functions as a negative regulator of the cystic fibrosis transmembrane regulator, CFTR [360, 361]. It is worth noting that, although now apparently refuted [362], altered Golgi acidification was thought to occur in cystic fibrosis [363]. Disruptions to sialylation, as identified by Condon and coworkers in the mouse Angelman syndrome model, have also been reported in cystic fibrosis [364, 365].

CAL also binds the CLC chloride channel, ClC-3 (CLCN3), and may coordinately regulate both CIC-3B and CFTR [366]. Both channels interact concurrently via the PDZK1 protein (*NHERF3*, Na+/H+ exchange regulatory cofactor NHE-RF3) and can bind the related PDZ protein, EBP50 protein (*NHERF1*, Na+/H+ exchange regulatory cofactor NHE-RF1). Further interactions occur between PDZK1 and EBP50 [367] and between EBP50 and CAL with several of the WNT pathway frizzled receptors [368, 369]. Whether frizzled receptor family members in the Golgi regulate lumenal acidification, via activity of CFTR and/or ClC3-B, is an interesting question that has yet to be addressed. Additionally, WNT function might indirectly impact a third anion channel, although at present this idea is more intuitively pleasing than factually supported. Surprisingly, the gene adjacent to *PDZK1* is *GPR89*, encoding the Golgi pH regulator GPHR [370]. The genomic adjacency of the PDZK1 and GPR89 genes has been maintained over the past 300 million years or so. However, early in vertebrate history (to the extent this may be inferred from extant teleost genomes), the genomic region bearing PDZK1 and GPR89 also housed the core WNT pathway regulator GSK3β. Although GSK3β is no longer syntenic, one wonders whether the WNT pathway might still exert an influence on this locus.

In summary, this assemblage of proteins (CFTR, ClC-3, PDZK1, EBP50, and perhaps GPHR) could participate in a macromolecular complex regulating Golgi acidification, perhaps influenced by/or influencing WNT pathway activity (as discussed below). This will need to be tested, but it is at least encouraging that the ultrastructural Golgi abnormalities reported by Condon et al. [356] in the mouse Angelman syndrome model bear a marked resemblance to the distended morphology of the Golgi in cells carrying mutations in GPHR [370]. Moreover, we suspect that disruption to the normal balance of GPHR and PDZK1 function and/or their adjacent paralogs at 1q21 might make a prominent contribution to the congenital anomalies and neurodevelopmental phenotypes, including mental retardation,

speech delay, autism, and schizophrenia, caused by microduplications and microdeletions at this locus [177, 371–373].

The second point of possible convergence between WNT signaling and E6-AP function centers on the downstream consequences of altered Golgi pH. Condon et al. [356] pointed out the phenotypic overlap of Angelman syndrome and patients with Christianson syndrome, which is due to loss-of-function mutations in NHE6 [374, 375], and the fact that both E6-AP and NHE6 impact organellar pH regulation. Of note, however, the general exocytotic process in *Ube3a*$^{m-/p+}$ neurons is not impaired. Rather, Condon and colleagues find evidence of defective sialylation and speculate this could be instrumental in neurocognitive deficits in Angelman syndrome [356]. Other research indicates that knockdown of *Ube3a* function in late fetal/early postnatal mouse development may perturb large-scale subcellular dynamics of Golgi function, inhibiting the recruitment of Golgi into proximal dendritic regions during the process of apical (but not basal) dendritic morphogenesis [376]. Thus, comparing data from the above groups, the degree to which bulk Golgi functions are affected by reduced *Ube3a* activity is perhaps not fully resolved.

Our interest here is whether WNT pathway function may especially be affected by altered organellar pH, perhaps providing insights (in so far as our model is concerned) on why speech and language are so devastatingly affected in Angelman syndrome. Of course, there is also severe intellectual disability as well. Altered sialylation seems well suited to produce such pan-neural effects and has been implicated as a contributing factor in Down syndrome [377]. However, if deficient sialylation is the main culprit in Angelman syndrome, at least as far as speech and language is concerned, this interpretation does not reveal a clear path toward connecting these Angelman syndrome research findings with other mechanisms of speech and language impairment (or autism). In fact, the same criticism could be leveled at the posited mechanisms of Wnt/E6-AP/chloride channel regulation of Golgi pH we have outlined above. Therefore, we suspect a more specific mechanism may be involved.

As we discussed earlier, mutations disrupting VPS35 function may result in an averbal phenotype similar to Angelman syndrome [268], and VPS35, along with SNX3, functions in the retrieval of the Wntless protein [263]. A conciliation of the findings of Ehlers and colleagues [356] with the framework developed here is that speech and language deficits associated with mutations altering Golgi pH may stem from disruption to Wntless-mediated WNT secretion. Inhibition of vacuolar acidification prevents secretion of functionally active WNTs [378]. Vacuolar acidification is additionally required for the WNT signaling process in recipient cells, establishing endosomal conditions conducive to phosphorylation of the LRP6 WNT co-receptor in the endocytosed signaling complex, and subsequent β-catenin activation [379]. Furthermore, how vesicular pH might impinge on the content, release, uptake, and signaling of Wntless containing synaptic exosomes remains an open question [284].

It is also noteworthy that, although the emphasis was placed on Golgi acidification, Ehlers and colleagues additionally report that pH in the ER was moderately elevated. Palmitoylation of newly synthesized WNT molecules occurs in the ER through the activity of the acyltransferase Porcupine (PORCN). In *Drosophila*,

palmitoylation was found to be essential for the WNT-mediated redistribution of Wntless (WLS) to the Golgi [380], and in mammalian cells WLS may even interact with lipidated WNT ligands in the ER, before they have exited that compartment [378]. WNT secretion is exquisitely sensitive to the level of PORCN activity [381], raising the question of whether elevated ER pH in Angelman syndrome might crucially impair WNT palmitoylation. This remains to be seen, but it is of some interest that patients with duplications at Xp11.23, encompassing PORCN, exhibit mental retardation, speech delay, and poor speech articulation [382]. Even in the case of the smallest of these duplications (0.8 Mb, but still involving PORCN), where psychomotor delay was slight and intellectual disability mild, speech was notably affected. It is not known (as far as we are aware) how sensitive the glycosylation of the WNT receptor is to alterations in the pH of the ER; however this could also be a point where WNT signaling is perturbed in Angelman syndrome. Interestingly, the MEST/PEG1 gene regulating this process [383] is adjacent in the genome to *Tsga14* (see Table 6.2) and has been implicated in autism [384, 385], as well as being a prominent site of genomic imprinting [386, 387].

Many issues still need to be worked out before the details of the research findings reviewed above are seamlessly united. Nonetheless, the emphasis we place here on the impact of Golgi/vacuolar pH regulation on WNT signaling provides specific testable hypotheses. For example, it would be valuable to examine the status of Golgi pH and WNT signaling in inducible pluripotent stem cell (iPSC)-derived neurons produced from autistic or CAS patients, particularly those individuals in which the candidate locus falls near to a gene encoding a likely regulator of organellar pH. Recent work from the laboratory of Ricardo Dolmetsch provides a fine example of the utility of this type of approach in the context of Timothy syndrome, which is associated with autism [388, 389]. Along such lines, mutations in the ATP6V0A2 gene (also mentioned by Condon et al. [356]), encoding the Golgi-localized alpha2 subunit of the H^+-vATPase, result in a skin disorder, cutis laxa, which frequently includes a variable degree of developmental delay [390]. Mutations in ATP6AP2, encoding an accessory subunit of the H^+-vATPase, disrupt WNT signaling, the planar cell polarity (PCP) pathway, vesicular acidification, and endolysosomal sorting [379, 391]. Mutations affecting splicing of this gene result in mild-to-moderate mental retardation and speech delay [392]. Suspiciously, one of the childhood speech apraxia copy number variations reported by Laffin et al., a duplication at 8q11.23 (directly involving only AK056897, a noncoding transcript of unknown function), falls next to the ATP6V1H gene encoding the regulatory H subunit of H^+-vATPase [210].

Finally, the last of our three points of convergence of WNT signaling with *Ube3a* function focuses on the possible involvement of *Ube3a* in potentially regulating aspects of cytokinesis and/or MVB formation (and, by extension, perhaps exosomal Wnt signaling). Genetic and proteomic studies conducted in *Drosophila* to identify targets of *UBE3A* activity reveal that the fly and human dUbe3a/E6-AP proteins interact with a conserved member of the Rho guanine nucleotide exchange factor family (Rho-GEFs/ARHGEFs), known in the fly as *pebble* (*pbl*) and in humans as epithelial cell-transforming sequence 2 oncogene (*ECT2*) [393]. ECT2

participates in the function of the Flemming body at the late stage of membrane abscission during cytokinesis [394, 395]. *Drosophila dUbe3a* also regulates the levels of the Eps15 protein [396], which is known to have roles in cytokinesis [397], endocytic recycling, and MVB biogenesis [398]. Very interestingly, ECT2 functions in a complex with plakophilin 4 (PKP4) and is required for the latter's localization and function at the midbody during cytokinesis [399, 400]. We note that PKP4 falls within the 2q24.1 microdeletion found by Laffin et al. in a case of CAS [210].

In summary, perhaps the critical contributions of UBE3A in Angelman syndrome include disruptions to WNT pathway function operating at two levels: alterations to Golgi (and/or ER) pH resulting in impaired Wntless-dependent WNT secretion, and impaired formation of MVBs required for sequestration of GSK3β [289] and exosomal packaging of Wntless [284], WNT ligands [283], and β-catenin [291].

Potocki–Lupski Syndrome: Another Case Where Trafficking May Go Awry

Potocki–Lupski syndrome (PTLS) and Smith–Magenis syndrome (SMS) are caused by reciprocal DNA rearrangements at 17p11.2, occurring as a consequence of non-allelic homologous recombination (NAHR) events during meiosis [401, 402]. In SMS, caused by microdeletion at 17p11.2 or mutations in the retinoic acid inducible 1 gene (*RAI1*), craniofacial and limb abnormalities are often evident, as well as mental retardation, sleep disturbances, hyperactivity, attention seeking, and a variety of self-directed/self-injurious behaviors [403]. The basis for PTLS, caused by duplication of the same region of 17p11.2, is less clear. The *RAI1* gene has been the principal suspect, given that the haploinsufficiency demonstrated in SMS indicates that this gene fulfills an important dosage-sensitive physiological function. The clinical phenotype in PTLS occasionally includes congenital anomalies, but the core features of the disorder are neurobehavioral, oftentimes including autism but—in all cases—speech and language deficits and variable degrees of cognitive impairment [403–405]. By examining deletions across 74 PTLS subjects, Zhang was able to define a minimal common region of duplication resulting in PTLS that spanned only 125 Kb (chr17: 17,510 to 17,635 k; NCBI build 36/hg18) and involved the *RAI1* gene (and the *SMCR5* gene) [405]. Zhang and colleagues conclude that RAI1 and surrounding regulatory sequences are most likely responsible for the clinical features typically observed in PTLS [405].

In songbirds, *RAI1* is differentially expressed in HVC relative to the subjacent brain region [406], and in addition, a comparison of zebra finch, chicken, lizard, and mammalian genomes indicates that *RAI1* has undergone positive evolutionary selection in the songbird [407]. Recalling the etymology of *RAI1*, it is also of interest that HVC is a site of retinoic acid synthesis in the songbird brain (occurring predominantly in $HVC_{(X)}$ neurons) and that inhibition of retinoic acid synthesis disrupts vocal learning [408–410]. In general, the contributions of retinoic acid signaling to the function of the juvenile and adult vertebrate brain are still poorly understood.

HVC represents a very attractive experimental target for refined manipulations of this signaling process, given that perturbation to HVC circuit dynamics in the adult bird would be anticipated to result in readily quantifiable alterations in syllabic timing during song production.

However, in the context of the framework we have advanced in this chapter, there is one additional note of interest regarding the human (and songbird) *RAI1* locus. Just 45 Kb from *RAI1*, separated by only one gene, is the target of myb1-like 2 gene, *TOM1L2*. The *Tom1* gene family is very highly conserved in evolution and, even in amoeba, participates in the formation of MVBs [411]. Similarly, in *Dictyostelium*, the single Tom1 protein functions in MVB biogenesis and interacts with the slime mold homologs of human EPS15 and TSG101 (respectively, EGF receptor pathway substrate 15 and tumor susceptibility gene 101) [412]. Recalling our discussion in the section entitled "Preferential Compromise of Speech Might Reflect Deficits in Exosomal Signaling," we proposed that TSG101 and other proteins involved in MVB formation (and midbody formation) might contribute to apraxia of speech. The vertebrate Tom1l1 protein indeed interacts with TSG101 and other proteins involved in MVB sorting, as well as at the midbody during cytokinesis [413, 414]. The Tom1l2 protein does not, however, possess the domain necessary for binding to TSG101 [415]. Nevertheless, Tom1l2 still binds Tollip, clathrin, and ubiquitin and is likely to be involved in post-Golgi trafficking processes, including the recruitment of clathrin onto endosomes [416]. It is feasible that Tom1l2 might participate in the complex life cycle of endosomes/MVBs/exosomes in neuronal dendrites and/or axons, but this has yet to be established.

Williams Syndrome and dup7: A WNT Receptor and MVB Regulator in the Critical Region

Williams syndrome (Williams–Beuren syndrome, WBS) represents a fascinating microdeletion disorder resulting from the removal of (typically) 1.55 Mb at 7q11.23, as a consequence of the same sort of NHAR mechanism involved in SMS/PTLS [417]. Deletion of this region results in loss of 26–28 genes, several of which are thought to make particular contributions to the suite of morphological and physiological features that characterize the syndrome [418]. The neurodevelopmental consequences of this deletion are of especial interest to researchers interested in the neurological mechanisms of social behavior, auditory perception, visuospatial processing, and speech and language [418, 419]. Average full-scale IQ in WBS adults falls in the 50–60 range, indicating mild-to-moderate cognitive impairment, but on this backdrop there are rather surprising areas of strengths and weaknesses in perceptual/cognitive function. The archetypal WBS subject is described as having a happy hypersocial disposition, although prone to anxiety. They have a strong liking for music, but a high proportion are said to develop a pronounced sensitivity to loud noises (hyperacusis).

A striking facet of the WBS perceptual/cognitive profile is a strong visuomotor propensity to focus attention on faces, fixating on faces for longer durations and

disengaging fixation from the eyes more slowly than do typically developing (TD) children [420, 421]. These visual scanning preferences for such highly salient social stimuli are a striking counterpoint to the situation observed in autism, wherein autistic toddlers tend to fixate on the mouth rather than eyes [422] and autistic children (7–13 years of age) have reduced levels of fixation on faces when viewing social scenes than do TD children [423]. As might be anticipated from the meiotic NHAR mechanism of 7q11.23 deletion (and by analogy to SMS/PTLS), duplications of the same region have also been identified (termed dup7). Duplication results in a largely opposite clinical profile that, for many patients, is said to be indistinguishable from autism [424], although contrastingly it has been argued that the social withdrawal observed in dup7 children might be a sequela of their speech and/or anxiety problems [425].

One of the most remarked upon aspects of WBS is the relative preservation of speech and language function despite other significant cognitive impairment. This portrayal has significantly moderated over the years in light of more comprehensive neuropsychological testing in WBS subjects. The current view is that the developmental timecourse of speech and language development in WBS may be delayed in early childhood, starting with a latter onset to both canonical babbling and another typical infant motor behavior, rhythmic hand banging [426]. As childhood progresses, speech and language development in WBS subjects seems to be largely in step with what would be expected given the chronological age and IQ of the individual [427]. Conversely, there seems to be good evidence that, in many cases, speech and language development is severely affected in cases of 7q11.23 duplication [428, 429]. Mervis and colleagues have studied the speech and language capabilities of 42 subjects with dup7 and report that, of 25 children between the age of 4 and 15, over 75 % met full criteria for CAS, as well as having other oromotor or speech sound symptoms [425].

The dosage-sensitive gene(s) in the 7q11.23 WBS/dup7 interval responsible for the neurodevelopmental features discussed above has yet to be identified with certainty. The LIM domain kinase 1 gene (*LIMK1*) is debatably the main candidate for visuospatial deficits and hyperacusis in WBS [430–432]. As yet, no clear culprit exists for the autism-like and speech and language deficits observed in people with dup7, although genes encoding CAP-GLY domain-containing linker protein 2 (*CLIP2*), General transcription factor II i (*GTF2i*), and Syntaxin 1A (*STX1A*) have been singled out for mention by various authors. Certainly, from the perspective of our hypothesis, the frizzled-9 gene (*FZD9*) in the WBS interval is of interest because of its obvious link to Wnt/β-catenin signaling and the fact that mice lacking *Fzd9* gene function have been reported to display visuospatial learning deficits [433]. Obviously, these mice do not speak to the possibility of whether *FZD9* duplication in dup7 subjects results in apraxia or other speech/language deficits. Most interesting, however, is the duplicated/deleted region also includes the gene encoding *VPS37D*, an ESCRT I complex subunit. The VPS37D protein physically interacts with the midbody protein CEP55 [394] and thus is implicated in membrane scission complex formation, such as is required for cytokinesis and the production of MVBs.

Potential Impacts of Synaptic WNT Signaling on Basal Ganglia Function

The preceding sections developed, we think, a prima facie case for the involvement of WNT signaling in speech and language disorders and, across taxa, in the molecular events that underlie basal ganglia/striatal participation in the production of birdsong. An important question we have yet to address is that of how WNT signaling might intersect with well-established molecular pathways of basal ganglia function. Prospectively, one would imagine there exists a very specific or delicate balance involved in WNT interactions with core molecular machinery. On the one hand, extensive research over many decades would presumably have been more successful in identifying molecular candidates for speech and language production if these were critical to phylogenetically conserved functions of the basal ganglia that, when disrupted, result in patently obvious behavioral deficits in rodents. Given this has not been the case, it seems reasonable to assume that a circumscribed signaling pathway is at play. On the other hand, however, such a pathway cannot exist in a vacuum, and—as evidenced in songbirds and humans—it must interact with molecular pathways that can powerfully regulate learning and social behavior.

Wnt5 Signaling in the Basal Ganglia Might Impact Molecular Mechanisms of Memory

The posited link between exosomal WNT signaling in Area X/basal ganglia and production of learned vocalizations potentially fulfills the first side of the above equation, but how might this pathway affect core molecular processes in this region of the brain? The presence of *Wnt5a*, *Wnt5b*, and other Wnt pathway genes in Table 6.2 raises the interesting possibility that song production is accompanied by changes to the functional status of dopamine and cAMP-regulated neuronal phosphoprotein 32 (DARPP32) and the cAMP response element-binding protein (CREB). DARPP32, also known as protein phosphatase 1, regulatory subunit 1B (*PPP1R1B*), is known to be a key regulator of the effect of dopamine on striatal medium spiny neurons [434, 435], which in songbirds and mammals prominently include FOXP2/dopamine D1 receptor-positive neurons [75, 88, 109]. CREB has a storied history in the molecular biology of learning and memory, being a central component of cyclic AMP (cAMP) and protein kinase A (PKA)-dependent memory formation in Aplysia, *Drosophila*, and vertebrates [436, 437]. Despite many decades of intensive research on CREB and DARPP32 function in the nervous system, insight on their recruitment during speech and language production may, paradoxically, come from a study of cell migration in a breast cancer cell line.

Hansen et al. [438] report that in MCF-7 cells, Wnt5a signaling through a frizzled-family receptor results in activation of the cAMP/PKA pathway, leading to phosphorylation of DARPP32 on amino acid Threonine-34. This phosphorylation converts DARPP32 into a potent inhibitor of Protein Phosphatase-1 (PP1). At the

same time, PKA also phosphorylates CREB on Serine-133, a modification that is critical for pCREB function as a transcriptional activator [439] and observed in association with memory formation [440, 441]. As pCREB is a substrate for dephosphorylation by PP1, the dual events triggered by Wnt5a could initiate a protracted period of pCREB-mediated transcription. Wnt5a signaling in neurons also activates a protein kinase C pathway to modulate synaptic NMDA receptor currents, facilitating the induction of long-term synaptic potentiation (LTP) [442]. Combined, these signaling mechanisms could contribute to Wnt5 alterations to synaptic plasticity and long-term memory [209]. Other Wnt-family members may also participate in this process or perhaps are preferentially utilized to similar ends elsewhere in the brain [443].

Birdsong Is the Opium of the Songbird, Speech the Opium of the People

As mentioned, CBGTC function has been implicated in motor learning, performance, and motivation [48]. Wnt5a acting through DARPP32, PKA, and PKC could contribute to aspects of learning mediated by the basal ganglia, but what mechanisms might connect the motivation to vocalize, with the act of producing the said vocalization? Disruption to such a link might be predicted to result in apraxia of speech. Songbirds represent an excellent model for exploring the motivation to vocalize [444]. Among vertebrates, songbirds are well known for their prolific vocal output, a feature that probably partially reflects their arboreal existence and possession of flight, and hence relative safety from many ground-bound predators. Songbird males (although in some species females sing too and may even duet) are motivated to sing in a variety of different contexts. The most obvious of these are female-directed song, involved in mate attraction, and male-directed song, typically motivated by agonistic interactions pertaining to territorial defense. However, they also engage in high rates of vocalization in the absence of any clearly intended recipient (being largely ignored by conspecifics), a context of vocal production termed undirected song. As we have discussed earlier, in zebra finches, there are acoustic differences between directed and undirected song, and these two modes of singing differentially engage Area X, as starkly revealed by the induction of *zenk* [68].

There are numerous complexities to understanding the basis for the motivation to sing, involving the interaction of steroids, dopamine, and opioids, and for more details the reader could consult the review by Riters [444]. Oversimplified and in a nutshell, sexually motivated directed song may depend particularly on dopaminergic signaling, whereas undirected song is more likely governed by the hedonic value of endogenous opioids. Where opioids in the brain mediate their effect on the motivation to vocalize remains uncertain. Most attention so far has been directed toward the medial preoptic nucleus (mPOA) and VTA. Projections from the VTA to Area X (Fig. 6.1) and surrounding striato-pallidum could conceivably be relevant [445, 446], but it is perhaps more likely that signaling endogenous to the striatum has a role to play. Four genes revealed in the Area X+VSP "core *FoxP2*" module (Table 6.2) capture our attention. The μ- and δ-opioid receptor-encoding genes,

Oprm1 and *Oprd1*, are both represented on this list, as is the gene located next to *Oprm1* in the genome, *Rgs17* (regulator of G-protein signaling 17, also termed *RGSZ2*). This seems not merely to be an inconsequential accident of genomic proximity. *Rgs17* functions in a complex between the μ- and δ-opioid receptors, the NMDA receptor, neural nitric oxide synthase (nNOS) [447, 448], and histidine triad nucleotide-binding protein 1 (HINT1), which is also a repressor of β-catenin signaling [449]. The workings of this complex are thought to constitute a zinc-redox switch, with major consequences for the regulation of numerous signaling cascades within neurons [450]. Notably, in tumor cells, *Rgs17* can activate the PKA/CREB pathway and induce *FoxP2* expression [451]. As a footnote, the *Cx3cl1* gene listed in Table 6.2 encodes fractalkine, a member of the chemokine family of ligands, which has been shown to negatively regulate morphine signaling via the μ-opioid receptor [452–454].

In songbirds, the μ- and δ-opioid receptors are expressed in several nuclei of the song system, including Area X (and surrounding striatum) [455]. Within Area X, the population of neurons expressing the μ-opioid receptor is heterogeneous and requires further characterization [455]. Significantly, the expression of the endogenous opioid, enkephalin, is induced in the Area X of birds singing undirected song [456]. Taken together, perhaps the Table 6.2 striatal *FoxP2* module genes in the opioid pathway mediate the pleasure songbirds (anthropomorphically) appear to derive from the act of singing. But at the synaptic level, how could opioid signaling alter the dynamics of the WNT pathway, which we propose here to be instrumental in the production of learned vocalizations? In passing, we just mentioned that HINT1, part of the μ-opioid/Rgs17 complex, inhibits β-catenin, but there may be even tighter links forged between Wnt and opioid signaling pathways in the striatum. Strikingly, mammalian Wntless has been identified as a μ-opioid receptor interacting protein [457], and, at the cellular level, the striatal neurons co-expressing Wntless and the μ-opioid receptor are enkephalinergic [458]. Moreover, stimulation of the opioid receptor results in a relocation of Wntless from the cytoplasm to the plasma membrane in striatal neuron dendrites [459]. Our level of understanding of these interactions is still too rudimentary to know how this process impacts vocalization, although it has been proposed that opioid agonist-induced redistribution of WLS might abrogate WNT signaling [459].

Whether opioid receptor function contributes to human speech and language remains a fascinating but thoroughly enigmatic issue. There are a few ethereal hints this may be so. *RGS17* and *OPMR1* at 6q25.2 are paralogs of *RGS20* and *OPRK1* at 8q11.23. In the former case, genome-wide linkage analysis has identified 6q25.2 as a locus linked to reading disability, with the peak marker falling in the intergenic region between *RGS17* and *OPMRI* [460]. In the latter case, the study of CNVs in CAS, by Shriberg and colleagues [210], identified a microduplication 8q11.23 that spans most of the region between OPRK1 and RGS20 (betwixt which fall the non-coding transcription unit AK056897 and the ATP6V1H gene, encoding the regulatory H subunit of H^+-vATPase, mentioned earlier). Finally, as an item of paleoanthropological interest, recent high-coverage sequencing of the archaic Denisovan genome (a Neandertal sister group, believed to have diverged from

present-day humans ~800,000 years ago) reveals that the OPRM1 gene of archaic humans possessed a stop codon in a 5′ exon of the gene, as do extant nonhuman great apes [461]. As 97 % of present-day humans do not possess this stop codon, allowing the production of an N-terminally extended form of the μ-opioid receptor, the single nucleotide change responsible may have had some consequence in the emergence of modern humans.

Conclusion

Exosomal WNT Signaling May Be Critical in Learned Vocalization

We have provided what we think is a singularly explanatory dissection of molecular and genetic contributions to speech and language. The crux of our hypothesis is that birdsong, speech, and language may be especially sensitive to disturbances to synaptic Wnt signaling processes (perhaps especially Wnt5-family mediated) that occur *during* the production of learned sounds. The complexity of the Wnt signaling process (including exosomal secretory and uptake mechanisms) appropriately parallels the wide diversity of genetic mutations that, in humans, affect speech and language. In general terms, this has a parallel with an emerging view on the cellular and genetic mechanisms that contribute to stuttering, where mutations in genes encoding lysosomal enzymes have been implicated [462]. Songbird studies developed as a model for stuttering (see Chap. 7 by Helekar) have not yet advanced to the point of exploring the contribution of these genes (*GNPTAB*, *GNPTG*, *NAGPA*) in song production, but this represents an exciting avenue for future research. Currently, we haven't explored mechanistic connections between the lysosomal processing pathway implicated in human genetic studies on stuttering and the synaptic Wnt/exosome-centric pathway we have posited to be involved in striatal contributions to speech. Ultimately, these cellular processes must cooperate in the production of speech, although not necessarily simultaneously in the same cells. Some evidence suggests the possibility that CNTNAP2, or genes in the vicinity, may also be involved in stuttering and/or Tourette syndrome [463–465].

Space limitations prevent us from discussing many other genes in Table 6.2 that could be of relevance to the hypothesis we have outlined here. A more useful exercise is to ask whether this WNT–exosome framework might lend additional support to genes currently implicated in speech and language deficits, provide new perspectives on their function, or unveil new candidates in loci that have been extensively researched, but for which a guilty party remains at large. A balanced translocation or deletions disrupting the *C10orf11* gene at 10q22 result in psychomotor delay, with patients carrying even the smallest deletions producing only a few words by age 4 [466, 467]. Very little is known about *C10orf11*, except that the gene is highly conserved, potentially regulated by p63, and, based on a loss-of-function genetic

screen in *Ciona intestinalis* (the tunicate sea squirt), encodes a novel component of the Wnt/β-catenin signaling pathway [468]. A copy number gain of *C10orf11* has also been noted in a patient with autism [469]. Thus, the threads we have pulled together in this chapter might provide a context for *C10orf11* functioning as a bona fide participant in speech and language production.

Several studies have identified the ankyrin repeat domain-containing protein 11 (*ANKRD11*) gene at 16q24.3 as involved in KBG syndrome, with typical speech delay, variable cognitive impairment, and autism-like features [470–472]. The framework advanced in this chapter provides an entrypoint to understand the mechanism of *ANKRD11*'s contribution to speech and cognitive function, given that the ANKRD11 protein may impinge on p53/p63 function [138, 473, 474]. Additionally, the gene encoding the frizzled-related protein, FRZB (a.k.a. secreted frizzled-related protein 3, *SFRP3*), falls in the critical region for 2q31.2q32.3 deletion syndrome, which features severe mental retardation and absent speech [475–478]. Secreted frizzled-related proteins are generally thought to function as inhibitors of Wnt signaling [479]. Perhaps of particular interest in the current context is the fact that SFRP3 may be mutually antagonistic with Wnt5a activity [480–482].

Does our hypothesis shed any light on the contribution of other genes implicated in speech and language deficits? The case for *CNTNAP2* involvement in autism, speech, and language seems well established (reviewed in Chap. 2). One particularly well-studied cellular function for this protein (also termed Caspr2 or Neurexin IV) is its role in organizing potassium channel protein complexes on the juxtaparanodal regions of CNS and PNS myelinated axons [483–485]. Recently, however, a distinct cell-autonomous function for CNTNAP2/Caspr2 has been identified in dendritic arbor growth and spine development [486]. With regard to this aspect of Caspr2 function, much less is certain. In the somatodendritic region, Caspr2 has been localized to endosomal structures [487], probably reflecting the peculiar transcytotic pathway that it and a few other proteins utilize to generate their polarized distribution in neurons [488]. Only ~40 % of Caspr2 + ve endosomes were labeled with the early endosomal marker EEA1 [487], and the provenance and destination of the other Caspr2 + ve endosomal structures remain undetermined. Thus, it remains an open question as to whether the latter might contribute to some representative of the menagerie of neuronal MVB types [489], such as endocytic marker negative MVBs participating in soma-to-dendrite or interdendritic transfer of material [489, 490]. In broad sweep, a role for Caspr2 in these processes is appealing, given the underdeveloped spine widths observed in Caspr2 knockdown neurons [486] and the role that MVBs proximal to a subset of spines may have in supplying material to adjacent spines within that segment of dendrite [491]. As a cautionary addendum [492] regarding the interpretation of CNVs in the vicinity of *CNTNAP2* (or mutations in *CNTNAP2* itself, if they do not seem to be obviously disruptive to function), it is worth noting that the gene encoding the vacuolar H^+-vATPase e2 subunit, *ATP6V0E2*, resides relatively close by (1.5 Mb telomeric).

Our observations perhaps highlight which of SHANK3's (SH3 and multiple ankyrin repeat domains 3) myriad potential postsynaptic interactions may be most relevant. In 22q13.3 deletion syndrome, also known as Phelan–McDermid

syndrome, *SHANK3* is lost, leading to minor dysmorphologies but major global delays in neurodevelopment, intellectual impairment, and severely delayed or absent speech [493]. In patients carrying the 22q13.3 deletion, behavior is autistic-like and similar behavioral features are observed in mouse models where the *Shank3* gene has been mutated [494, 495]. In particular, striatal synapses and cortico-striatal circuits are impaired in these mice [494]. Of the many biochemical contributions SHANK3 makes at the synapse, what may be of most interest to explore is its association in a complex with LAPSER1 (*LZTS2*) and β-catenin, transducing a signal that, upon NMDA receptor activation, leads to β-catenin nuclear translocation from the postsynaptic density [496].

Remarkably, there is perhaps now a glimmer of the synaptic Achilles' heel of autism, speech, and language deficits. SHANK3 interacts with latrophilins [497] and the latrophilin *LPHN1* may contribute to psychomotor and language delay arising from microdeletion of a region of 19p13.12 [498]. LPHN1 protein in turn binds to the Neurexin NRXN1 [499]. Mutations in NRXN1 are almost invariably associated with severe intellectual disability and speech deficits [500]. LPHN1 also binds tightly in a transsynaptic complex with its ligand, teneurin-2 (encoded by the *ODZ2* gene) [501]. It should be noted here that both Nrxn1 and Odz2 feature in Table 6.2. Quite surprisingly, a recent analysis of *CNTNAP2* and *NRXN1* contributions to reorganizing synaptic morphology in *Drosophila* indicates that both proteins specify the presynaptic concentration of the active zone protein, *bruchpilot* [502]. Remembering back to our earlier discussion, the adjacency of *brp*/*ERC1* to a *WNT* ligand gene has been maintained for ~600–800 million years. Might transsynaptic WNT signaling be at the heart of this agglomeration of proteins encoded by candidate autism/speech and language genes? We speculate that an important function of the ERC1 presynaptic protein might be to deliver WNT ligand (and perhaps other exosomally packaged WNT pathway components) for synaptic release, such that postsynaptically the signal can be transduced by SHANK3/β-catenin. Additionally, ERC1-dependent transport of the relevant molecules to the active zone may be abetted by the interaction of ERC1 with LL5β [503], the product of the *Phldb2* gene (again see Table 6.2) as well, perhaps, as Bassoon and Piccolo, given their involvement with ERC2 in post-Golgi vesicular trafficking [504].

Going Out on a Limb

Many of the genes mentioned in this chapter have homologs that are very highly conserved, some fulfilling the same functions in cytokinesis and MVB formation whether in human cells or in single-cell organisms. Similarly, Wnt signaling is a venerable method of cell–cell communication. We would therefore anticipate that other organ systems/structures in the body also rely on the pathway we have attempted to delineate, either for their development, function, or both. As a corollary, we might also expect that genetic lesions affecting such structures would frequently be associated with speech and language deficits. If this comorbidity occurred

in a recurring fashion, across multiple loci in the genome, this would provide valuable support for the hypothesis we have proposed in this chapter.

There appears to be a nonrandom association of ectrodactyly (split hand/foot malformation, SHFM) with intellectual disability, often with pronounced speech and language deficits being singled out in clinical descriptions. This form of congenital limb malformation arises as a consequence of disruption to the apical ectodermal ridge (AER) during the outgrowth of the developing limb bud and formation of the digits (Fig. 6.5) [505, 506]. The initial establishment of the AER is not compromised and the proximal portions of the limb; the stylopod (humerus/femur) and zeugopod (radius ulna/tibia fibula) are unaffected. However, during the development of the autopod (wrist, ankle, digits), the median portion of the AER is not maintained, leading to a failure in the specification of the central digits. The most obvious molecular theme to emerge from studies of SHFM and other disorders featuring ectrodactyly is the involvement of the *TP63* gene (related to TP53, discussed earlier), the Wnt pathway, and distal-less homeobox genes (*DLX1*, *2*, *5*, and *6*) [506, 507]. There is a telling concordance between genes potentially involved in speech and language (some of which we have already introduced in this chapter) and those implicated in SHFM.

The *SNX3* gene has been linked to microcephaly, microphthalmia, ectrodactyly, and prognathism (MMEP), and deletions of the 6q21 region involving *SNX3* almost invariably lead to moderate-to-severe mental retardation, so in this instance speech and language are not preferentially targeted [264, 508]. In the preceding section, mention was made of *LZTS2*. It may be relevant to note also that *LZTS2* and *BTRC* (encoding β-TRCP, a key component of the β-catenin destruction complex [200]) are located together at 10q24.3 in the SHFM3 locus [509, 510]. The *CRK* gene listed in Table 6.2 also resides at a locus implicated in SHFM [511–513] and has been identified alongside Wnt pathway components in exosomes produced from some cell types [282].

Connecting speech to SHFM, porcupine (*PORCN*), encoding the acyltransferase that palmitoylates Wnt ligands, is causative of focal dermal hypoplasia/Goltz syndrome and sometimes associated with SHFM [514–516]. As mentioned earlier, small deletions at Xp11.23 involving *PORCN* result in speech deficits [382]. Recessive Robinow syndrome results from mutation of the *ROR2* gene and is occasionally accompanied by SHFM [506]. In the section "Involvement of the *WNT* Pathway in Birdsong and Speech," we introduced ROR2 as a receptor for Wnt5a and deletion of a region including *ROR2* results in dysarthria [171]. Additionally, the gene encoding the kinase TGF-beta-activated kinase 1 (*TAK1*/*MAP3K7*), interacting with ROR2 [517], represents an interesting candidate for absent speech but intact receptive language capabilities (although the deletions in this patient encompass many other genes as well) [518, 519].

In this chapter, we introduced the possibility that regulators of cytokinesis and MVB formation may contribute to speech and language. In this respect, it may be noteworthy that an *EPS15*-related gene, *EPS15L1*, is a strong candidate for SHFM at 19p13.11 [520, 521]. EPS15 is associated with MVBs [398], and *EPS1* and *EPS15L1* are highly homologous and might be able to compensate for one another

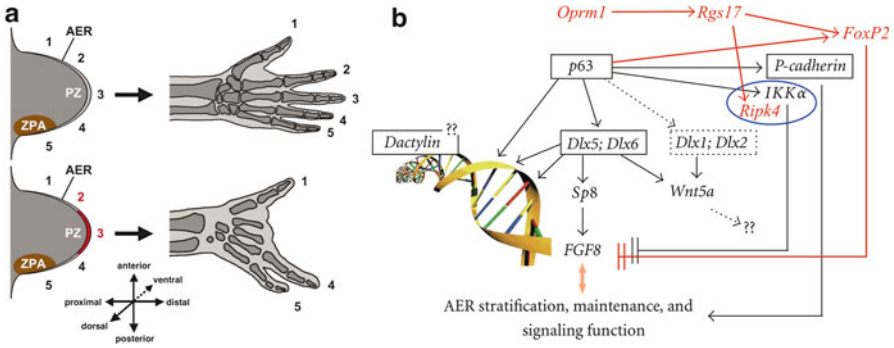


Fig. 6.5 Speech and maintenance of the limb bud medial apical ectodermal ridge may involve similar molecular pathways. In (**a**) the morphogenetic basis for the ectrodactyly phenotype stems from a deficiency in the maintenance of the median AER required for the development of the central digits. The *top two panels* show normal development of the autopod, with the position of presumptive digit formation (at *left*, numbered 1–5) relative to the AER located at the distal tip of the limb bud. Reciprocal interactions between the AER, the zone of polarizing activity (ZPA) and the progress zone (PZ) coordinate the specification of positional information, differentiation, and growth. The *bottom* part of the figure shows the ectrodactyly malformation arising by failure to maintain the median AER (shown in *red*), leading to the absence of the central rays (digits 2 & 3). (**b**) Schematic representation of p63-centered pathways relevant to ectrodactyly (shown in *black*). Genes for which there is strong evidence of involvement in ectrodactyly are framed in *solid boxes*, and putative disease genes (*Dlx1*/*Dlx2*) are framed by a *dashed box*. Similarly, highly probable interactions are given by *solid lines* and putative regulatory interactions by *dashed lines*. Fibroblast growth factor 8 (FGF8) is an essential signaling molecule produced by the AER and required for its maintenance and signaling function (this mutual dependence is indicated by the *double arrow*). This pathway, in whole or part, may also be active in the striatum during vocal production, additionally engaging *FoxP2* and other genes mentioned in this chapter (as shown in *red*). In the pathway depicted, singing-regulated expression of the endogenous opioid enkephalin could signal through the μ-opioid receptor to stimulate *Rgs17* activity. In tumor cells, *Rgs17* has been demonstrated to activate the cAMP–PKA–CREB pathway and, through this mechanism, *FoxP2* [450]. *FoxP2* is also likely to be transcriptionally regulated by p63 [534]. *Rgs17* signaling also leads, via pCREB, to *Ripk4* transcriptional upregulation. *Ripk4* is encircled in *blue* with *IKKα*, as both cause pterygium syndromes (moreover, the overlap of disease phenotypes caused by *Ripk4* and *TP63* mutations is noted in section "Going Out on a Limb"). Most interestingly, *FGF8* was one of the very few genes identified by Spiteri et al. [108] as being bound by FOXP2 in ChIP-chip assays from both human fetal basal ganglia and inferior frontal cortex tissue. Whether FOXP2 represses *FGF8* transcription in the limb bud is not known. Crosstalk with Wnt signaling pathways could occur at multiple points in the schematic, but are omitted for clarity. In (**a**) is reprinted with permission from Duijf PH, van Bokhoven H, Brunner HG. Pathogenesis of split-hand/split-foot malformation. Hum Mol Genet. 2003 Apr 1;12 Spec No 1:R51–60 [506]. In (**b**) is reprinted from Guerrini L, Costanzo A, Merlo GR. A symphony of regulations centered on p63 to control development of ectoderm-derived structures. J Biomed Biotechnol. 2011;2011:864904 [507] with modifications, under Creative Commons license. Copyright © 2011 Luisa Guerrini et al.

in the brain [522]. In the case study by Aten et al., the affected boy was initially suspected on having Angelman syndrome, with speech and language being severely impaired (4 to 5 words by age 6) [520]. The girl studied by Bens and colleagues displayed milder speech impairment [521].

The p53-related gene, *TP63*, is implicated in Hay–Wells syndrome and is the causative agent at the SHFM4 locus [523]. TP63 protein functions as an antagonist to WNT-induced transcription [524]. Hay–Wells shares some phenotypic overlap with Bartsocas–Papas syndrome [525], which has been identified as caused by mutations in the *RIPK4* gene [526]. The molecular basis for this phenotypic overlap likely reflects the fact that RIPK4 functions in the same pathway, as it is transcriptionally regulated by p63 and phosphorylates DVL2 in canonical WNT signaling (also see Fig. 6.5) [526, 527]. A child with features of Hay–Wells syndrome was recently found to carry a homozygous mutation resulting in a likely pathogenic single amino acid change RIPK4. Of interest, although her cognitive development was quite normal, the clinical description includes articulatory problems and possible oromotor speech disorder [525].

It is also noteworthy that in the case of the CAS patients studied by Laffin et al. [210], one of the subjects had a deletion involving the *DLX1* and *DLX2* genes at 2q31. The *DLX1* and *DLX2* genes have been offered as candidates for SHFM5, although there is some disagreement [528, 529]. In contrast, there is convincing evidence to suggest that the *DLX5* and *DLX6* (or perhaps DSS1) genes are responsible for SHFM1 at 7q21.3 [530, 531]. Disruption of *DLX5*/*DLX6* does not appear to typically result in severe speech deficits, although hearing loss and developmental delay are sometimes reported, as is autism [532, 533]. What is rather fascinating from the mechanistic perspective is that FoxP2 strongly regulates the expression of the large noncoding RNA in the region, *Shhrs*/*Evf-2* [109], which cooperates with the DLX2 protein to regulate a *DLX5*/*6* gene enhancer element [534].

Motivated by the accumulated evidence in favor of a role for *TP63* in SHFM, Kouwenhoven and colleagues conducted a genome-wide ChIP-seq analysis of p63 binding sites in primary human keratinocytes [535]. Examining a subset of loci selected for their association with SHFM, they identified p63 binding sites in the *TP63* gene itself, as well as proximal to the *SNX3* and *PORCN* genes and in the *DLX5*/*DLX6*/*DSS1* (SHFM1) locus. Most surprisingly, excepting *DSS1*, the gene with the most p63 binding sites in its vicinity was *FOXP2* [535]. It would be most interesting to know whether p63 also regulates *FOXP2* (and other genes in Table 6.2) in Area X and VSP during the act of singing. Moreover, FOXP2 prominently binds the *FGF8* promoter region in basal ganglia and inferior frontal cortex [108], and FGF8 is essential for AER growth regulatory and patterning functions [536, 537].

In the aggregate, the findings discussed above suggest shared molecular pathways utilized in maintenance of the median AER and events happening in the striatum during vocal production. In the developing limb, a major role has emerged for *Wnt5*/*ROR2* in regulating a PCP pathway that controls the orientation of cell division and cell migration in the limb bud [505]. Numerous studies have implicated the PCP pathway in axonal and dendritic morphogenesis. Of more immediate interest, a revisiting of the work of Konopka et al. [107] reveals that some of the FOXP2 and FOXP2^{chimp} differentially regulated "hub" genes identified also fit with the story developed above, in particular *ROR2* and *DLX5* (both SHFM candidates). Almost mystically, it is interesting to contemplate the parallelism between the putative

shared molecular substructure of limb development and learned vocalization, as adumbrated here, and the notion that language may have evolved from a gestural communication system [538].

Relevance of the Proposed Model to Autism

Summarizing our central hypothesis, we propose that the core molecular deficit in multiple disorders of speech and language development (and many cases of autism) is one of defective Wnt signaling (perhaps specifically Wnt5a/Wnt5b). Transsynaptic delivery of WNT ligands likely occurs through a specialized mechanism involving MVB formation and exosomal release. In support of our model, we note the three following points. (1) In songbirds, *FoxP2* module genes modulated by the act of singing include both *Wnt5* family members and several genes involved in MVB formation (or its proxy in this argument, midbody formation during cytokinesis). These include *IQGAP1*, the endophilins, and *TBC1D15*. In Table 6.2, numerous other Wnt pathway genes are also represented. (2) We have highlighted from the human genetics literature numerous cases of speech delay and/or autism (frequently both) that can be mechanistically united based on this posited reliance on Wnt signaling. (3) We have explored possible molecular contributions to three syndromes of intense research interest and for which profound speech delays are a common feature. Remarkably, we find manifestations in the genome leading us to propose that Angelman syndrome, PTLS, and Williams (dup7) syndrome compromise MVB formation and exosomal Wnt release (whether occurring pre- or postsynaptically, or both, is not certain). Moreover, the involvement of the same vesicular trafficking process is also supported by genes in the vicinity of CNVs identified by Shriberg and colleagues in non-syndromic forms of CAS [210].

The hypothesis we have framed here is the most explicit attempt yet to delineate the core molecular processes underlaying the production of speech and language. Given the frequent overlap of genomic regions implicated in speech and language disorders and those implicated in autism, it seems reasonable to propose that some of the genetic complexity of autism might similarly arise from disruption to the complex process of Wnt ligand synthesis, packaging, synaptic release, and signal transduction. Guided by the songbird studies of Hilliard et al. [111, 113], we place the emphasis on synaptic signaling during the production of learned social behavior, rather than embryonic/fetal deficits in establishing neural connectivity. However, it should be quite obvious that highly similar processes will be involved during neural development. Indeed, the utilization of the same Wnt signaling mechanisms may underlay often reported neuroanatomical abnormalities in autism. For example, WNT5A/RYK activity regulates neuron fasciculation and guidance across the corpus callosum [539, 540] and so potentially could contribute to inter- and intra-hemispheric connectional alterations detected in the cortex of people with autism [541]. Similarly, Wnt/β-catenin signaling is critical to cerebellar development [186, 542] and could underlay the structural/cellular deficits often observed in the cerebellum of autistic subjects [543].

Nonneural structural biases in people with autism may also reflect their differential responsivity to the signaling pathways we have highlighted here. Perhaps most significant of these is 2D:4D digit ratio (mentioned earlier with respect to contentious dyslexia data). There has been much discussion centered around the finding that autistic subjects have an altered 2D:4D digit ratio relative to the normal population, an observation that has been interpreted as evidence that in utero exposure to steroidal hormones influences the risk for developing autism [544–547]. It is instructive to consider which genes function as key regulators in establishing the 2D:4D ratio during development of the autopod. A recent study examining how the 2D:4D ratio is affected by androgen and estrogen levels in neonatal mice suggests that Wnt5a might be a key component of translating these steroidal signals into differential digit growth [548]. The assumption continues to be that, during the narrow window of developmental time that digit ratio is specified, prenatal androgens and estrogens are influencing brain development, producing the observed correlations between 2D:4D ratio and a wide variety of eventual behavioral phenotypes. An alternative interpretation, suggested by the model we have elaborated in this chapter, is that the 2D:4D ratio indeed reflects the individual's genomic and cellular responsivity to steroids and Wnt5 signaling while in utero. However, we posit it is that same idiosyncratic responsivity to those very same signals (estrogen and Wnt) occurring at the synapse, pre-, postnatally, and into adulthood, which has an additional (and perhaps the dominant) effect on the behavioral phenotype of the individual. As indicated in section "Estrogenic Contributions to Perception, Vocal Production, and…Literacy," the 16p11.2 region might have a role to play in setting this responsivity, with implications for some conceptualizations of autism [179, 544].

The point here is that if developmental miswiring is not to blame, deficits in moment-to-moment synaptic WNT signaling in autism or speech and language deficits (e.g., CAS) may obtain, perhaps making these conditions more amenable to pharmacological strategies than otherwise anticipated. Above, we mentioned the impact aberrant Wnt signaling might have on the development of the cerebellum. This could, for example, compromise olivocerebellar connections involved in regulation of gaze [549], thereby providing a neuroanatomical basis for the altered gaze preferences in autistic subjects, as discussed earlier [422, 423]. However, arguing for the existence of important postnatal processes, neuroanatomic changes in the inferior olive of autistic subjects have been described as progressive [550]. Cerebellar vermal lobules VI and VII receive massive projections from the inferior olive, and whereas theses vermal lobules are hypoplastic in autism, they are hyperplastic in Williams syndrome [551] (the neuroanatomical condition of the inferior olive in Williams syndrome warrants attention). Both *FOXP1* and *FOXP2* are expressed in the inferior olive [102, 552], and the former, at least, has been identified as contributing to autism [106]. Bringing these pieces of evidence together, and arguing by analogy from the songbird experiments of Hilliard and colleagues, we propose that it may be valuable to test whether *FoxP1* and *FoxP2* (and other Table 6.2 genes) expression in the inferior olive might be modulated by visual/social experience or during the socially directed commission of behavior (such as fixation of gaze on the face or eyes).

Taken together, the observations made in this chapter provide a framework that will hopefully prove useful in working toward an understanding of the molecular and genetic control of speech and language. Other complex social behaviors might also be founded on the same molecular mechanisms, and vocal learning is but one form of imitation engaging the striatum [553]. Deficits in imitative learning are suggested to contribute to autism [554, 555], so a shared genetic liability with speech and language deficits seems parsimonious. The next few years of research will either confirm or reject our hypothesis, which we also hope will prompt other researchers to develop alternative explanations that have similar reach. Research in the songbird could have a major role to play in testing such ideas.

References

1. Krough A (1929) The progress of physiology. Am J Physiol 90(2):243–251
2. Krebs HA (1975) The August Krogh principle: "For many problems there is an animal on which it can be most conveniently studied". J Exp Zool 194(1):221–226
3. Morris D (1954) The reproductive behaviour of the zebra finch (Poephila guttata), with special reference to pseudofemale behaviour and displacement activities. Behaviour 6:271–322
4. Goldman SA, Nottebohm F (1983) Neuronal production, migration, and differentiation in a vocal control nucleus of the adult female canary brain. Proc Natl Acad Sci U S A 80(8):2390–2394
5. Barnea A (2010) Wild neurogenesis. Brain Behav Evol 75(2):86–87
6. Pytte CL, George S, Korman S, David E, Bogdan D, Kirn JR (2012) Adult neurogenesis is associated with the maintenance of a stereotyped, learned motor behavior. J Neurosci 32(20):7052–7057
7. Brainard MS, Doupe AJ (2002) What songbirds teach us about learning. Nature 417(6886):351–358
8. Kao MH, Doupe AJ, Brainard MS (2005) Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. Nature 433(7026):638–643
9. 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
10. Keller GB, Hahnloser RH (2009) Neural processing of auditory feedback during vocal practice in a songbird. Nature 457(7226):187–190
11. Dave AS, Margoliash D (2000) Song replay during sleep and computational rules for sensorimotor vocal learning. Science 290(5492):812–816
12. Deregnaucourt S, Mitra PP, Feher O, Pytte C, Tchernichovski O (2005) How sleep affects the developmental learning of bird song. Nature 433(7027):710–716
13. Nottebohm F, Arnold AP (1976) Sexual dimorphism in vocal control areas of songbird brain. Science 194(4261):211–213
14. Agate RJ, Grisham W, Wade J, Mann S, Wingfield J, Schanen C et al (2003) Neural, not gonadal, origin of brain sex differences in a gynandromorphic finch. Proc Natl Acad Sci U S A 100(8):4873–4878
15. Gil D, Graves J, Hazon N, Wells A (1999) Male attractiveness and differential testosterone investment in Zebra Finch eggs. Science 286(5437):126–128
16. Blount JD, Metcalfe NB, Birkhead TR, Surai PF (2003) Carotenoid modulation of immune function and sexual attractiveness in zebra finches. Science 300(5616):125–127
17. Goodson JL, Schrock SE, Klatt JD, Kabelik D, Kingsbury MA (2009) Mesotocin and nonapeptide receptors promote estrildid flocking behavior. Science 325(5942):862–866
18. 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(6697):67–71

19. Tchernichovski O, Mitra PP, Lints T, Nottebohm F (2001) Dynamics of the vocal imitation process: how a zebra finch learns its song. Science 291(5513):2564–2569
20. Vignal C, Mathevon N, Mottin S (2004) Audience drives male songbird response to partner's voice. Nature 430(6998):448–451
21. Feher O, Wang H, Saar S, Mitra PP, Tchernichovski O (2009) De novo establishment of wild-type song culture in the zebra finch. Nature 459(7246):564–568
22. Warren WC, Clayton DF, Ellegren H, Arnold AP, Hillier LW, Kunstner A et al (2010) The genome of a songbird. Nature 464(7289):757–762
23. Scharff C, Adam I (2013) Neurogenetics of birdsong. Curr Opin Neurobiol 23(1):29–36
24. 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
25. Forstmeier W, Burger C, Temnow K, Deregnaucourt S (2009) The genetic basis of zebra finch vocalizations. Evolution 63(8):2114–2130
26. Okanoya K, Dooling RJ (1987) Strain differences in auditory thresholds in the canary (Serinus canarius). J Comp Psychol 101(2):213–215
27. Wright TF, Brittan-Powell EF, Dooling RJ, Mundinger PC (2004) Sex-linked inheritance of hearing and song in the Belgian Waterslager canary. Proc Biol Sci 271(suppl 6):S409–S412
28. Brittan-Powell EF, Dooling RJ, Ryals B, Gleich O (2010) Electrophysiological and morphological development of the inner ear in Belgian Waterslager canaries. Hear Res 269(1–2):56–69
29. Bolund E, Martin K, Kempenaers B, Forstmeier W (2010) Inbreeding depression of sexually selected traits and attractiveness in the zebra finch. Anim Behav 79(4):947–955
30. Hemmings NL, Slate J, Birkhead TR (2012) Inbreeding causes early death in a passerine bird. Nat Commun 3:863
31. Backstrom N, Forstmeier W, Schielzeth H, Mellenius H, Nam K, Bolund E et al (2010) The recombination landscape of the zebra finch Taeniopygia guttata genome. Genome Res 20(4):485–495
32. Oleksyk TK, Pombert JF, Siu D, Mazo-Vargas A, Ramos B, Guiblet W et al (2012) A locally funded Puerto Rican parrot (Amazona vittata) genome sequencing project increases avian data and advances young researcher education. Gigascience 1(1):14
33. Zhang G, Cowled C, Shi Z, Huang Z, Bishop-Lilly KA, Fang X et al (2013) Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. Science 339(6118):456–460
34. Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ, Washietl S et al (2011) A high-resolution map of human evolutionary constraint using 29 mammals. Nature 478(7370): 476–482
35. Sun YB, Zhou WP, Liu HQ, Irwin DM, Shen YY, Zhang YP (2013) Genome-wide scans for candidate genes involved in the aquatic adaptation of dolphins. Genome Biol Evol 5(1):130–139
36. White SA, Fisher SE, Geschwind DH, Scharff C, Holy TE (2006) Singing mice, songbirds, and more: models for FOXP2 function and dysfunction in human speech and language. J Neurosci 26(41):10376–10379
37. Bolhuis JJ, Okanoya K, Scharff C (2010) Twitter evolution: converging mechanisms in birdsong and human speech. Nat Rev Neurosci 11(11):747–759
38. Nottebohm F, Stokes TM, Leonard CM (1976) Central control of song in canary, Serinus-Canarius. J Comp Neurol 165(4):457–486
39. Long MA, Fee MS (2008) Using temperature to analyse temporal dynamics in the songbird motor pathway. Nature 456(7219):189–194
40. Jarvis ED, Gunturkun O, Bruce L, Csillag A, Karten H, Kuenzel W et al (2005) Avian brains and a new understanding of vertebrate brain evolution. Nat Rev Neurosci 6(2):151–159
41. Yu AC, Margoliash D (1996) Temporal hierarchical control of singing in birds. Science 273(5283):1871–1875
42. Leonardo A, Fee MS (2005) Ensemble coding of vocal control in birdsong. J Neurosci 25(3):652–661

43. Reiner A, Perkel DJ, Bruce LL, Butler AB, Csillag A, Kuenzel W et al (2004) Revised nomenclature for avian telencephalon and some related brainstem nuclei. J Comp Neurol 473(3):377–414
44. Doupe AJ, Perkel DJ, Reiner A, Stern EA (2005) Birdbrains could teach basal ganglia research a new song. Trends Neurosci 28(7):353–363
45. Graybiel AM (1998) The basal ganglia and chunking of action repertoires. Neurobiol Learn Mem 70(1–2):119–136
46. Graybiel AM (2004) Network-level neuroplasticity in cortico-basal ganglia pathways. Parkinsonism Relat Disord 10(5):293–296
47. Maia TV, Frank MJ (2011) From reinforcement learning models to psychiatric and neurological disorders. Nat Neurosci 14(2):154–162
48. Liljeholm M, O'Doherty JP (2012) Contributions of the striatum to learning, motivation, and performance: an associative account. Trends Cogn Sci 16(9):467–475
49. Feenders G, Liedvogel M, Rivas M, Zapka M, Horita H, Hara E et al (2008) Molecular mapping of movement-associated areas in the avian brain: a motor theory for vocal learning origin. PLoS One 3(3):e1768
50. Bottjer SW, Miesner EA, Arnold AP (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. Science 224(4651):901–903
51. Garcia-Calero E, Scharff C (2013) Calbindin expression in developing striatum of zebra finches and its relation to the formation of area X. J Comp Neurol 521(2):326–341
52. Jarvis ED, Nottebohm F (1997) Motor-driven gene expression. Proc Natl Acad Sci U S A 94(8):4097–4102
53. 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(9):3776–3787
54. 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(20):7088–7098
55. Carrillo GD, Doupe AJ (2004) Is the songbird Area X striatal, pallidal, or both? An anatomical study. J Comp Neurol 473(3):415–437
56. 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(16):5730–5743
57. Luo M, Perkel DJ (1999) A GABAergic, strongly inhibitory projection to a thalamic nucleus in the zebra finch song system. J Neurosci 19(15):6700–6711
58. Thorn CA, Atallah H, Howe M, Graybiel AM (2010) Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. Neuron 66(5):781–795
59. 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
60. Bottjer SW, Altenau B (2010) Parallel pathways for vocal learning in basal ganglia of songbirds. Nat Neurosci 13(2):153–155
61. Olveczky BP (2011) Motoring ahead with rodents. Curr Opin Neurobiol 21(4):571–578
62. 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
63. 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
64. Aronov D, Andalman AS, Fee MS (2008) A specialized forebrain circuit for vocal babbling in the juvenile songbird. Science 320(5876):630–634
65. Olveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS Biol 3(5):e153

66. 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
67. Veit L, Aronov D, Fee MS (2011) Learning to breathe and sing: development of respiratory-vocal coordination in young songbirds. J Neurophysiol 106(4):1747–1765
68. 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
69. Hessler NA, Doupe AJ (1999) Social context modulates singing-related neural activity in the songbird forebrain. Nat Neurosci 2(3):209–211
70. Kojima S, Doupe AJ (2011) Social performance reveals unexpected vocal competency in young songbirds. Proc Natl Acad Sci U S A 108(4):1687–1692
71. 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(49):13232–13247
72. Stepanek L, Doupe AJ (2010) Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. J Neurophysiol 104(5):2474–2486
73. Kimpo RR, Doupe AJ (1997) FOS is induced by singing in distinct neuronal populations in a motor network. Neuron 18(2):315–325
74. Kubikova L, Turner EA, Jarvis ED (2007) The pallial basal ganglia pathway modulates the behaviorally driven gene expression of the motor pathway. Eur J Neurosci 25(7):2145–2160
75. Haesler S, Wada K, Nshdejan A, Morrisey EE, Lints T, Jarvis ED et al (2004) FoxP2 expression in avian vocal learners and non-learners. J Neurosci 24(13):3164–3175
76. 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
77. Takahashi K, Liu FC, Hirokawa K, Takahashi H (2003) Expression of Foxp2, a gene involved in speech and language, in the developing and adult striatum. J Neurosci Res 73(1):61–72
78. Habas C, Guillevin R, Abanou A (2010) In vivo structural and functional imaging of the human rubral and inferior olivary nuclei: a mini-review. Cerebellum 9(2):167–173
79. Llinas RR (2011) Cerebellar motor learning versus cerebellar motor timing: the climbing fibre story. J Physiol 589(Pt 14):3423–3432
80. Rauschecker AM, Pringle A, Watkins KE (2008) Changes in neural activity associated with learning to articulate novel auditory pseudowords by covert repetition. Hum Brain Mapp 29(11):1231–1242
81. Watkins KE, Smith SM, Davis S, Howell P (2008) Structural and functional abnormalities of the motor system in developmental stuttering. Brain 131(Pt 1):50–59
82. Hicks TP, Onodera S (2012) The mammalian red nucleus and its role in motor systems, including the emergence of bipedalism and language. Prog Neurobiol 96(2):165–175
83. Gale SD, Perkel DJ (2010) Anatomy of a songbird basal ganglia circuit essential for vocal learning and plasticity. J Chem Neuroanat 39(2):124–131
84. Teramitsu I, Poopatanapong A, Torrisi S, White SA (2010) Striatal FoxP2 is actively regulated during songbird sensorimotor learning. PLoS One 5(1):e8548
85. Roper A, Zann R (2006) The onset of song learning and song tutor selection in fledgling zebra finches. Ethology 112(5):458–470
86. Bottjer SW, Glaessner SL, Arnold AP (1985) Ontogeny of brain nuclei controlling song learning and behavior in zebra finches. J Neurosci 5(6):1556–1562
87. Nordeen EJ, Nordeen KW (1988) Sex and regional differences in the incorporation of neurons born during song learning in zebra finches. J Neurosci 8(8):2869–2874
88. Rochefort C, He X, Scotto-Lomassese S, Scharff C (2007) Recruitment of FoxP2-expressing neurons to area X varies during song development. Dev Neurobiol 67(6):809–817
89. Thompson CK, Schwabe F, Schoof A, Mendoza E, Gampe J, Rochefort C et al (2013) Young and intense: FoxP2 immunoreactivity in Area X varies with age, song stereotypy, and singing in male zebra finches. Front Neural Circuits 7:24
90. Teramitsu I, White SA (2006) FoxP2 regulation during undirected singing in adult songbirds. J Neurosci 26(28):7390–7394

91. Miller JE, Spiteri E, Condro MC, Dosumu-Johnson RT, Geschwind DH, White SA (2008) Birdsong decreases protein levels of FoxP2, a molecule required for human speech. J Neurophysiol 100(4):2015–2025
92. Miller JE, Hilliard AT, White SA (2010) Song practice promotes acute vocal variability at a key stage of sensorimotor learning. PLoS One 5(1):e8592
93. 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
94. Tchernichovski O, Nottebohm F (1998) Social inhibition of song imitation among sibling male zebra finches. Proc Natl Acad Sci U S A 95(15):8951–8956
95. Schulz SB, Haesler S, Scharff C, Rochefort C (2010) Knockdown of FoxP2 alters spine density in Area X of the zebra finch. Genes Brain Behav 9(7):732–740
96. Foster EF, Bottjer SW (1998) Axonal connections of the high vocal center and surrounding cortical regions in juvenile and adult male zebra finches. J Comp Neurol 397(1):118–138
97. Alarcon M, Abrahams BS, Stone JL, Duvall JA, Perederiy JV, Bomar JM et al (2008) Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. Am J Hum Genet 82(1):150–159
98. Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, Groszer M et al (2008) A functional genetic link between distinct developmental language disorders. N Engl J Med 359(22):2337–2345
99. 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
100. Wang B, Lin D, Li C, Tucker P (2003) Multiple domains define the expression and regulatory properties of Foxp1 forkhead transcriptional repressors. J Biol Chem 278(27):24259–24268
101. Li S, Weidenfeld J, Morrisey EE (2004) Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions. Mol Cell Biol 24(2):809–822
102. Ferland RJ, Cherry TJ, Preware PO, Morrisey EE, Walsh CA (2003) Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain. J Comp Neurol 460(2):266–279
103. Hamdan FF, Daoud H, Rochefort D, Piton A, Gauthier J, Langlois M et al (2010) De novo mutations in FOXP1 in cases with intellectual disability, autism, and language impairment. Am J Hum Genet 87(5):671–678
104. Horn D, Kapeller J, Rivera-Brugues N, Moog U, Lorenz-Depiereux B, Eck S et al (2010) Identification of FOXP1 deletions in three unrelated patients with mental retardation and significant speech and language deficits. Hum Mutat 31(11):E1851–E1860
105. Palumbo O, D'Agruma L, Minenna AF, Palumbo P, Stallone R, Palladino T et al (2013) 3p14.1 de novo microdeletion involving the FOXP1 gene in an adult patient with autism, severe speech delay and deficit of motor coordination. Gene 516(1):107–113
106. Bacon C, Rappold GA (2012) The distinct and overlapping phenotypic spectra of FOXP1 and FOXP2 in cognitive disorders. Hum Genet 131(11):1687–1698
107. Konopka G, Bomar JM, Winden K, Coppola G, Jonsson ZO, Gao F et al (2009) Human-specific transcriptional regulation of CNS development genes by FOXP2. Nature 462(7270):213–217
108. Spiteri E, Konopka G, Coppola G, Bomar J, Oldham M, Ou J et al (2007) Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. Am J Hum Genet 81(6):1144–1157
109. Vernes SC, Oliver PL, Spiteri E, Lockstone HE, Puliyadi R, Taylor JM et al (2011) Foxp2 regulates gene networks implicated in neurite outgrowth in the developing brain. PLoS Genet 7(7):e1002145
110. Enard W, Gehre S, Hammerschmidt K, Holter SM, Blass T, Somel M et al (2009) A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. Cell 137(5):961–971

111. Hilliard AT, Miller JE, Fraley ER, Horvath S, White SA (2012) Molecular microcircuitry underlies functional specification in a basal ganglia circuit dedicated to vocal learning. Neuron 73(3):537–552
112. Zhao W, Langfelder P, Fuller T, Dong J, Li A, Hovarth S (2010) Weighted gene coexpression network analysis: state of the art. J Biopharm Stat 20(2):281–300
113. Hilliard AT, Miller JE, Horvath S, White SA (2012) Distinct neurogenomic states in basal ganglia subregions relate differently to singing behavior in songbirds. PLoS Comput Biol 8(11):e1002773
114. Williams H (2001) Choreography of song, dance and beak movements in the zebra finch (Taeniopygia guttata). J Exp Biol 204(Pt 20):3497–3506
115. Drnevich J, Replogle KL, Lovell P, Hahn TP, Johnson F, Mast TG et al (2012) Impact of experience-dependent and -independent factors on gene expression in songbird brain. Proc Natl Acad Sci U S A 109(suppl 2):17245–17252
116. Lodewyckx L, Cailotto F, Thysen S, Luyten FP, Lories RJ (2012) Tight regulation of wingless-type signaling in the articular cartilage—subchondral bone biomechanical unit: transcriptomics in Frzb-knockout mice. Arthritis Res Ther 14(1):R16
117. Desplats PA, Lambert JR, Thomas EA (2008) Functional roles for the striatal-enriched transcription factor, Bcl11b, in the control of striatal gene expression and transcriptional dysregulation in Huntington's disease. Neurobiol Dis 31(3):298–308
118. Bonnet C, Andrieux J, Beri-Dexheimer M, Leheup B, Boute O, Manouvrier S et al (2010) Microdeletion at chromosome 4q21 defines a new emerging syndrome with marked growth restriction, mental retardation and absent or severely delayed speech. J Med Genet 47(6):377–384
119. Rosenstierne MW, Vinther J, Mittler G, Larsen L, Mann M, Norrild B (2008) Conserved CPEs in the p53 3′ untranslated region influence mRNA stability and protein synthesis. Anticancer Res 28(5A):2553–2559
120. White EJ, Brewer G, Wilson GM (2013) Post-transcriptional control of gene expression by AUF1: mechanisms, physiological targets, and regulation. Biochim Biophys Acta 1829(6–7):680–688
121. Kantamneni S, Holman D, Wilkinson KA, Correa SA, Feligioni M, Ogden S et al (2008) GISP binding to TSG101 increases GABA receptor stability by down-regulating ESCRT-mediated lysosomal degradation. J Neurochem 107(1):86–95
122. Soliman MA, Riabowol K (2007) After a decade of study-ING, a PHD for a versatile family of proteins. Trends Biochem Sci 32(11):509–519
123. Thalappilly S, Feng X, Pastyryeva S, Suzuki K, Muruve D, Larocque D et al (2011) The p53 tumor suppressor is stabilized by inhibitor of growth 1 (ING1) by blocking polyubiquitination. PLoS One 6(6):e21065
124. Quemener-Redon S, Benech C, Audebert-Bellanger S, Friocourt G, Planes M, Parent P et al (2013) A small de novo 16q24.1 duplication in a woman with severe clinical features. Eur J Med Genet 56(4):211–215
125. Paschou P, Stylianopoulou E, Karagiannidis I, Rizzo R, Tarnok Z, Wolanczyk T et al (2012) Evaluation of the LIM homeobox genes LHX6 and LHX8 as candidates for Tourette syndrome. Genes Brain Behav 11(4):444–451
126. Zhao Y, Marin O, Hermesz E, Powell A, Flames N, Palkovits M et al (2003) The LIM-homeobox gene Lhx8 is required for the development of many cholinergic neurons in the mouse forebrain. Proc Natl Acad Sci U S A 100(15):9005–9010
127. Abdelmoity AT, Hall JJ, Bittel DC, Yu S (2011) 1.39 Mb inherited interstitial deletion in 12p13.33 associated with developmental delay. Eur J Med Genet 54(2):198–203
128. Tymanskyj SR, Lin S, Gordon-Weeks PR (2010) Evolution of the spatial distribution of MAP1B phosphorylation sites in vertebrate neurons. J Anat 216(6):692–704
129. Shao M, Lin Y, Liu Z, Zhang Y, Wang L, Liu C et al (2012) GSK-3 activity is critical for the orientation of the cortical microtubules and the dorsoventral axis determination in zebrafish embryos. PLoS One 7(5):e36655

130. Veerappa AM, Saldanha M, Padakannaya P, Ramachandra NB (2013) Family-based genome-wide copy number scan identifies five new genes of dyslexia involved in dendritic spinal plasticity. J Hum Genet, May 16, 2013; doi:10.1038/jhg.2013.47
131. Hoshina N, Tanimura A, Yamasaki M, Inoue T, Fukabori R, Kuroda T et al (2013) Protocadherin 17 regulates presynaptic assembly in topographic corticobasal ganglia circuits. Neuron 78:839–854
132. Zhou L, Jones EV, Murai KK (2012) EphA signaling promotes actin-based dendritic spine remodeling through slingshot phosphatase. J Biol Chem 287(12):9346–9359
133. Hendriks WJ, Dilaver G, Noordman YE, Kremer B, Fransen JA (2009) PTPRR protein tyrosine phosphatase isoforms and locomotion of vesicles and mice. Cerebellum 8(2):80–88
134. Krieg AJ, Hammond EM, Giaccia AJ (2006) Functional analysis of p53 binding under differential stresses. Mol Cell Biol 26(19):7030–7045
135. Maiuri MC, Malik SA, Morselli E, Kepp O, Criollo A, Mouchel PL et al (2009) Stimulation of autophagy by the p53 target gene Sestrin2. Cell Cycle 8(10):1571–1576
136. Matsuki T, Matthews RT, Cooper JA, van der Brug MP, Cookson MR, Hardy JA et al (2010) Reelin and stk25 have opposing roles in neuronal polarization and dendritic Golgi deployment. Cell 143(5):826–836
137. Buonincontri R, Bache I, Silahtaroglu A, Elbro C, Nielsen AM, Ullmann R et al (2011) A cohort of balanced reciprocal translocations associated with dyslexia: identification of two putative candidate genes at DYX1. Behav Genet 41(1):125–133
138. Neilsen PM, Noll JE, Suetani RJ, Schulz RB, Al-Ejeh F, Evdokiou A et al (2011) Mutant p53 uses p63 as a molecular chaperone to alter gene expression and induce a pro-invasive secretome. Oncotarget 2(12):1203–1217
139. Brady CA, Jiang D, Mello SS, Johnson TM, Jarvis LA, Kozak MM et al (2011) Distinct p53 transcriptional programs dictate acute DNA-damage responses and tumor suppression. Cell 145(4):571–583
140. Nusse R, Varmus H (2012) Three decades of Wnts: a personal perspective on how a scientific field developed. EMBO J 31(12):2670–2684
141. Rubin JS, Bottaro DP (2007) Loss of secreted frizzled-related protein-1 expression in renal cell carcinoma reveals a critical tumor suppressor function. Clin Cancer Res 13(16):4660–4663
142. Gao C, Chen YG (2010) Dishevelled: the hub of Wnt signaling. Cell Signal 22(5):717–727
143. Gibson CJ, Gruen JR (2008) The human lexinome: genes of language and reading. J Commun Disord 41(5):409–420
144. Newbury DF, Monaco AP (2010) Genetic advances in the study of speech and language disorders. Neuron 68(2):309–320
145. Poelmans G, Buitelaar JK, Pauls DL, Franke B (2011) A theoretical molecular network for dyslexia: integrating available genetic findings. Mol Psychiatry 16(4):365–382
146. van Amerongen R, Fuerer C, Mizutani M, Nusse R (2012) Wnt5a can both activate and repress Wnt/beta-catenin signaling during mouse embryonic development. Dev Biol 369(1):101–114
147. Yang Y (2012) Wnt signaling in development and disease. Cell Biosci 2(1):14
148. Thevenon J, Callier P, Andrieux J, Delobel B, David A, Sukno S et al (2013) 12p13.33 microdeletion including ELKS/ERC1, a new locus associated with childhood apraxia of speech. Eur J Hum Genet 21(1):82–88
149. Yamamoto A, Nagano T, Takehara S, Hibi M, Aizawa S (2005) Shisa promotes head formation through the inhibition of receptor protein maturation for the caudalizing factors, Wnt and FGF. Cell 120(2):223–235
150. Xie Z, Chen Y, Li Z, Bai G, Zhu Y, Yan R et al (2011) Smad6 promotes neuronal differentiation in the intermediate zone of the dorsal neural tube by inhibition of the Wnt/beta-catenin pathway. Proc Natl Acad Sci U S A 108(29):12119–12124
151. Hu Y, Galkin AV, Wu C, Reddy V, Su AI (2011) CAFEt algorithm reveals Wnt/PCP signature in lung squamous cell carcinoma. PLoS One 6(10):e25807
152. Glesne D, Huberman E (2006) Smad6 is a protein kinase X phosphorylation substrate and is required for HL-60 cell differentiation. Oncogene 25(29):4086–4098

153. Carlin D, Sepich D, Grover VK, Cooper MK, Solnica-Krezel L, Inbal A (2012) Six3 cooperates with Hedgehog signaling to specify ventral telencephalon by promoting early expression of Foxg1a and repressing Wnt signaling. Development 139(14):2614–2624
154. Lagutin OV, Zhu CC, Kobayashi D, Topczewski J, Shimamura K, Puelles L et al (2003) Six3 repression of Wnt signaling in the anterior neuroectoderm is essential for vertebrate forebrain development. Genes Dev 17(3):368–379
155. Gerard B, Sanders MA, Visscher DW, Tait L, Shekhar MP (2012) Lysine 394 is a novel Rad6B-induced ubiquitination site on beta-catenin. Biochim Biophys Acta 1823(10): 1686–1696
156. Shekhar MP, Tait L, Gerard B (2006) Essential role of T-cell factor/beta-catenin in regulation of Rad6B: a potential mechanism for Rad6B overexpression in breast cancer cells. Mol Cancer Res 4(10):729–745
157. Shimokawa T, Furukawa Y, Sakai M, Li M, Miwa N, Lin YM et al (2003) Involvement of the FGF18 gene in colorectal carcinogenesis, as a novel downstream target of the beta-catenin/T--cell factor complex. Cancer Res 63(19):6116–6120
158. Reinhold MI, Naski MC (2007) Direct interactions of Runx2 and canonical Wnt signaling induce FGF18. J Biol Chem 282(6):3653–3663
159. Morris LG, Kaufman AM, Gong Y, Ramaswami D, Walsh LA, Turcan S et al (2013) Recurrent somatic mutation of FAT1 in multiple human cancers leads to aberrant Wnt activation. Nat Genet 45(3):253–261
160. Fukata M, Kuroda S, Nakagawa M, Kawajiri A, Itoh N, Shoji I et al (1999) Cdc42 and Rac1 regulate the interaction of IQGAP1 with beta-catenin. J Biol Chem 274(37):26044–26050
161. Sharma M, Henderson BR (2007) IQ-domain GTPase-activating protein 1 regulates beta-catenin at membrane ruffles and its role in macropinocytosis of N-cadherin and adenomatous polyposis coli. J Biol Chem 282(11):8545–8556
162. Goto T, Sato A, Shimizu M, Adachi S, Satoh K, Iemura S et al (2013) IQGAP1 functions as a modulator of dishevelled nuclear localization in Wnt signaling. PLoS One 8(4):e60865
163. Liu CF, Parker K, Yao HH (2010) WNT4/beta-catenin pathway maintains female germ cell survival by inhibiting activin betaB in the mouse fetal ovary. PLoS One 5(4):e10382
164. Wuebben EL, Mallanna SK, Cox JL, Rizzino A (2012) Musashi2 is required for the self-renewal and pluripotency of embryonic stem cells. PLoS One 7(4):e34827
165. Lee HS, Park MH, Yang SJ, Park KC, Kim NS, Kim YS et al (2007) Novel candidate targets of Wnt/beta-catenin signaling in hepatoma cells. Life Sci 80(7):690–698
166. Sese M, Corominas M, Stocker H, Heino TI, Hafen E, Serras F (2006) The Cdi/TESK1 kinase is required for Sevenless signaling and epithelial organization in the Drosophila eye. J Cell Sci 119(Pt 24):5047–5056
167. Klapholz-Brown Z, Walmsley GG, Nusse YM, Nusse R, Brown PO (2007) Transcriptional program induced by Wnt protein in human fibroblasts suggests mechanisms for cell cooperativity in defining tissue microenvironments. PLoS One 2(9):e945
168. Vernes SC, Spiteri E, Nicod J, Groszer M, Taylor JM, Davies KE et al (2007) High-throughput analysis of promoter occupancy reveals direct neural targets of FOXP2, a gene mutated in speech and language disorders. Am J Hum Genet 81(6):1232–1250
169. Bonkowsky JL, Wang X, Fujimoto E, Lee JE, Chien CB, Dorsky RI (2008) Domain-specific regulation of foxP2 CNS expression by lef1. BMC Dev Biol 8:103
170. O'Connell MP, Weeraratna AT (2009) Hear the Wnt Ror: how melanoma cells adjust to changes in Wnt. Pigment Cell Melanoma Res 22(6):724–739
171. Siggberg L, Peippo M, Sipponen M, Miikkulainen T, Shimojima K, Yamamoto T et al (2011) 9q22 Deletion—first familial case. Orphanet J Rare Dis 6:45
172. Lin PI, Chien YL, Wu YY, Chen CH, Gau SS, Huang YS et al (2012) The WNT2 gene polymorphism associated with speech delay inherent to autism. Res Dev Disabil 33(5): 1533–1540
173. Allou L, Lambert L, Amsallem D, Bieth E, Edery P, Destree A et al (2012) 14q12 and severe Rett-like phenotypes: new clinical insights and physical mapping of FOXG1-regulatory elements. Eur J Hum Genet 20(12):1216–1223

174. Kortum F, Das S, Flindt M, Morris-Rosendahl DJ, Stefanova I, Goldstein A et al (2011) The core FOXG1 syndrome phenotype consists of postnatal microcephaly, severe mental retardation, absent language, dyskinesia, and corpus callosum hypogenesis. J Med Genet 48(6):396–406
175. Danesin C, Houart C (2012) A Fox stops the Wnt: implications for forebrain development and diseases. Curr Opin Genet Dev 22(4):323–330
176. Carri AD, Onorati M, Lelos MJ, Castiglioni V, Faedo A, Menon R et al (2013) Developmentally coordinated extrinsic signals drive human pluripotent stem cell differentiation toward authentic DARPP-32+ medium-sized spiny neurons. Development 140(2):301–312
177. Crespi BJ, Crofts HJ (2012) Association testing of copy number variants in schizophrenia and autism spectrum disorders. J Neurodev Disord 4(1):15
178. Malhotra D, Sebat J (2012) CNVs: harbingers of a rare variant revolution in psychiatric genetics. Cell 148(6):1223–1241
179. Crespi B, Badcock C (2008) Psychosis and autism as diametrical disorders of the social brain. Behav Brain Sci 31(3):241–261; discussion 61–320
180. Kalkman HO (2012) A review of the evidence for the canonical Wnt pathway in autism spectrum disorders. Mol Autism 3(1):10
181. Okerlund ND, Cheyette BN (2011) Synaptic Wnt signaling-a contributor to major psychiatric disorders? J Neurodev Disord 3(2):162–174
182. Bowers JM, Konopka G (2012) The role of the FOXP family of transcription factors in ASD. Dis Markers 33(5):251–260
183. O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S et al (2011) Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. Nat Genet 43(6):585–589
184. Lancaster MA, Gopal DJ, Kim J, Saleem SN, Silhavy JL, Louie CM et al (2011) Defective Wnt-dependent cerebellar midline fusion in a mouse model of Joubert syndrome. Nat Med 17(6):726–731
185. Pei Y, Brun SN, Markant SL, Lento W, Gibson P, Taketo MM et al (2012) WNT signaling increases proliferation and impairs differentiation of stem cells in the developing cerebellum. Development 139(10):1724–1733
186. Selvadurai HJ, Mason JO (2011) Wnt/beta-catenin signalling is active in a highly dynamic pattern during development of the mouse cerebellum. PLoS One 6(8):e23012
187. French CA, Groszer M, Preece C, Coupe AM, Rajewsky K, Fisher SE (2007) Generation of mice with a conditional Foxp2 null allele. Genesis 45(7):440–446
188. Fujita E, Tanabe Y, Shiota A, Ueda M, Suwa K, Momoi MY et al (2008) Ultrasonic vocalization impairment of Foxp2 (R552H) knockin mice related to speech-language disorder and abnormality of Purkinje cells. Proc Natl Acad Sci U S A 105(8):3117–3122
189. Fatemi SH, Aldinger KA, Ashwood P, Bauman ML, Blaha CD, Blatt GJ et al (2012) Consensus paper: pathological role of the cerebellum in autism. Cerebellum 11(3):777–807
190. McGonigle-Chalmers M, Alderson-Day B, Fleming J, Monsen K (2013) Profound expressive language impairment in low functioning children with autism: an investigation of syntactic awareness using a computerised learning task. J Autism Dev Disord 43(9):2062–2081
191. O'Connor K (2012) Auditory processing in autism spectrum disorder: a review. Neurosci Biobehav Rev 36(2):836–854
192. Pickles A, Simonoff E, Conti-Ramsden G, Falcaro M, Simkin Z, Charman T et al (2009) Loss of language in early development of autism and specific language impairment. J Child Psychol Psychiatry 50(7):843–852
193. Ackermann H (2008) Cerebellar contributions to speech production and speech perception: psycholinguistic and neurobiological perspectives. Trends Neurosci 31(6):265–272
194. Guy J, Gan J, Selfridge J, Cobb S, Bird A (2007) Reversal of neurological defects in a mouse model of Rett syndrome. Science 315(5815):1143–1147
195. McGraw CM, Samaco RC, Zoghbi HY (2011) Adult neural function requires MeCP2. Science 333(6039):186

196. Nguyen MV, Du F, Felice CA, Shan X, Nigam A, Mandel G et al (2012) MeCP2 is critical for maintaining mature neuronal networks and global brain anatomy during late stages of postnatal brain development and in the mature adult brain. J Neurosci 32(29):10021–10034
197. Silva AJ, Ehninger D (2009) Adult reversal of cognitive phenotypes in neurodevelopmental disorders. J Neurodev Disord 1(2):150–157
198. Cross I, Delhanty J, Chapman P, Bowles LV, Griffin D, Wolstenholme J et al (1992) An intrachromosomal insertion causing 5q22 deletion and familial adenomatous polyposis coli in two generations. J Med Genet 29(3):175–179
199. Hodgson SV, Coonar AS, Hanson PJ, Cottrell S, Scriven PN, Jones T et al (1993) Two cases of 5q deletions in patients with familial adenomatous polyposis: possible link with Caroli's disease. J Med Genet 30(5):369–375
200. Stamos JL, Weis WI (2013) The beta-catenin destruction complex. Cold Spring Harb Perspect Biol 5(1):a007898
201. Fukuyama R, Niculaita R, Ng KP, Obusez E, Sanchez J, Kalady M et al (2008) Mutated in colorectal cancer, a putative tumor suppressor for serrated colorectal cancer, selectively represses beta-catenin-dependent transcription. Oncogene 27(46):6044–6055
202. Ishiguro H, Furukawa Y, Daigo Y, Miyoshi Y, Nagasawa Y, Nishiwaki T et al (2000) Isolation and characterization of human NBL4, a gene involved in the beta-catenin/tcf signaling pathway. Jpn J Cancer Res 91(6):597–603
203. Pratt T, Davey JW, Nowakowski TJ, Raasumaa C, Rawlik K, McBride D et al (2012) The expression and activity of beta-catenin in the thalamus and its projections to the cerebral cortex in the mouse embryo. BMC Neurosci 13:20
204. Hu J, Prinz WA, Rapoport TA (2011) Weaving the web of ER tubules. Cell 147(6):1226–1231
205. Renvoise B, Blackstone C (2010) Emerging themes of ER organization in the development and maintenance of axons. Curr Opin Neurobiol 20(5):531–537
206. Budnik V, Salinas PC (2011) Wnt signaling during synaptic development and plasticity. Curr Opin Neurobiol 21(1):151–159
207. Inestrosa NC, Arenas E (2010) Emerging roles of Wnts in the adult nervous system. Nat Rev Neurosci 11(2):77–86
208. Pino D, Choe Y, Pleasure SJ (2011) Wnt5a controls neurite development in olfactory bulb interneurons. ASN Neuro 3(3):e00059
209. Tabatadze N, Tomas C, McGonigal R, Lin B, Schook A, Routtenberg A (2012) Wnt transmembrane signaling and long-term spatial memory. Hippocampus 22(6):1228–1241
210. Laffin JJ, Raca G, Jackson CA, Strand EA, Jakielski KJ, Shriberg LD (2012) Novel candidate genes and regions for childhood apraxia of speech identified by array comparative genomic hybridization. Genet Med 14(11):928–936
211. Atcha FA, Syed A, Wu B, Hoverter NP, Yokoyama NN, Ting JH et al (2007) A unique DNA binding domain converts T-cell factors into strong Wnt effectors. Mol Cell Biol 27(23):8352–8363
212. Hoverter NP, Ting JH, Sundaresh S, Baldi P, Waterman ML (2012) A WNT/p21 circuit directed by the C-clamp, a sequence-specific DNA binding domain in TCFs. Mol Cell Biol 32(18):3648–3662
213. Sakakibara S, Nakamura Y, Satoh H, Okano H (2001) Rna-binding protein Musashi2: developmentally regulated expression in neural precursor cells and subpopulations of neurons in mammalian CNS. J Neurosci 21(20):8091–8107
214. Wallmen B, Schrempp M, Hecht A (2012) Intrinsic properties of Tcf1 and Tcf4 splice variants determine cell-type-specific Wnt/beta-catenin target gene expression. Nucleic Acids Res 40(19):9455–9469
215. Raca G, Baas BS, Kirmani S, Laffin JJ, Jackson CA, Strand EA et al (2013) Childhood Apraxia of Speech (CAS) in two patients with 16p11.2 microdeletion syndrome. Eur J Hum Genet 21(4):455–459
216. Kumar RA, Marshall CR, Badner JA, Babatz TD, Mukamel Z, Aldinger KA et al (2009) Association and mutation analyses of 16p11.2 autism candidate genes. PLoS One 4(2):e4582

217. Rosenfeld JA, Coppinger J, Bejjani BA, Girirajan S, Eichler EE, Shaffer LG et al (2010) Speech delays and behavioral problems are the predominant features in individuals with developmental delays and 16p11.2 microdeletions and microduplications. J Neurodev Disord 2(1):26–38
218. Zufferey F, Sherr EH, Beckmann ND, Hanson E, Maillard AM, Hippolyte L et al (2012) A 600 kb deletion syndrome at 16p11.2 leads to energy imbalance and neuropsychiatric disorders. J Med Genet 49(10):660–668
219. Blaker-Lee A, Gupta S, McCammon JM, De Rienzo G, Sive H (2012) Zebrafish homologs of genes within 16p11.2, a genomic region associated with brain disorders, are active during brain development, and include two deletion dosage sensor genes. Dis Model Mech 5(6):834–851
220. Golzio C, Willer J, Talkowski ME, Oh EC, Taniguchi Y, Jacquemont S et al (2012) KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. Nature 485(7398):363–367
221. Schlange T, Matsuda Y, Lienhard S, Huber A, Hynes NE (2007) Autocrine WNT signaling contributes to breast cancer cell proliferation via the canonical WNT pathway and EGFR transactivation. Breast Cancer Res 9(5):R63
222. Wang Z, Chen H (2009) Amino acid limitation induces down-regulation of WNT5a at transcriptional level. Biochem Biophys Res Commun 378(4):789–794
223. Hofmann M, Schuster-Gossler K, Watabe-Rudolph M, Aulehla A, Herrmann BG, Gossler A (2004) WNT signaling, in synergy with T/TBX6, controls Notch signaling by regulating Dll1 expression in the presomitic mesoderm of mouse embryos. Genes Dev 18(22):2712–2717
224. Bauer A, Chauvet S, Huber O, Usseglio F, Rothbacher U, Aragnol D et al (2000) Pontin52 and reptin52 function as antagonistic regulators of beta-catenin signalling activity. EMBO J 19(22):6121–6130
225. Lorain S, Quivy JP, Monier-Gavelle F, Scamps C, Lecluse Y, Almouzni G et al (1998) Core histones and HIRIP3, a novel histone-binding protein, directly interact with WD repeat protein HIRA. Mol Cell Biol 18(9):5546–5556
226. Ye X, Zerlanko B, Kennedy A, Banumathy G, Zhang R, Adams PD (2007) Downregulation of Wnt signaling is a trigger for formation of facultative heterochromatin and onset of cell senescence in primary human cells. Mol Cell 27(2):183–196
227. Kim DS, Hahn Y (2011) Identification of novel phosphorylation modification sites in human proteins that originated after the human-chimpanzee divergence. Bioinformatics 27(18):2494–2501
228. Hilger M, Mann M (2012) Triple SILAC to determine stimulus specific interactions in the Wnt pathway. J Proteome Res 11(2):982–994
229. Lee KH, Johmura Y, Yu LR, Park JE, Gao Y, Bang JK et al (2012) Identification of a novel Wnt5a-CK1varepsilon-Dvl2-Plk1-mediated primary cilia disassembly pathway. EMBO J 31(14):3104–3117
230. Zhang Y, Wang F, Han L, Wu Y, Li S, Yang X et al (2011) GABARAPL1 negatively regulates Wnt/beta-catenin signaling by mediating Dvl2 degradation through the autophagy pathway. Cell Physiol Biochem 27(5):503–512
231. Kumar P, Yadav VK, Baral A, Saha D, Chowdhury S (2011) Zinc-finger transcription factors are associated with guanine quadruplex motifs in human, chimpanzee, mouse and rat promoters genome-wide. Nucleic Acids Res 39(18):8005–8016
232. Li J, Yang Y, Jiang B, Zhang X, Zou Y, Gong Y (2010) Sp1 and KLF15 regulate basal transcription of the human LRP5 gene. BMC Genet 11:12
233. Mikels AJ, Nusse R (2006) Purified Wnt5a protein activates or inhibits beta-catenin-TCF signaling depending on receptor context. PLoS Biol 4(4):e115
234. Balikova A, Jaager K, Viil J, Maimets T, Kadaja-Saarepuu L (2012) Leukocyte marker CD43 promotes cell growth in co-operation with beta-catenin in non-hematopoietic cancer cells. Int J Oncol 41(1):299–309
235. Hsu C, Morohashi Y, Yoshimura S, Manrique-Hoyos N, Jung S, Lauterbach MA et al (2010) Regulation of exosome secretion by Rab35 and its GTPase-activating proteins TBC1D10A-C. J Cell Biol 189(2):223–232

236. Mulholland DJ, Dedhar S, Coetzee GA, Nelson CC (2005) Interaction of nuclear receptors with the Wnt/beta-catenin/Tcf signaling axis: Wnt you like to know? Endocr Rev 26(7):898–915
237. Scott EL, Brann DW (2013) Estrogen regulation of Dkk1 and Wnt/beta-Catenin signaling in neurodegenerative disease. Brain Res 1514:63–74
238. Wandosell F, Varea O, Arevalo MA, Garcia-Segura LM (2012) Oestradiol regulates beta-catenin-mediated transcription in neurones. J Neuroendocrinol 24(1):191–194
239. Schlinger BA, Saldanha CJ (2005) Songbirds: a novel perspective on estrogens and the aging brain. Age (Dordr) 27(4):287–296
240. Ball GF, Balthazart J (2010) Seasonal and hormonal modulation of neurotransmitter systems in the song control circuit. J Chem Neuroanat 39(2):82–95
241. Grisham W, Arnold AP (1995) A direct comparison of the masculinizing effects of testosterone, androstenedione, estrogen, and progesterone on the development of the zebra finch song system. J Neurobiol 26(2):163–170
242. Grisham W, Lee J, Park SH, Mankowski JL, Arnold AP (2008) A dose–response study of estradiol's effects on the developing zebra finch song system. Neurosci Lett 445(2):158–161
243. Remage-Healey L, Saldanha CJ, Schlinger BA (2011) Estradiol synthesis and action at the synapse: evidence for "synaptocrine" signaling. Front Endocrinol (Lausanne) 2:28
244. Kelly MJ, Ronnekleiv OK (2009) Control of CNS neuronal excitability by estrogens via membrane-initiated signaling. Mol Cell Endocrinol 308(1–2):17–25
245. Laughlin GA, Kritz-Silverstein D, Barrett-Connor E (2010) Endogenous oestrogens predict 4-year decline in verbal fluency in postmenopausal women: the Rancho Bernardo Study. Clin Endocrinol (Oxf) 72(1):99–106
246. Rees JM, Regunath G, Whiteside SP, Wadnerkar MB, Cowell PE (2008) Adaptation of wavelet transform analysis to the investigation of biological variations in speech signals. Med Eng Phys 30(7):865–871
247. Anthoni H, Sucheston L, Lewis B, Tapia-Páez I, Fan X, Zucchelli M et al (2012) The Aromatase Gene CYP19A1: several genetic and functional lines of evidence supporting a role in reading, speech and language. Behav Genet 42(4):509–527
248. Schumacher J, Hoffmann P, Schmal C, Schulte-Korne G, Nothen MM (2007) Genetics of dyslexia: the evolving landscape. J Med Genet 44(5):289–297
249. Taipale M, Kaminen N, Nopola-Hemmi J, Haltia T, Myllyluoma B, Lyytinen H et al (2003) A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in brain. Proc Natl Acad Sci 100(20):11553–11558
250. Massinen S, Tammimies K, Tapia-Páez I, Matsson H, Hokkanen M-E, Söderberg O et al (2009) Functional interaction of DYX1C1 with estrogen receptors suggests involvement of hormonal pathways in dyslexia. Hum Mol Genet 18(15):2802–2812
251. Tammimies K, Tapia-Paez I, Ruegg J, Rosin G, Kere J, Gustafsson JA et al (2012) The rs3743205 SNP is important for the regulation of the dyslexia candidate gene DYX1C1 by estrogen receptor beta and DNA methylation. Mol Endocrinol 26(4):619–629
252. Manning JT, Bundred PE (2000) The ratio of 2nd to 4th digit length: a new predictor of disease predisposition? Med Hypotheses 54(5):855–857
253. Stoodley CJ, Stein JF (2013) Cerebellar function in developmental dyslexia. Cerebellum 12(2):267–276
254. Tsutsui K (2008) Neurosteroids in the Purkinje cell: biosynthesis, mode of action and functional significance. Mol Neurobiol 37(2–3):116–125
255. Boets B, De Smedt B, Wouters J, Lemay K, Ghesquiere P (2007) No relation between 2D:4D fetal testosterone marker and dyslexia. Neuroreport 18(14):1487–1491
256. van Gelder M, Tijms J, Hoeks J (2005) Second to fourth digit ratio and dyslexia: no evidence for an association between reading disabilities and the 2D:4D ratio. Dev Med Child Neurol 47(10):718; author reply 9
257. Hanson BJ, Hong W (2003) Evidence for a role of SNX16 in regulating traffic between the early and later endosomal compartments. J Biol Chem 278(36):34617–34630
258. Brankatschk B, Pons V, Parton RG, Gruenberg J (2011) Role of SNX16 in the dynamics of tubulo-cisternal membrane domains of late endosomes. PLoS One 6(7):e21771

259. Rodal AA, Blunk AD, Akbergenova Y, Jorquera RA, Buhl LK, Littleton JT (2011) A presynaptic endosomal trafficking pathway controls synaptic growth signaling. J Cell Biol 193(1): 201–217
260. Seaman MN (2005) Recycle your receptors with retromer. Trends Cell Biol 15(2):68–75
261. Bonifacino JS, Hurley JH (2008) Retromer. Curr Opin Cell Biol 20(4):427–436
262. Cullen PJ, Korswagen HC (2012) Sorting nexins provide diversity for retromer-dependent trafficking events. Nat Cell Biol 14(1):29–37
263. Zhang P, Wu Y, Belenkaya TY, Lin X (2011) SNX3 controls Wingless/Wnt secretion through regulating retromer-dependent recycling of Wntless. Cell Res 21(12):1677–1690
264. Vervoort VS, Viljoen D, Smart R, Suthers G, DuPont BR, Abbott A et al (2002) Sorting nexin 3 (SNX3) is disrupted in a patient with a translocation t(6;13)(q21;q12) and microcephaly, microphthalmia, ectrodactyly, prognathism (MMEP) phenotype. J Med Genet 39(12): 893–899
265. Harbour ME, Breusegem SY, Seaman MN (2012) Recruitment of the endosomal WASH complex is mediated by the extended 'tail' of Fam21 binding to the retromer protein Vps35. Biochem J 442(1):209–220
266. Harterink M, Port F, Lorenowicz MJ, McGough IJ, Silhankova M, Betist MC et al (2011) A SNX3-dependent retromer pathway mediates retrograde transport of the Wnt sorting receptor Wntless and is required for Wnt secretion. Nat Cell Biol 13(8):914–923
267. Zerem A, Vinkler C, Michelson M, Leshinsky-Silver E, Lerman-Sagie T, Lev D (2011) Mosaic marker chromosome 16 resulting in 16q11.2-q12.1 gain in a child with intellectual disability, microcephaly, and cerebellar cortical dysplasia. Am J Med Genet A 155A(12):2991–2996
268. Finelli P, Sirchia SM, Masciadri M, Crippa M, Recalcati MP, Rusconi D et al (2012) Juxtaposition of heterochromatic and euchromatic regions by chromosomal translocation mediates a heterochromatic long-range position effect associated with a severe neurological phenotype. Mol Cytogenet 5:16
269. Payen E, Verkerk T, Michalovich D, Dreyer SD, Winterpacht A, Lee B et al (1998) The centromeric/nucleolar chromatin protein ZFP-37 may function to specify neuronal nuclear domains. J Biol Chem 273(15):9099–9109
270. Tebbenkamp AT, Borchelt DR (2010) Analysis of chaperone mRNA expression in the adult mouse brain by meta analysis of the Allen Brain Atlas. PLoS One 5(10):e13675
271. Nakajima O, Nakamura F, Yamashita N, Tomita Y, Suto F, Okada T et al (2006) FKBP133: a novel mouse FK506-binding protein homolog alters growth cone morphology. Biochem Biophys Res Commun 346(1):140–149
272. Jia D, Gomez TS, Billadeau DD, Rosen MK (2012) Multiple repeat elements within the FAM21 tail link the WASH actin regulatory complex to the retromer. Mol Biol Cell 23(12):2352–2361
273. Voineagu I, Huang L, Winden K, Lazaro M, Haan E, Nelson J et al (2012) CCDC22: a novel candidate gene for syndromic X-linked intellectual disability. Mol Psychiatry 17(1):4–7
274. Hettema EH, Lewis MJ, Black MW, Pelham HR (2003) Retromer and the sorting nexins Snx4/41/42 mediate distinct retrieval pathways from yeast endosomes. EMBO J 22(3): 548–557
275. Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS et al (2003) Global analysis of protein localization in budding yeast. Nature 425(6959):686–691
276. Oliveira DL, Nakayasu ES, Joffe LS, Guimaraes AJ, Sobreira TJ, Nosanchuk JD et al (2010) Characterization of yeast extracellular vesicles: evidence for the participation of different pathways of cellular traffic in vesicle biogenesis. PLoS One 5(6):e11113
277. Caballe A, Martin-Serrano J (2011) ESCRT machinery and cytokinesis: the road to daughter cell separation. Traffic 12(10):1318–1326
278. Harding CV, Heuser JE, Stahl PD (2013) Exosomes: looking back three decades and into the future. J Cell Biol 200(4):367–371
279. Hanson PI, Cashikar A (2012) Multivesicular body morphogenesis. Annu Rev Cell Dev Biol 28:337–362

280. Katoh K, Shibata H, Hatta K, Maki M (2004) CHMP4b is a major binding partner of the ALG-2-interacting protein Alix among the three CHMP4 isoforms. Arch Biochem Biophys 421(1):159–165
281. McCullough J, Fisher RD, Whitby FG, Sundquist WI, Hill CP (2008) ALIX-CHMP4 interactions in the human ESCRT pathway. Proc Natl Acad Sci U S A 105(22):7687–7691
282. Tauro BJ, Greening DW, Mathias RA, Ji H, Mathivanan S, Scott AM et al (2012) Comparison of ultracentrifugation, density gradient separation, and immunoaffinity capture methods for isolating human colon cancer cell line LIM1863-derived exosomes. Methods 56(2): 293–304
283. Gross JC, Chaudhary V, Bartscherer K, Boutros M (2012) Active Wnt proteins are secreted on exosomes. Nat Cell Biol 14(10):1036–1045
284. Koles K, Nunnari J, Korkut C, Barria R, Brewer C, Li Y et al (2012) Mechanism of evenness interrupted (Evi)-exosome release at synaptic boutons. J Biol Chem 287(20):16820–16834
285. Korkut C, Ataman B, Ramachandran P, Ashley J, Barria R, Gherbesi N et al (2009) Trans-synaptic transmission of vesicular Wnt signals through Evi/Wntless. Cell 139(2):393–404
286. Chatellard-Causse C, Blot B, Cristina N, Torch S, Missotten M, Sadoul R (2002) Alix (ALG-2-interacting protein X), a protein involved in apoptosis, binds to endophilins and induces cytoplasmic vacuolization. J Biol Chem 277(32):29108–29115
287. Boucrot E, Pick A, Camdere G, Liska N, Evergren E, McMahon HT et al (2012) Membrane fission is promoted by insertion of amphipathic helices and is restricted by crescent BAR domains. Cell 149(1):124–136
288. Palumbo O, Palumbo P, Palladino T, Stallone R, Miroballo M, Piemontese MR et al (2012) An emerging phenotype of interstitial 15q25.2 microdeletions: clinical report and review. Am J Med Genet A 158A(12):3182–3189
289. Taelman VF, Dobrowolski R, Plouhinec JL, Fuentealba LC, Vorwald PP, Gumper I et al (2010) Wnt signaling requires sequestration of glycogen synthase kinase 3 inside multivesicular endosomes. Cell 143(7):1136–1148
290. Niehrs C, Acebron SP (2010) Wnt signaling: multivesicular bodies hold GSK3 captive. Cell 143(7):1044–1046
291. Chairoungdua A, Smith DL, Pochard P, Hull M, Caplan MJ (2010) Exosome release of beta-catenin: a novel mechanism that antagonizes Wnt signaling. J Cell Biol 190(6):1079–1091
292. Verweij FJ, Middeldorp JM, Pegtel DM (2012) Intracellular signaling controlled by the endosomal-exosomal pathway. Commun Integr Biol 5(1):88–93
293. Fumoto K, Kikuchi K, Gon H, Kikuchi A (2012) Wnt5a signaling controls cytokinesis by correctly positioning ESCRT-III at the midbody. J Cell Sci 125(Pt 20):4822–4832
294. Capalbo L, Montembault E, Takeda T, Bassi ZI, Glover DM, D'Avino PP (2012) The chromosomal passenger complex controls the function of endosomal sorting complex required for transport-III Snf7 proteins during cytokinesis. Open Biol 2(5):120070
295. Carlton JG, Caballe A, Agromayor M, Kloc M, Martin-Serrano J (2012) ESCRT-III governs the Aurora B-mediated abscission checkpoint through CHMP4C. Science 336(6078): 220–225
296. Petronczki M, Uhlmann F (2012) Cell biology. ESCRTing DNA at the cleavage site during cytokinesis. Science 336(6078):166–167
297. Li WM, Webb SE, Lee KW, Miller AL (2006) Recruitment and SNARE-mediated fusion of vesicles in furrow membrane remodeling during cytokinesis in zebrafish embryos. Exp Cell Res 312(17):3260–3275
298. Krepischi-Santos AC, Rajan D, Temple IK, Shrubb V, Crolla JA, Huang S et al (2009) Constitutional haploinsufficiency of tumor suppressor genes in mentally retarded patients with microdeletions in 17p13.1. Cytogenet Genome Res 125(1):1–7
299. Shlien A, Baskin B, Achatz MI, Stavropoulos DJ, Nichols KE, Hudgins L et al (2010) A common molecular mechanism underlies two phenotypically distinct 17p13.1 microdeletion syndromes. Am J Hum Genet 87(5):631–642
300. Yu X, Harris SL, Levine AJ (2006) The regulation of exosome secretion: a novel function of the p53 protein. Cancer Res 66(9):4795–4801

301. Yu X, Riley T, Levine AJ (2009) The regulation of the endosomal compartment by p53 the tumor suppressor gene. FEBS J 276(8):2201–2212
302. Honegger A, Leitz J, Bulkescher J, Hoppe-Seyler K, Hoppe-Seyler F (2013) Silencing of human papillomavirus (HPV) E6/E7 oncogene expression affects both the contents and the amounts of extracellular microvesicles released from HPV-positive cancer cells. Int J Cancer 133:1631–1642
303. Devillard F, Guinchat V, Moreno-De-Luca D, Tabet AC, Gruchy N, Guillem P et al (2010) Paracentric inversion of chromosome 2 associated with cryptic duplication of 2q14 and deletion of 2q37 in a patient with autism. Am J Med Genet A 152A(9):2346–2354
304. Loomis CA, Kimmel RA, Tong CX, Michaud J, Joyner AL (1998) Analysis of the genetic pathway leading to formation of ectopic apical ectodermal ridges in mouse Engrailed-1 mutant limbs. Development 125(6):1137–1148
305. Bachar-Dahan L, Goltzmann J, Yaniv A, Gazit A (2006) Engrailed-1 negatively regulates beta-catenin transcriptional activity by destabilizing beta-catenin via a glycogen synthase kinase-3beta-independent pathway. Mol Biol Cell 17(6):2572–2580
306. Sillitoe RV, Vogel MW, Joyner AL (2010) Engrailed homeobox genes regulate establishment of the cerebellar afferent circuit map. J Neurosci 30(30):10015–10024
307. Kuemerle B, Gulden F, Cherosky N, Williams E, Herrup K (2007) The mouse Engrailed genes: a window into autism. Behav Brain Res 176(1):121–132
308. Inuzuka T, Inokawa A, Chen C, Kizu K, Narita H, Shibata H et al (2013) ALG-2-interacting Tubby-like protein superfamily member PLSCR3 is secreted by exosomal pathway and taken up by recipient cultured cells. Biosci Rep 33(2):art:e00026.doi:10.1042/BSR20120123
309. Mathivanan S, Lim JW, Tauro BJ, Ji H, Moritz RL, Simpson RJ (2010) Proteomics analysis of A33 immunoaffinity-purified exosomes released from the human colon tumor cell line LIM1215 reveals a tissue-specific protein signature. Mol Cell Proteomics 9(2):197–208
310. Shibata H, Suzuki H, Kakiuchi T, Inuzuka T, Yoshida H, Mizuno T et al (2008) Identification of Alix-type and Non-Alix-type ALG-2-binding sites in human phospholipid scramblase 3: differential binding to an alternatively spliced isoform and amino acid-substituted mutants. J Biol Chem 283(15):9623–9632
311. Gonzalez LJ, Gibbons E, Bailey RW, Fairbourn J, Nguyen T, Smith SK et al (2009) The influence of membrane physical properties on microvesicle release in human erythrocytes. PMC Biophys 2(1):7
312. Schneider A, Simons M (2013) Exosomes: vesicular carriers for intercellular communication in neurodegenerative disorders. Cell Tissue Res 352(1):33–47
313. Close P, Gillard M, Ladang A, Jiang Z, Papuga J, Hawkes N et al (2012) DERP6 (ELP5) and C3ORF75 (ELP6) regulate tumorigenicity and migration of melanoma cells as subunits of Elongator. J Biol Chem 287(39):32535–32545
314. Creppe C, Buschbeck M (2011) Elongator: an ancestral complex driving transcription and migration through protein acetylation. J Biomed Biotechnol 2011:924898
315. Rahl PB, Chen CZ, Collins RN (2005) Elp1p, the yeast homolog of the FD disease syndrome protein, negatively regulates exocytosis independently of transcriptional elongation. Mol Cell 17(6):841–853
316. Yuan J, Tang W, Luo K, Chen X, Gu X, Wan B et al (2006) Cloning and characterization of the human gene DERP6, which activates transcriptional activities of p53. Mol Biol Rep 33(3):151–158
317. Tanaka SS, Nakane A, Yamaguchi YL, Terabayashi T, Abe T, Nakao K et al (2013) Dullard/Ctdnep1 modulates WNT signalling activity for the formation of primordial germ cells in the mouse embryo. PLoS One 8(3):e57428
318. Han S, Bahmanyar S, Zhang P, Grishin N, Oegema K, Crooke R et al (2012) Nuclear envelope phosphatase 1-regulatory subunit 1 (formerly TMEM188) is the metazoan Spo7p ortholog and functions in the lipin activation pathway. J Biol Chem 287(5):3123–3137
319. Kim Y, Gentry MS, Harris TE, Wiley SE, Lawrence JC Jr, Dixon JE (2007) A conserved phosphatase cascade that regulates nuclear membrane biogenesis. Proc Natl Acad Sci U S A 104(16):6596–6601

320. Karanasios E, Han GS, Xu Z, Carman GM, Siniossoglou S (2010) A phosphorylation-regulated amphipathic helix controls the membrane translocation and function of the yeast phosphatidate phosphatase. Proc Natl Acad Sci U S A 107(41):17539–17544
321. Michaillat L, Mayer A (2013) Identification of genes affecting vacuole membrane fragmentation in Saccharomyces cerevisiae. PLoS One 8(2):e54160
322. Sasser T, Qiu QS, Karunakaran S, Padolina M, Reyes A, Flood B et al (2012) Yeast lipin 1 orthologue pah1p regulates vacuole homeostasis and membrane fusion. J Biol Chem 287(3):2221–2236
323. Cha YH, Kim NH, Park C, Lee I, Kim HS, Yook JI (2012) MiRNA-34 intrinsically links p53 tumor suppressor and Wnt signaling. Cell Cycle 11(7):1273–1281
324. Kim NH, Kim HS, Kim NG, Lee I, Choi HS, Li XY et al (2011) p53 and microRNA-34 are suppressors of canonical Wnt signaling. Sci Signal 4(197):ra71
325. Williams CA, Driscoll DJ, Dagli AI (2010) Clinical and genetic aspects of Angelman syndrome. Genet Med 12(7):385–395
326. Beaudenon S, Huibregtse JM (2008) HPV E6, E6AP and cervical cancer. BMC Biochem 9(suppl 1):S4
327. Levav-Cohen Y, Wolyniec K, Alsheich-Bartok O, Chan AL, Woods SJ, Jiang YH et al (2012) E6AP is required for replicative and oncogene-induced senescence in mouse embryo fibroblasts. Oncogene 31(17):2199–2209
328. Gunther EJ, Moody SE, Belka GK, Hahn KT, Innocent N, Dugan KD et al (2003) Impact of p53 loss on reversal and recurrence of conditional Wnt-induced tumorigenesis. Genes Dev 17(4):488–501
329. Kim NH, Kim HS, Li XY, Lee I, Choi HS, Kang SE et al (2011) A p53/miRNA-34 axis regulates Snail1-dependent cancer cell epithelial-mesenchymal transition. J Cell Biol 195(3):417–433
330. Lichtig H, Gilboa DA, Jackman A, Gonen P, Levav-Cohen Y, Haupt Y et al (2010) HPV16 E6 augments Wnt signaling in an E6AP-dependent manner. Virology 396(1):47–58
331. Perez-Plasencia C, Duenas-Gonzalez A, Alatorre-Tavera B (2008) Second hit in cervical carcinogenesis process: involvement of wnt/beta catenin pathway. Int Arch Med 1(1):10
332. Sapp M, Bienkowska-Haba M (2009) Viral entry mechanisms: human papillomavirus and a long journey from extracellular matrix to the nucleus. FEBS J 276(24):7206–7216
333. Straight SW, Herman B, McCance DJ (1995) The E5 oncoprotein of human papillomavirus type 16 inhibits the acidification of endosomes in human keratinocytes. J Virol 69(5):3185–3192
334. Suprynowicz FA, Krawczyk E, Hebert JD, Sudarshan SR, Simic V, Kamonjoh CM et al (2010) The human papillomavirus type 16 E5 oncoprotein inhibits epidermal growth factor trafficking independently of endosome acidification. J Virol 84(20):10619–10629
335. Scheffer KD, Gawlitza A, Spoden GA, Zhang XA, Lambert C, Berditchevski F et al (2013) Tetraspanin CD151 mediates papillomavirus type 16 endocytosis. J Virol 87(6):3435–3446
336. Spoden G, Freitag K, Husmann M, Boller K, Sapp M, Lambert C et al (2008) Clathrin- and caveolin-independent entry of human papillomavirus type 16—involvement of tetraspanin-enriched microdomains (TEMs). PLoS One 3(10):e3313
337. Bergant M, Banks L (2013) SNX17 facilitates infection with diverse papillomavirus types. J Virol 87(2):1270–1273
338. Bergant Marusic M, Ozbun MA, Campos SK, Myers MP, Banks L (2012) Human papillomavirus L2 facilitates viral escape from late endosomes via sorting nexin 17. Traffic 13(3):455–467
339. Hasegawa M, Furuya M, Kasuya Y, Nishiyama M, Sugiura T, Nikaido T et al (2007) CD151 dynamics in carcinoma-stroma interaction: integrin expression, adhesion strength and proteolytic activity. Lab Invest 87(9):882–892
340. Roberts JN, Buck CB, Thompson CD, Kines R, Bernardo M, Choyke PL, Lowy DR, Schiller JT (2007) Genital transmission of HPV in a mouse model is potentiated by nonoxynol-9 and inhibited by carrageenan. Nat Med 13(7):857–861
341. Sterk LM, Geuijen CA, Oomen LC, Calafat J, Janssen H, Sonnenberg A (2000) The tetraspan molecule CD151, a novel constituent of hemidesmosomes, associates with the integrin

alpha6beta4 and may regulate the spatial organization of hemidesmosomes. J Cell Biol 149(4):969–982
342. Donoso M, Cancino J, Lee J, van Kerkhof P, Retamal C, Bu G et al (2009) Polarized traffic of LRP1 involves AP1B and SNX17 operating on Y-dependent sorting motifs in different pathways. Mol Biol Cell 20(1):481–497
343. Farfan P, Lee J, Larios J, Sotelo P, Bu G, Marzolo MP (2013) A sorting nexin 17-binding domain within the LRP1 cytoplasmic tail mediates receptor recycling through the basolateral sorting endosome. Traffic 14(7):823–838
344. McCabe KL, Bronner M (2011) Tetraspanin, CD151, is required for maintenance of trigeminal placode identity. J Neurochem 117(2):221–230
345. Horton AC, Ehlers MD (2003) Neuronal polarity and trafficking. Neuron 40(2):277–295
346. Lasiecka ZM, Winckler B (2011) Mechanisms of polarized membrane trafficking in neurons—focusing in on endosomes. Mol Cell Neurosci 48(4):278–287
347. Terrand J, Bruban V, Zhou L, Gong W, El Asmar Z, May P et al (2009) LRP1 controls intracellular cholesterol storage and fatty acid synthesis through modulation of Wnt signaling. J Biol Chem 284(1):381–388
348. Woldt E, Terrand J, Mlih M, Matz RL, Bruban V, Coudane F et al (2012) The nuclear hormone receptor PPARgamma counteracts vascular calcification by inhibiting Wnt5a signalling in vascular smooth muscle cells. Nat Commun 3:1077
349. Krawczyk E, Hanover JA, Schlegel R, Suprynowicz FA (2008) Karyopherin beta3: a new cellular target for the HPV-16 E5 oncoprotein. Biochem Biophys Res Commun 371(4): 684–688
350. Hansen CN, Ketabi Z, Rosenstierne MW, Palle C, Boesen HC, Norrild B (2009) Expression of CPEB, GAPDH and U6snRNA in cervical and ovarian tissue during cancer development. APMIS 117(1):53–59
351. Chao HW, Lai YT, Lu YL, Lin CL, Mai W, Huang YS (2012) NMDAR signaling facilitates the IPO5-mediated nuclear import of CPEB3. Nucleic Acids Res 40(17):8484–8498
352. Peng SC, Lai YT, Huang HY, Huang HD, Huang YS (2010) A novel role of CPEB3 in regulating EGFR gene transcription via association with Stat5b in neurons. Nucleic Acids Res 38(21):7446–7457
353. Spangle JM, Munger K (2013) The HPV16 E6 oncoprotein causes prolonged receptor protein tyrosine kinase signaling and enhances internalization of phosphorylated receptor species. PLoS Pathog 9(3):e1003237
354. Pavlopoulos E, Trifilieff P, Chevaleyre V, Fioriti L, Zairis S, Pagano A et al (2011) Neuralized1 activates CPEB3: a function for nonproteolytic ubiquitin in synaptic plasticity and memory storage. Cell 147(6):1369–1383
355. Vogler C, Spalek K, Aerni A, Demougin P, Muller A, Huynh KD et al (2009) CPEB3 is associated with human episodic memory. Front Behav Neurosci 3:4
356. Condon KH, Ho J, Robinson CG, Hanus C, Ehlers MD (2013) The Angelman syndrome protein Ube3a/E6AP is required for Golgi acidification and surface protein sialylation. J Neurosci 33(9):3799–3814
357. Ohgaki R, van IJzendoorn SC, Matsushita M, Hoekstra D, Kanazawa H (2011) Organellar Na+/H+ exchangers: novel players in organelle pH regulation and their emerging functions. Biochemistry 50(4):443–450
358. Schwede M, Garbett K, Mirnics K, Geschwind DH, Morrow EM (2013) Genes for endosomal NHE6 and NHE9 are misregulated in autism brains. Mol Psychiatry advance online publication 19 March 2013; doi:10.1038/mp.2013.28
359. Jeong KW, Kim HZ, Kim S, Kim YS, Choe J (2007) Human papillomavirus type 16 E6 protein interacts with cystic fibrosis transmembrane regulator-associated ligand and promotes E6-associated protein-mediated ubiquitination and proteasomal degradation. Oncogene 26(4):487–499
360. Cheng J, Cebotaru V, Cebotaru L, Guggino WB (2010) Syntaxin 6 and CAL mediate the degradation of the cystic fibrosis transmembrane conductance regulator. Mol Biol Cell 21(7):1178–1187

361. Cheng J, Moyer BD, Milewski M, Loffing J, Ikeda M, Mickle JE et al (2002) A Golgi-associated PDZ domain protein modulates cystic fibrosis transmembrane regulator plasma membrane expression. J Biol Chem 277(5):3520–3529
362. Haggie PM, Verkman AS (2009) Defective organellar acidification as a cause of cystic fibrosis lung disease: reexamination of a recurring hypothesis. Am J Physiol Lung Cell Mol Physiol 296(6):L859–L867
363. Barasch J, Kiss B, Prince A, Saiman L, Gruenert D, al-Awqati Q (1991) Defective acidification of intracellular organelles in cystic fibrosis. Nature 352(6330):70–73
364. Dosanjh A, Muchmore EA (2009) Expression of DeltaF508 cystic fibrosis transmembrane regulator (CFTR) decreases membrane sialylation. Open Respir Med J 3:79–84
365. Rhim AD, Stoykova L, Glick MC, Scanlin TF (2001) Terminal glycosylation in cystic fibrosis (CF): a review emphasizing the airway epithelial cell. Glycoconj J 18(9):649–659
366. Gentzsch M, Cui L, Mengos A, Chang XB, Chen JH, Riordan JR (2003) The PDZ-binding chloride channel ClC-3B localizes to the Golgi and associates with cystic fibrosis transmembrane conductance regulator-interacting PDZ proteins. J Biol Chem 278(8):6440–6449
367. LaLonde DP, Bretscher A (2009) The scaffold protein PDZK1 undergoes a head-to-tail intramolecular association that negatively regulates its interaction with EBP50. Biochemistry 48(10):2261–2271
368. Wheeler DS, Barrick SR, Grubisha MJ, Brufsky AM, Friedman PA, Romero G (2011) Direct interaction between NHERF1 and Frizzled regulates beta-catenin signaling. Oncogene 30(1):32–42
369. Yao R, Maeda T, Takada S, Noda T (2001) Identification of a PDZ domain containing Golgi protein, GOPC, as an interaction partner of frizzled. Biochem Biophys Res Commun 286(4):771–778
370. Maeda Y, Ide T, Koike M, Uchiyama Y, Kinoshita T (2008) GPHR is a novel anion channel critical for acidification and functions of the Golgi apparatus. Nat Cell Biol 10(10):1135–1145
371. Brunetti-Pierri N, Berg JS, Scaglia F, Belmont J, Bacino CA, Sahoo T et al (2008) Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. Nat Genet 40(12):1466–1471
372. Mefford HC, Sharp AJ, Baker C, Itsara A, Jiang Z, Buysse K et al (2008) Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. N Engl J Med 359(16):1685–1699
373. Rosenfeld JA, Traylor RN, Schaefer GB, McPherson EW, Ballif BC, Klopocki E et al (2012) Proximal microdeletions and microduplications of 1q21.1 contribute to variable abnormal phenotypes. Eur J Hum Genet 20(7):754–761
374. Gilfillan GD, Selmer KK, Roxrud I, Smith R, Kyllerman M, Eiklid K et al (2008) SLC9A6 mutations cause X-linked mental retardation, microcephaly, epilepsy, and ataxia, a phenotype mimicking Angelman syndrome. Am J Hum Genet 82(4):1003–1010
375. Schroer RJ, Holden KR, Tarpey PS, Matheus MG, Griesemer DA, Friez MJ et al (2010) Natural history of Christianson syndrome. Am J Med Genet A 152A(11):2775–2783
376. Miao S, Chen R, Ye J, Tan GH, Li S, Zhang J et al (2013) The Angelman syndrome protein Ube3a is required for polarized dendrite morphogenesis in pyramidal neurons. J Neurosci 33(1):327–333
377. Gulesserian T, Engidawork E, Fountoulakis M, Lubec G (2007) Manifold decrease of sialic acid synthase in fetal Down syndrome brain. Amino Acids 32(1):141–144
378. Coombs GS, Yu J, Canning CA, Veltri CA, Covey TM, Cheong JK et al (2010) WLS-dependent secretion of WNT3A requires Ser209 acylation and vacuolar acidification. J Cell Sci 123(Pt 19):3357–3367
379. Cruciat CM, Ohkawara B, Acebron SP, Karaulanov E, Reinhard C, Ingelfinger D et al (2010) Requirement of prorenin receptor and vacuolar H+−ATPase-mediated acidification for Wnt signaling. Science 327(5964):459–463
380. Herr P, Basler K (2012) Porcupine-mediated lipidation is required for Wnt recognition by Wls. Dev Biol 361(2):392–402

381. Proffitt KD, Virshup DM (2012) Precise regulation of porcupine activity is required for physiological Wnt signaling. J Biol Chem 287(41):34167–34178
382. Giorda R, Bonaglia MC, Beri S, Fichera M, Novara F, Magini P et al (2009) Complex segmental duplications mediate a recurrent dup(X)(p11.22-p11.23) associated with mental retardation, speech delay, and EEG anomalies in males and females. Am J Hum Genet 85(3):394–400
383. Jung H, Lee SK, Jho EH (2011) Mest/Peg1 inhibits Wnt signalling through regulation of LRP6 glycosylation. Biochem J 436(2):263–269
384. Schneider E, Mayer S, El Hajj N, Jensen LR, Kuss AW, Zischler H et al (2012) Methylation and expression analyses of the 7q autism susceptibility locus genes MEST, COPG2, and TSGA14 in human and anthropoid primate cortices. Cytogenet Genome Res 136(4):278–287
385. Korvatska O, Estes A, Munson J, Dawson G, Bekris LM, Kohen R et al (2011) Mutations in the TSGA14 gene in families with autism spectrum disorders. Am J Med Genet B Neuropsychiatr Genet 156B(3):303–311
386. Yamada T, Mitsuya K, Kayashima T, Yamasaki K, Ohta T, Yoshiura K et al (2004) Imprinting analysis of 10 genes and/or transcripts in a 1.5-Mb MEST-flanking region at human chromosome 7q32. Genomics 83(3):402–412
387. MacIsaac JL, Bogutz AB, Morrissy AS, Lefebvre L (2012) Tissue-specific alternative polyadenylation at the imprinted gene Mest regulates allelic usage at Copg2. Nucleic Acids Res 40(4):1523–1535
388. Krey JF, Pasca SP, Shcheglovitov A, Yazawa M, Schwemberger R, Rasmusson R et al (2013) Timothy syndrome is associated with activity-dependent dendritic retraction in rodent and human neurons. Nat Neurosci 16(2):201–209
389. Pasca SP, Portmann T, Voineagu I, Yazawa M, Shcheglovitov A, Pasca AM et al (2011) Using iPSC-derived neurons to uncover cellular phenotypes associated with Timothy syndrome. Nat Med 17(12):1657–1662
390. Fischer B, Dimopoulou A, Egerer J, Gardeitchik T, Kidd A, Jost D et al (2012) Further characterization of ATP6V0A2-related autosomal recessive cutis laxa. Hum Genet 131(11):1761–1773
391. Hermle T, Guida MC, Beck S, Helmstadter S, Simons M (2013) Drosophila ATP6AP2/VhaPRR functions both as a novel planar cell polarity core protein and a regulator of endosomal trafficking. EMBO J 32(2):245–259
392. Ramser J, Abidi FE, Burckle CA, Lenski C, Toriello H, Wen G et al (2005) A unique exonic splice enhancer mutation in a family with X-linked mental retardation and epilepsy points to a novel role of the renin receptor. Hum Mol Genet 14(8):1019–1027
393. Reiter LT, Seagroves TN, Bowers M, Bier E (2006) Expression of the Rho-GEF Pbl/ECT2 is regulated by the UBE3A E3 ubiquitin ligase. Hum Mol Genet 15(18):2825–2835
394. Morita E, Sandrin V, Chung HY, Morham SG, Gygi SP, Rodesch CK et al (2007) Human ESCRT and ALIX proteins interact with proteins of the midbody and function in cytokinesis. EMBO J 26(19):4215–4227
395. Simon GC, Schonteich E, Wu CC, Piekny A, Ekiert D, Yu X et al (2008) Sequential Cyk-4 binding to ECT2 and FIP3 regulates cleavage furrow ingression and abscission during cytokinesis. EMBO J 27(13):1791–1803
396. Jensen L, Farook MF, Reiter LT (2013) Proteomic profiling in Drosophila reveals potential dube3a regulation of the actin cytoskeleton and neuronal homeostasis. PLoS One 8(4):e61952
397. Smith CM, Chircop M (2012) Clathrin-mediated endocytic proteins are involved in regulating mitotic progression and completion. Traffic 13(12):1628–1641
398. Tjota M, Lee SK, Wu J, Williams JA, Khanna MR, Thomas GH (2011) Annexin B9 binds to beta(H)-spectrin and is required for multivesicular body function in Drosophila. J Cell Sci 124(Pt 17):2914–2926
399. Keil R, Kiessling C, Hatzfeld M (2009) Targeting of p0071 to the midbody depends on KIF3. J Cell Sci 122(Pt 8):1174–1183
400. Wolf A, Keil R, Gotzl O, Mun A, Schwarze K, Lederer M et al (2006) The armadillo protein p0071 regulates Rho signalling during cytokinesis. Nat Cell Biol 8(12):1432–1440

401. Liu P, Carvalho CM, Hastings PJ, Lupski JR (2012) Mechanisms for recurrent and complex human genomic rearrangements. Curr Opin Genet Dev 22(3):211–220
402. Liu P, Lacaria M, Zhang F, Withers M, Hastings PJ, Lupski JR (2011) Frequency of nonallelic homologous recombination is correlated with length of homology: evidence that ectopic synapsis precedes ectopic crossing-over. Am J Hum Genet 89(4):580–588
403. Carmona-Mora P, Walz K (2010) Retinoic acid induced 1, RAI1: a dosage sensitive gene related to neurobehavioral alterations including autistic behavior. Curr Genomics 11(8):607–617
404. Gulhan Ercan-Sencicek A, Davis Wright NR, Frost SJ, Fulbright RK, Felsenfeld S, Hart L et al (2012) Searching for Potocki-Lupski syndrome phenotype: a patient with language impairment and no autism. Brain Dev 34(8):700–703
405. Zhang F, Potocki L, Sampson JB, Liu P, Sanchez-Valle A, Robbins-Furman P et al (2010) Identification of uncommon recurrent Potocki-Lupski syndrome-associated duplications and the distribution of rearrangement types and mechanisms in PTLS. Am J Hum Genet 86(3):462–470
406. Lovell PV, Clayton DF, Replogle KL, Mello CV (2008) Birdsong "transcriptomics": neurochemical specializations of the oscine song system. PLoS One 3(10):e3440
407. Nam K, Mugal C, Nabholz B, Schielzeth H, Wolf JB, Backstrom N et al (2010) Molecular evolution of genes in avian genomes. Genome Biol 11(6):R68
408. Denisenko-Nehrbass NI, Jarvis E, Scharff C, Nottebohm F, Mello CV (2000) Site-specific retinoic acid production in the brain of adult songbirds. Neuron 27(2):359–370
409. Denisenko-Nehrbass NI, Mello CV (2001) Molecular targets of disulfiram action on song maturation in zebra finches. Brain Res Mol Brain Res 87(2):246–250
410. Olson CR, Rodrigues PV, Jeong JK, Prahl DJ, Mello CV (2011) Organization and development of zebra finch HVC and paraHVC based on expression of zRalDH, an enzyme associated with retinoic acid production. J Comp Neurol 519(1):148–161
411. Herman EK, Walker G, van der Giezen M, Dacks JB (2011) Multivesicular bodies in the enigmatic amoeboflagellate Breviata anathema and the evolution of ESCRT 0. J Cell Sci 124(Pt 4):613–621
412. Blanc C, Charette SJ, Mattei S, Aubry L, Smith EW, Cosson P et al (2009) Dictyostelium Tom1 participates to an ancestral ESCRT-0 complex. Traffic 10(2):161–171
413. Puertollano R (2005) Interactions of TOM1L1 with the multivesicular body sorting machinery. J Biol Chem 280(10):9258–9264
414. Yanagida-Ishizaki Y, Takei T, Ishizaki R, Imakagura H, Takahashi S, Shin HW et al (2008) Recruitment of Tom1L1/Srcasm to endosomes and the midbody by Tsg101. Cell Struct Funct 33(1):91–100
415. Wang T, Liu NS, Seet LF, Hong W (2010) The emerging role of VHS domain-containing Tom1, Tom1L1 and Tom1L2 in membrane trafficking. Traffic 11(9):1119–1128
416. Katoh Y, Imakagura H, Futatsumori M, Nakayama K (2006) Recruitment of clathrin onto endosomes by the Tom1-Tollip complex. Biochem Biophys Res Commun 341(1): 143–149
417. Bayes M, Magano LF, Rivera N, Flores R, Perez Jurado LA (2003) Mutational mechanisms of Williams-Beuren syndrome deletions. Am J Hum Genet 73(1):131–151
418. Pober BR (2010) Williams-Beuren syndrome. N Engl J Med 362(3):239–252
419. Haas BW, Reiss AL (2012) Social brain development in williams syndrome: the current status and directions for future research. Front Psychol 3:186
420. Porter MA, Shaw TA, Marsh PJ (2010) An unusual attraction to the eyes in Williams-Beuren syndrome: a manipulation of facial affect while measuring face scanpaths. Cogn Neuropsychiatry 15(6):505–530
421. Riby DM, Hancock PJ (2008) Viewing it differently: social scene perception in Williams syndrome and autism. Neuropsychologia 46(11):2855–2860
422. Jones W, Carr K, Klin A (2008) Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. Arch Gen Psychiatry 65(8):946–954

423. Rice K, Moriuchi JM, Jones W, Klin A (2012) Parsing heterogeneity in autism spectrum disorders: visual scanning of dynamic social scenes in school-aged children. J Am Acad Child Adolesc Psychiatry 51(3):238–248
424. Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D et al (2011) Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. Neuron 70(5):863–885
425. Velleman SL, Mervis CB (2011) Children with 7q11.23 duplication syndrome: speech, language, cognitive, and behavioral characteristics and their implications for intervention. Perspect Lang Learn Educ 18(3):108–116
426. Masataka N (2001) Why early linguistic milestones are delayed in children with Williams syndrome: late onset of hand banging as a possible rate–limiting constraint on the emergence of canonical babbling. Dev Sci 4(2):158–164
427. Osborne LR, Mervis CB (2007) Rearrangements of the Williams-Beuren syndrome locus: molecular basis and implications for speech and language development. Expert Rev Mol Med 9(15):1–16
428. Berg JS, Brunetti-Pierri N, Peters SU, Kang SH, Fong CT, Salamone J et al (2007) Speech delay and autism spectrum behaviors are frequently associated with duplication of the 7q11.23 Williams-Beuren syndrome region. Genet Med 9(7):427–441
429. Somerville MJ, Mervis CB, Young EJ, Seo EJ, del Campo M, Bamforth S et al (2005) Severe expressive-language delay related to duplication of the Williams-Beuren locus. N Engl J Med 353(16):1694–1701
430. Frangiskakis JM, Ewart AK, Morris CA, Mervis CB, Bertrand J, Robinson BF et al (1996) LIM-kinase1 hemizygosity implicated in impaired visuospatial constructive cognition. Cell 86(1):59–69
431. Gray V, Karmiloff-Smith A, Funnell E, Tassabehji M (2006) In-depth analysis of spatial cognition in Williams syndrome: a critical assessment of the role of the LIMK1 gene. Neuropsychologia 44(5):679–685
432. Matsumoto N, Kitani R, Kalinec F (2011) Linking LIMK1 deficiency to hyperacusis and progressive hearing loss in individuals with Williams syndrome. Commun Integr Biol 4(2):208–210
433. Zhao C, Aviles C, Abel RA, Almli CR, McQuillen P, Pleasure SJ (2005) Hippocampal and visuospatial learning defects in mice with a deletion of frizzled 9, a gene in the Williams syndrome deletion interval. Development 132(12):2917–2927
434. Walaas SI, Hemmings HC Jr, Greengard P, Nairn AC (2011) Beyond the dopamine receptor: regulation and roles of serine/threonine protein phosphatases. Front Neuroanat 5:50
435. Yger M, Girault JA (2011) DARPP-32, Jack of All Trades… Master of Which? Front Behav Neurosci 5:56
436. Davis RL (1996) Physiology and biochemistry of Drosophila learning mutants. Physiol Rev 76(2):299–317
437. Kandel ER (2012) The molecular biology of memory: cAMP, PKA, CRE, CREB-1, CREB-2, and CPEB. Mol Brain 5:14
438. Hansen C, Howlin J, Tengholm A, Dyachok O, Vogel WF, Nairn AC et al (2009) Wnt-5a-induced phosphorylation of DARPP-32 inhibits breast cancer cell migration in a CREB-dependent manner. J Biol Chem 284(40):27533–27543
439. Mayr B, Montminy M (2001) Transcriptional regulation by the phosphorylation-dependent factor CREB. Nat Rev Mol Cell Biol 2(8):599–609
440. Hotte M, Thuault S, Lachaise F, Dineley KT, Hemmings HC, Nairn AC et al (2006) D1 receptor modulation of memory retrieval performance is associated with changes in pCREB and pDARPP-32 in rat prefrontal cortex. Behav Brain Res 171(1):127–133
441. Viola H, Furman M, Izquierdo LA, Alonso M, Barros DM, de Souza MM et al (2000) Phosphorylated cAMP response element-binding protein as a molecular marker of memory processing in rat hippocampus: effect of novelty. J Neurosci 20(23):RC112
442. Cerpa W, Gambrill A, Inestrosa NC, Barria A (2011) Regulation of NMDA-receptor synaptic transmission by Wnt signaling. J Neurosci 31(26):9466–9471

443. Maguschak KA, Ressler KJ (2011) Wnt signaling in amygdala-dependent learning and memory. J Neurosci 31(37):13057–13067
444. Riters LV (2012) The role of motivation and reward neural systems in vocal communication in songbirds. Front Neuroendocrinol 33(2):194–209
445. Jarjour S, Bai L, Gianoulakis C (2009) Effect of acute ethanol administration on the release of opioid peptides from the midbrain including the ventral tegmental area. Alcohol Clin Exp Res 33(6):1033–1043
446. Leriche M, Cote-Velez A, Mendez M (2007) Presence of pro-opiomelanocortin mRNA in the rat medial prefrontal cortex, nucleus accumbens and ventral tegmental area: studies by RT-PCR and in situ hybridization techniques. Neuropeptides 41(6):421–431
447. Rodriguez-Munoz M, de la Torre-Madrid E, Sanchez-Blazquez P, Wang JB, Garzon J (2008) NMDAR-nNOS generated zinc recruits PKCgamma to the HINT1-RGS17 complex bound to the C terminus of Mu-opioid receptors. Cell Signal 20(10):1855–1864
448. Sanchez-Blazquez P, Rodriguez-Munoz M, Bailon C, Garzon J (2012) GPCRs promote the release of zinc ions mediated by nNOS/NO and the redox transducer RGSZ2 protein. Antioxid Redox Signal 17(9):1163–1177
449. Huber O, Weiske J (2008) Beta-catenin takes a HIT. Cell Cycle 7(10):1326–1331
450. Rodriguez-Munoz M, Garzon J (2013) Nitric oxide and zinc-mediated protein assemblies involved in Mu opioid receptor signaling. Mol Neurobiol, doi:10.1007/s12035-013-8465-z
451. James MA, Lu Y, Liu Y, Vikis HG, You M (2009) RGS17, an overexpressed gene in human lung and prostate cancer, induces tumor cell proliferation through the cyclic AMP-PKA-CREB pathway. Cancer Res 69(5):2108–2116
452. Chen X, Geller EB, Rogers TJ, Adler MW (2007) The chemokine CX3CL1/fractalkine interferes with the antinociceptive effect induced by opioid agonists in the periaqueductal grey of rats. Brain Res 1153:52–57
453. Heinisch S, Palma J, Kirby LG (2011) Interactions between chemokine and mu-opioid receptors: anatomical findings and electrophysiological studies in the rat periaqueductal grey. Brain Behav Immun 25(2):360–372
454. Suzuki M, El-Hage N, Zou S, Hahn YK, Sorrell ME, Sturgill JL et al (2011) Fractalkine/CX3CL1 protects striatal neurons from synergistic morphine and HIV-1 Tat-induced dendritic losses and death. Mol Neurodegener 6:78
455. Khurshid N, Agarwal V, Iyengar S (2009) Expression of mu- and delta-opioid receptors in song control regions of adult male zebra finches (Taenopygia guttata). J Chem Neuroanat 37(3):158–169
456. Wada K, Howard JT, McConnell P, Whitney O, Lints T, Rivas MV et al (2006) A molecular neuroethological approach for identifying and characterizing a cascade of behaviorally regulated genes. Proc Natl Acad Sci U S A 103(41):15212–15217
457. Jin J, Kittanakom S, Wong V, Reyes BA, Van Bockstaele EJ, Stagljar I et al (2010) Interaction of the mu-opioid receptor with GPR177 (Wntless) inhibits Wnt secretion: potential implications for opioid dependence. BMC Neurosci 11:33
458. Reyes AR, Levenson R, Berrettini W, Van Bockstaele EJ (2010) Ultrastructural relationship between the mu opioid receptor and its interacting protein, GPR177, in striatal neurons. Brain Res 1358:71–80
459. Reyes BA, Vakharia K, Ferraro TN, Levenson R, Berrettini WH, Van Bockstaele EJ (2012) Opiate agonist-induced re-distribution of Wntless, a mu-opioid receptor interacting protein, in rat striatal neurons. Exp Neurol 233(1):205–213
460. Strug LJ, Addis L, Chiang T, Baskurt Z, Li W, Clarke T et al (2012) The genetics of reading disability in an often excluded sample: novel loci suggested for reading disability in Rolandic epilepsy. PLoS One 7(7):e40696
461. Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S et al (2012) A high-coverage genome sequence from an archaic Denisovan individual. Science 338(6104):222–226
462. Kang C, Drayna D (2012) A role for inherited metabolic deficits in persistent developmental stuttering. Mol Genet Metab 107(3):276–280

463. Clarke RA, Lee S, Eapen V (2012) Pathogenetic model for Tourette syndrome delineates overlap with related neurodevelopmental disorders including Autism. Transl Psychiatry 2:e158
464. Petrin AL, Giacheti CM, Maximino LP, Abramides DV, Zanchetta S, Rossi NF et al (2010) Identification of a microdeletion at the 7q33-q35 disrupting the CNTNAP2 gene in a Brazilian stuttering case. Am J Med Genet A 152A(12):3164–3172
465. Verkerk AJ, Mathews CA, Joosse M, Eussen BH, Heutink P, Oostra BA (2003) CNTNAP2 is disrupted in a family with Gilles de la Tourette syndrome and obsessive compulsive disorder. Genomics 82(1):1–9
466. Tzschach A, Bisgaard AM, Kirchhoff M, Graul-Neumann LM, Neitzel H, Page S et al (2010) Chromosome aberrations involving 10q22: report of three overlapping interstitial deletions and a balanced translocation disrupting C10orf11. Eur J Hum Genet 18(3):291–295
467. Tzschach A, Krause-Plonka I, Menzel C, Knoblauch A, Toennies H, Hoeltzenbein M et al (2006) Molecular cytogenetic analysis of a de novo interstitial chromosome 10q22 deletion. Am J Med Genet A 140(10):1108–1110
468. Wada S, Hamada M, Kobayashi K, Satoh N (2008) Novel genes involved in canonical Wnt/beta-catenin signaling pathway in early Ciona intestinalis embryos. Dev Growth Differ 50(4):215–227
469. Prasad A, Merico D, Thiruvahindrapuram B, Wei J, Lionel AC, Sato D et al (2012) A discovery resource of rare copy number variations in individuals with autism spectrum disorder. G3 (Bethesda) 2(12):1665–1685
470. Lo-Castro A, Brancati F, Digilio MC, Garaci FG, Bollero P, Alfieri P et al (2013) Neurobehavioral phenotype observed in KBG syndrome caused by ANKRD11 mutations. Am J Med Genet B Neuropsychiatr Genet 162(1):17–23
471. Sacharow S, Li D, Fan YS, Tekin M (2012) Familial 16q24.3 microdeletion involving ANKRD11 causes a KBG-like syndrome. Am J Med Genet A 158A(3):547–552
472. Willemsen MH, Fernandez BA, Bacino CA, Gerkes E, de Brouwer AP, Pfundt R et al (2010) Identification of ANKRD11 and ZNF778 as candidate genes for autism and variable cognitive impairment in the novel 16q24.3 microdeletion syndrome. Eur J Hum Genet 18(4): 429–435
473. Neilsen PM, Cheney KM, Li CW, Chen JD, Cawrse JE, Schulz RB et al (2008) Identification of ANKRD11 as a p53 coactivator. J Cell Sci 121(Pt 21):3541–3552
474. Noll JE, Jeffery J, Al-Ejeh F, Kumar R, Khanna KK, Callen DF et al (2012) Mutant p53 drives multinucleation and invasion through a process that is suppressed by ANKRD11. Oncogene 31(23):2836–2848
475. Cocchella A, Malacarne M, Forzano F, Marciano C, Pierluigi M, Perroni L et al (2010) The refinement of the critical region for the 2q31.2q32.3 deletion syndrome indicates candidate genes for mental retardation and speech impairment. Am J Med Genet B Neuropsychiatr Genet 153B(7):1342–1346
476. Mencarelli MA, Caselli R, Pescucci C, Hayek G, Zappella M, Renieri A et al (2007) Clinical and molecular characterization of a patient with a 2q31.2-32.3 deletion identified by array-CGH. Am J Med Genet A 143A(8):858–865
477. Prontera P, Bernardini L, Stangoni G, Capalbo A, Rogaia D, Ardisia C et al (2009) 2q31.2q32.3 deletion syndrome: report of an adult patient. Am J Med Genet A 149A(4): 706–712
478. Van Buggenhout G, Van Ravenswaaij-Arts C, Mc Maas N, Thoelen R, Vogels A, Smeets D et al (2005) The del(2)(q32.2q33) deletion syndrome defined by clinical and molecular characterization of four patients. Eur J Med Genet 48(3):276–289
479. Bovolenta P, Esteve P, Ruiz JM, Cisneros E, Lopez-Rios J (2008) Beyond Wnt inhibition: new functions of secreted Frizzled-related proteins in development and disease. J Cell Sci 121(Pt 6):737–746
480. Ekstrom EJ, Sherwood V, Andersson T (2011) Methylation and loss of Secreted Frizzled-Related Protein 3 enhances melanoma cell migration and invasion. PLoS One 6(4):e18674

481. Qian D, Jones C, Rzadzinska A, Mark S, Zhang X, Steel KP et al (2007) Wnt5a functions in planar cell polarity regulation in mice. Dev Biol 306(1):121–133
482. Yamada A, Iwata T, Yamato M, Okano T, Izumi Y (2013) Diverse functions of secreted frizzled-related proteins in the osteoblastogenesis of human multipotent mesenchymal stromal cells. Biomaterials 34(13):3270–3278
483. Poliak S, Gollan L, Martinez R, Custer A, Einheber S, Salzer JL et al (1999) Caspr2, a new member of the neurexin superfamily, is localized at the juxtaparanodes of myelinated axons and associates with K+ channels. Neuron 24(4):1037–1047
484. Rasband MN (2011) Composition, assembly, and maintenance of excitable membrane domains in myelinated axons. Semin Cell Dev Biol 22(2):178–184
485. Savvaki M, Panagiotaropoulos T, Stamatakis A, Sargiannidou I, Karatzioula P, Watanabe K et al (2008) Impairment of learning and memory in TAG-1 deficient mice associated with shorter CNS internodes and disrupted juxtaparanodes. Mol Cell Neurosci 39(3):478–490
486. Anderson GR, Galfin T, Xu W, Aoto J, Malenka RC, Sudhof TC (2012) Candidate autism gene screen identifies critical role for cell-adhesion molecule CASPR2 in dendritic arborization and spine development. Proc Natl Acad Sci U S A 109(44):18120–18125
487. Bel C, Oguievetskaia K, Pitaval C, Goutebroze L, Faivre-Sarrailh C (2009) Axonal targeting of Caspr2 in hippocampal neurons via selective somatodendritic endocytosis. J Cell Sci 122(Pt 18):3403–3413
488. von Bartheld CS (2004) Axonal transport and neuronal transcytosis of trophic factors, tracers, and pathogens. J Neurobiol 58(2):295–314
489. Von Bartheld CS, Altick AL (2011) Multivesicular bodies in neurons: distribution, protein content, and trafficking functions. Prog Neurobiol 93(3):313–340
490. Saito N, Okada Y, Noda Y, Kinoshita Y, Kondo S, Hirokawa N (1997) KIFC2 is a novel neuron-specific C-terminal type kinesin superfamily motor for dendritic transport of multivesicular body-like organelles. Neuron 18(3):425–438
491. Cooney JR, Hurlburt JL, Selig DK, Harris KM, Fiala JC (2002) Endosomal compartments serve multiple hippocampal dendritic spines from a widespread rather than a local store of recycling membrane. J Neurosci 22(6):2215–2224
492. Birnbaum RY, Clowney EJ, Agamy O, Kim MJ, Zhao J, Yamanaka T et al (2012) Coding exons function as tissue-specific enhancers of nearby genes. Genome Res 22(6):1059–1068
493. Phelan K, McDermid HE (2012) The 22q13.3 deletion syndrome (Phelan-McDermid Syndrome). Mol Syndromol 2(3–5):186–201
494. Peca J, Feliciano C, Ting JT, Wang W, Wells MF, Venkatraman TN et al (2011) Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. Nature 472(7344): 437–442
495. Uchino S, Waga C (2013) SHANK3 as an autism spectrum disorder-associated gene. Brain Dev 35(2):106–110
496. Schmeisser MJ, Grabrucker AM, Bockmann J, Boeckers TM (2009) Synaptic cross-talk between N-methyl-D-aspartate receptors and LAPSER1-beta-catenin at excitatory synapses. J Biol Chem 284(42):29146–29157
497. Tobaben S, Sudhof TC, Stahl B (2000) The G protein-coupled receptor CL1 interacts directly with proteins of the Shank family. J Biol Chem 275(46):36204–36210
498. Bonaglia MC, Marelli S, Novara F, Commodaro S, Borgatti R, Minardo G et al (2010) Genotype-phenotype relationship in three cases with overlapping 19p13.12 microdeletions. Eur J Hum Genet 18(12):1302–1309
499. Boucard AA, Ko J, Sudhof TC (2012) High affinity neurexin binding to cell adhesion G-protein-coupled receptor CIRL1/latrophilin-1 produces an intercellular adhesion complex. J Biol Chem 287(12):9399–9413
500. Bena F, Bruno DL, Eriksson M, van Ravenswaaij-Arts C, Stark Z, Dijkhuizen T et al (2013) Molecular and clinical characterization of 25 individuals with exonic deletions of NRXN1 and comprehensive review of the literature. Am J Med Genet B Neuropsychiatr Genet 162(4): 388–403

501. Silva JP, Lelianova VG, Ermolyuk YS, Vysokov N, Hitchen PG, Berninghausen O et al (2011) Latrophilin 1 and its endogenous ligand Lasso/teneurin-2 form a high-affinity transsynaptic receptor pair with signaling capabilities. Proc Natl Acad Sci U S A 108(29):12113–12118
502. Zweier C, de Jong EK, Zweier M, Orrico A, Ousager LB, Collins AL et al (2009) CNTNAP2 and NRXN1 are mutated in autosomal-recessive Pitt-Hopkins-like mental retardation and determine the level of a common synaptic protein in Drosophila. Am J Hum Genet 85(5):655–666
503. Lansbergen G, Grigoriev I, Mimori-Kiyosue Y, Ohtsuka T, Higa S, Kitajima I et al (2006) CLASPs attach microtubule plus ends to the cell cortex through a complex with LL5beta. Dev Cell 11(1):21–32
504. Maas C, Torres VI, Altrock WD, Leal-Ortiz S, Wagh D, Terry-Lorenzo RT et al (2012) Formation of Golgi-derived active zone precursor vesicles. J Neurosci 32(32):11095–11108
505. Barrow J (2011) Wnt/planar cell polarity signaling: an important mechanism to coordinate growth and patterning in the limb. Organogenesis 7(4):260–266
506. Duijf PH, van Bokhoven H, Brunner HG (2003) Pathogenesis of split-hand/split-foot malformation. Hum Mol Genet 12 Spec No 1:R51–R60
507. Guerrini L, Costanzo A, Merlo GR (2011) A symphony of regulations centered on p63 to control development of ectoderm-derived structures. J Biomed Biotechnol 2011:864904
508. Niedrist D, Lurie IW, Schinzel A (2009) 4q32-q35 and 6q16-q22 are valuable candidate regions for split hand/foot malformation. Eur J Hum Genet 17(8):1086–1091
509. Dai L, Deng Y, Li N, Xie L, Mao M, Zhu J (2013) Discontinuous microduplications at chromosome 10q24.31 identified in a Chinese family with split hand and foot malformation. BMC Med Genet 14:45
510. Lyle R, Radhakrishna U, Blouin JL, Gagos S, Everman DB, Gehrig C et al (2006) Split-hand/split-foot malformation 3 (SHFM3) at 10q24, development of rapid diagnostic methods and gene expression from the region. Am J Med Genet A 140(13):1384–1395
511. Hyon C, Marlin S, Chantot-Bastaraud S, Mabboux P, Beaujard MP, Al Ageeli E et al (2011) A new 17p13.3 microduplication including the PAFAH1B1 and YWHAE genes resulting from an unbalanced X;17 translocation. Eur J Med Genet 54(3):287–291
512. Lezirovitz K, Maestrelli SR, Cotrim NH, Otto PA, Pearson PL, Mingroni-Netto RC (2008) A novel locus for split-hand/foot malformation associated with tibial hemimelia (SHFLD syndrome) maps to chromosome region 17p13.1-17p13.3. Hum Genet 123(6):625–631
513. Ostergaard JR, Graakjaer J, Brandt C, Birkebaek NH (2012) Further delineation of 17p13.3 microdeletion involving CRK. The effect of growth hormone treatment. Eur J Med Genet 55(1):22–26
514. Sutton VR, Van den Veyver IB (1993) Focal dermal hypoplasia. In: Pagon RA, Adam MP, Bird TD, Dolan CR, Fong CT, Stephens K, editors. GeneReviews™ [Internet]. Seattle (WA): University of Washington, Seattle; 1993–2013
515. Grzeschik KH, Bornholdt D, Oeffner F, Konig A, del Carmen Boente M, Enders H et al (2007) Deficiency of PORCN, a regulator of Wnt signaling, is associated with focal dermal hypoplasia. Nat Genet 39(7):833–835
516. Wang X, Reid Sutton V, Omar Peraza-Llanes J, Yu Z, Rosetta R, Kou YC et al (2007) Mutations in X-linked PORCN, a putative regulator of Wnt signaling, cause focal dermal hypoplasia. Nat Genet 39(7):836–838
517. Winkel A, Stricker S, Tylzanowski P, Seiffart V, Mundlos S, Gross G et al (2008) Wnt-ligand-dependent interaction of TAK1 (TGF-beta-activated kinase-1) with the receptor tyrosine kinase Ror2 modulates canonical Wnt-signalling. Cell Signal 20(11):2134–2144
518. Passarge E (2000) A distinctive phenotype associated with an interstitial deletion 6q14 contained within a de novo pericentric inversion 6 (p11.2q15). Cytogenet Cell Genet 91(1–4):192–198

519. Poot M, van't Slot R, Leupert R, Beyer V, Passarge E, Haaf T (2009) Three de novo losses and one insertion within a pericentric inversion of chromosome 6 in a patient with complete absence of expressive speech and reduced pain perception. Eur J Med Genet 52(1):27–30
520. Aten E, den Hollander N, Ruivenkamp C, Knijnenburg J, van Bokhoven H, den Dunnen J et al (2009) Split hand-foot malformation, tetralogy of Fallot, mental retardation and a 1 Mb 19p deletion-evidence for further heterogeneity? Am J Med Genet A 149A(5):975–981
521. Bens S, Haake A, Tonnies H, Vater I, Stephani U, Holterhus PM et al (2011) A de novo 1.1Mb microdeletion of chromosome 19p13.11 provides indirect evidence for EPS15L1 to be a strong candidate for split hand split foot malformation. Eur J Med Genet 54(5):e501–e504
522. Pozzi B, Amodio S, Lucano C, Sciullo A, Ronzoni S, Castelletti D et al (2012) The endocytic adaptor Eps15 controls marginal zone B cell numbers. PLoS One 7(11):e50818
523. Ianakiev P, Kilpatrick MW, Toudjarska I, Basel D, Beighton P, Tsipouras P (2000) Split-hand/split-foot malformation is caused by mutations in the p63 gene on 3q27. Am J Hum Genet 67(1):59–66
524. Drewelus I, Gopfert C, Hippel C, Dickmanns A, Damianitsch K, Pieler T et al (2010) p63 antagonizes Wnt-induced transcription. Cell Cycle 9(3):580–587
525. Gripp KW, Ennis S, Napoli J (2013) Exome analysis in clinical practice: expanding the phenotype of bartsocas-papas syndrome. Am J Med Genet A 161(5):1058–1063
526. Mitchell K, O'Sullivan J, Missero C, Blair E, Richardson R, Anderson B et al (2012) Exome sequence identifies RIPK4 as the Bartsocas-Papas syndrome locus. Am J Hum Genet 90(1):69–75
527. Huang X, McGann JC, Liu BY, Hannoush RN, Lill JR, Pham V et al (2013) Phosphorylation of Dishevelled by protein kinase RIPK4 regulates Wnt signaling. Science 339(6126):1441–1445
528. Dimitrov B, Balikova I, de Ravel T, Van Esch H, De Smedt M, Baten E et al (2011) 2q31.1 microdeletion syndrome: redefining the associated clinical phenotype. J Med Genet 48(2):98–104
529. Goodman FR, Majewski F, Collins AL, Scambler PJ (2002) A 117-kb microdeletion removing HOXD9-HOXD13 and EVX2 causes synpolydactyly. Am J Hum Genet 70(2):547–555
530. Robledo RF, Rajan L, Li X, Lufkin T (2002) The Dlx5 and Dlx6 homeobox genes are essential for craniofacial, axial, and appendicular skeletal development. Genes Dev 16(9):1089–1101
531. van Silfhout AT, van den Akker PC, Dijkhuizen T, Verheij JB, Olderode-Berends MJ, Kok K et al (2009) Split hand/foot malformation due to chromosome 7q aberrations(SHFM1): additional support for functional haploinsufficiency as the causative mechanism. Eur J Hum Genet 17(11):1432–1438
532. Feenstra I, Hanemaaijer N, Sikkema-Raddatz B, Yntema H, Dijkhuizen T, Lugtenberg D et al (2011) Balanced into array: genome-wide array analysis in 54 patients with an apparently balanced de novo chromosome rearrangement and a meta-analysis. Eur J Hum Genet 19(11):1152–1160
533. Nakashima N, Yamagata T, Mori M, Kuwajima M, Suwa K, Momoi MY (2010) Expression analysis and mutation detection of DLX5 and DLX6 in autism. Brain Dev 32(2):98–104
534. Feng J, Bi C, Clark BS, Mady R, Shah P, Kohtz JD (2006) The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. Genes Dev 20(11):1470–1484
535. Kouwenhoven EN, van Heeringen SJ, Tena JJ, Oti M, Dutilh BE, Alonso ME et al (2010) Genome-wide profiling of p63 DNA-binding sites identifies an element that regulates gene expression during limb development in the 7q21 SHFM1 locus. PLoS Genet 6(8):e1001065
536. Lewandoski M, Sun X, Martin GR (2000) Fgf8 signalling from the AER is essential for normal limb development. Nat Genet 26(4):460–463
537. Mariani FV, Ahn CP, Martin GR (2008) Genetic evidence that FGFs have an instructive role in limb proximal-distal patterning. Nature 453(7193):401–405
538. Corballis MC (2009) Language as gesture. Hum Mov Sci 28(5):556–565
539. Kalil K, Li L, Hutchins BI (2011) Signaling mechanisms in cortical axon growth, guidance, and branching. Front Neuroanat 5:62

540. Keeble TR, Halford MM, Seaman C, Kee N, Macheda M, Anderson RB et al (2006) The Wnt receptor Ryk is required for Wnt5a-mediated axon guidance on the contralateral side of the corpus callosum. J Neurosci 26(21):5840–5848
541. Thomas C, Humphreys K, Jung KJ, Minshew N, Behrmann M (2011) The anatomy of the callosal and visual-association pathways in high-functioning autism: a DTI tractography study. Cortex 47(7):863–873
542. Wen J, Yang HB, Zhou B, Lou HF, Duan S (2013) beta-Catenin is critical for cerebellar foliation and lamination. PLoS One 8(5):e64451
543. Rogers TD, McKimm E, Dickson PE, Goldowitz D, Blaha CD, Mittleman G (2013) Is autism a disease of the cerebellum? An integration of clinical and pre-clinical research. Front Syst Neurosci 7:15
544. Baron-Cohen S (2010) Empathizing, systemizing, and the extreme male brain theory of autism. Prog Brain Res 186:167–175
545. Breedlove SM (2010) Minireview: organizational hypothesis: instances of the fingerpost. Endocrinology 151(9):4116–4122
546. Manning JT, Baron-Cohen S, Wheelwright S, Sanders G (2001) The 2nd to 4th digit ratio and autism. Dev Med Child Neurol 43(3):160–164
547. Teatero ML, Netley C (2013) A critical review of the research on the extreme male brain theory and digit ratio (2D:4D). J Autism Dev Disord, doi:10.1007/s10803-013-1819-6
548. Zheng Z, Cohn MJ (2011) Developmental basis of sexually dimorphic digit ratios. Proc Natl Acad Sci U S A 108(39):16289–16294
549. Wegiel J, Kuchna I, Nowicki K, Imaki H, Yong Ma S, Azmitia EC et al (2013) Contribution of olivofloccular circuitry developmental defects to atypical gaze in autism. Brain Res 1512:106–122
550. Bauman ML, Kemper TL (2005) Neuroanatomic observations of the brain in autism: a review and future directions. Int J Dev Neurosci 23(2–3):183–187
551. Schmitt JE, Eliez S, Warsofsky IS, Bellugi U, Reiss AL (2001) Enlarged cerebellar vermis in Williams syndrome. J Psychiatr Res 35(4):225–229
552. Hashimoto M, Ito R, Kitamura N, Namba K, Hisano Y (2012) Epha4 controls the midline crossing and contralateral axonal projections of inferior olive neurons. J Comp Neurol 520(8):1702–1720
553. Losin EA, Iacoboni M, Martin A, Dapretto M (2012) Own-gender imitation activates the brain's reward circuitry. Soc Cogn Affect Neurosci 7(7):804–810
554. Cossu G, Boria S, Copioli C, Bracceschi R, Giuberti V, Santelli E et al (2012) Motor representation of actions in children with autism. PLoS One 7(9):e44779
555. Williams JH, Whiten A, Singh T (2004) A systematic review of action imitation in autistic spectrum disorder. J Autism Dev Disord 34(3):285–299

Chapter 7
Stuttered Birdsong

Santosh A. Helekar, Delanthi Salgado-Commissariat, David B. Rosenfield, and Henning U. Voss

Abstract Birdsong is analogous to speech in terms of its role in communication, vocal motor control, auditory perception, and development. Songbirds such as zebra finches can therefore be used to model speech motor control disorders. In this chapter, we describe our efforts at developing a variant form of zebra finch song containing syllable repetitions that resemble part-word repetitions of developmental stuttering. We further discuss functional magnetic resonance imaging experiments that reveal changes in neural activations produced by song stimuli in syllable repeater birds. Finally, we present findings and review data to propose that synaptic plasticity and neuromodulatory mechanisms might play a role in the development of repetitive or oscillatory vocal output.

Keywords Dysfluency • Stuttering • Syllable repetition • Vocal motor control • Auditory perception • Synaptic plasticity • Neuromodulation • Songbird • fMRI

S.A. Helekar, M.B.B.S., Ph.D. (✉)
Department of Neurology and Neuroscience, The Methodist Hospital Research Institute and Weill Medical College of Cornell University, 6565 Fannin Street, F770A, Houston, TX 77030, USA
e-mail: sahelekar@tmhs.org

D. Salgado-Commissariat, B.S., Ph.D.
Comparative Medicine Program, The Methodist Hospital Research Institute, Houston, TX, USA

D.B. Rosenfield, M.D.
Department of Neurology, The Methodist Hospital Neurological Institute, 6560 Fannin Street, Ste 802, Houston, TX 77030, USA

H.U. Voss, Ph.D.
Department of Radiology, Weill Medical College of Cornell University, New York, NY, USA

S.A. Helekar (ed.), *Animal Models of Speech and Language Disorders*,
DOI 10.1007/978-1-4614-8400-4_7,

Introduction

Developmental stuttering is the most common disorder of speech motor control. More than 1 % of the adults and 4 % of the children in the United States are affected by it [1]. It begins during early childhood years coinciding with rapid language, speech, and neuromuscular maturation. It is characterized by intermittent fluency failures that take the form of blocks, repetitions, and prolongation of initial sounds or syllables of words. The majority of children recover without any therapy. There is a great deal of published research on speech production in human stuttering [2]. A theoretical basis for the motor control deficit in stuttering has been proposed by Rosenfield and coworkers [3–5]. However, the neurobiological basis of this disorder is poorly understood. Because of the essential dependence of speech on language, a uniquely human attribute, there is a complete lack of animal models of speech disorders. Such a model, albeit simplistic and devoid of the linguistic component of speech, might still represent the vocal motor component of speech and therefore serve to provide insights into the neural mechanisms underlying the vocal motor aspects of a dysfunction such as stuttering.

In this chapter, we describe our work on the first such model in the songbird zebra finch. Since persistent modification/modulation of vocal motor control circuits is causally significant to oscillatory or repetitive motor output that might be at the crux of the underlying problem in stuttering, we also review relevant findings in the neurophysiology of synaptic plasticity of song control pathways in the songbird brain. We do not discuss the neurobiology of the songbird song and auditory systems, apart from mentioning some facts directly relevant to our results, because it has been reviewed in detail in Chaps. 3–6 of this book.

Possible Minimal Model of Stuttering in the Songbird

Our laboratory experiments in zebra finches show that variant forms of their song pattern, consisting of song syllable repetitions, may provide us with a minimal model of developmental stuttering. Song syllable repetitions, at least superficially, resemble part-word repetitions, which are a common feature of stuttering, essentially because both types of repetition are characterized by involuntary or compulsive oscillations of vocal output and both can be induced by delayed auditory feedback in subjects that otherwise produce normal vocal output [6, 7]. As in children learning to speak, zebra finches depend on a long critical period of tutor-based learning for the acquisition of birdsong [8, 9]. Each adult male zebra finch produces stereotyped song that is spectrally and temporally distinct [10]. However, studies done in our laboratory among others have identified a significant amount of variability in the acoustic profiles of adult zebra finch song [11, 12]. The extent to which this variability is generated or influenced by learning remains to be determined.

Work in several laboratories investigating songbird behavior and neurobiology has shown that various experimental manipulations can cause zebra finches to produce abnormal stuttered or repeated song syllables. Leonardo and Konishi were

able to do this by altering the auditory feedback [7]. Tchernichovski and coworkers observed this phenomenon by overstimulation with a tutor song model during developmental learning [13] (personal communication). Recording of air sac pressure changes after partial muting or decrease in phonation in adult birds showed stuttering of expiratory pulses inserted into the song motif [14]. A detailed study of the time course of song deterioration immediately after deafening has also shown stuttering of syllables in birds that had some syllable repetitions in their song motifs to begin with [15]. Our work has focused on studying the development of stuttered song by manipulating the tutoring environment of young zebra finches, to attempt to develop a possible animal model of human stuttering.

Variable Song Syllable Repetitions and Their Development

We have observed that ~7 % of normally raised adult male zebra finches produce a deviant pattern of song consisting of motifs with abnormal repetitions of syllables [16]. We called these birds spontaneous repeaters because they were not tutored by any adults who were repeaters themselves. A syllable repetition episode within a motif consists of sequences of 3–16 repeated occurrences of a single syllable, as determined by its spectral and temporal profile (Fig. 7.1). We observed that in more than 72 % of the 20 min song recording sessions of spontaneous repeaters, 10–92 % of their song motifs contained at least one repetition episode. The interval between the end of an adjacent syllable in a motif and the beginning of a repetition episode is at least ten times smaller than for distance calls that sometimes accompany songs. Furthermore, syllable repetitions differ from introductory notes, a common feature of the initiation of singing, in that, unlike the former, the latter show highly variable inter-note intervals. These observations indicate that in laboratory-bred zebra finches, songs containing syllable repetitions might be a distinctive developmental variant of the temporal pattern of song.

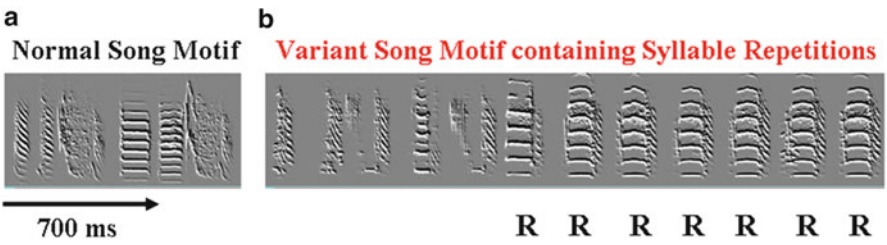


Fig. 7.1 Representative spectrograms of female-directed song motifs from a normal singer and a song syllable repeater. (**a**) Spectrogram of a normal non-repeater zebra finch. The abscissa represents time in seconds and the ordinate frequency in kilohertz. (**b**) Spectrogram of a syllable repeater. R's underneath the spectrogram denote repeated syllables. *From Voss HU, Salgado-Commissariat D, Helekar SA. Altered auditory BOLD response to conspecific birdsong in zebra finches with stuttered syllables. PLoS One 2010 Dec 23;5(12):e14415 (32) Creative common license*

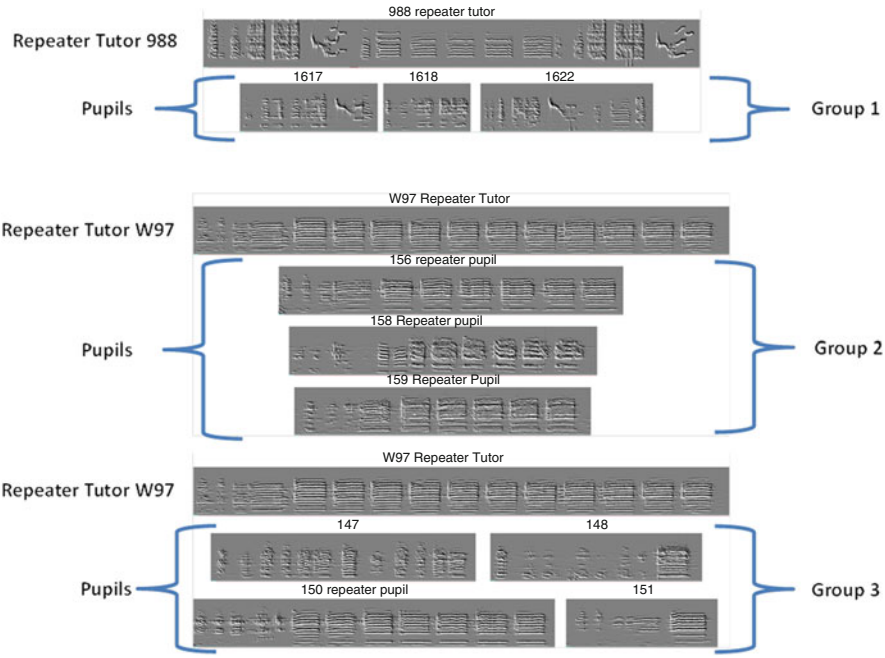


Fig. 7.2 Representative spectrograms of song motifs in three groups of tutored repeater and non-repeater pupils and their repeater tutors. In group 1, all three pupils are non-repeaters. In group 2, all three are repeaters, and in group 3, two are repeaters and two non-repeaters

Recordings of female-directed songs carried out at ~120 days in normal birds that were tutored during the developmental critical period by repeater tutors reveal that ~60 % of them produce syllable repetitions as integral components of their song motifs. We call these birds tutored repeaters. In eight groups of siblings consisting of 2–4 pupils tutored by repeater tutors, the fractions of birds that became tutored repeaters ranged from 33 to 66 %. Figure 7.2 shows representative song spectrograms from three groups of pupils of two repeater tutors. In tutored repeaters, syllable repetitions occur in almost 100 % of the motifs that they produce. Their repeated syllables show <40 % similarity (measured using Sound Analysis 3) [17] with the syllable repeated by their tutor, indicating that it is unlikely that the repeated syllable is simply copied from tutors in every case. Instead, in many cases, it is the tendency to repeat that is learned, and it persists into adulthood.

Adult-Phase Song Plasticity Involving Syllable Repetitions

In tutored repeaters that were exposed on a long-term basis to songs of non-repeaters after the critical period of learning, we found a progressive reduction of the mean number and variance of repeated syllables per song motif in 60 % of such birds over a period of 3–8 weeks. This demonstrates that song plasticity in adulthood can

produce recovery of songs to a normal pattern. Notably, spontaneous repeaters when exposed to normal singers as adults for the same duration do not show this reversal to normalcy. This change is similar to that occurring with recovery from abnormal alterations in auditory feedback [7]. A reduction in mean number of repeated syllables per motif is also seen in birds that were acoustically isolated, but to a lesser extent (16.4 %), than birds exposed to normal singers (39.7 %), suggesting that exposure to non-repeater songs might influence this restorative plasticity. In contrast, 40 % of birds exposed to non-repeater songs show no such reduction in number and variance of syllable repetitions. The latter birds therefore are similar to spontaneous repeaters and might share common underlying neural mechanisms with them, possibly involving a deficiency in the ability to undergo synaptic plasticity in the adult or post-critical period phase of life and/or a similar deficiency in song learning [16]. An analysis of the sequence of syllables within the motifs of syllable repeaters and non-repeaters shows that spontaneous repeaters do not possess highly variable motif sequences. Observations over an extended period of 18 months in tutored repeaters also show no significant changes in measured song sequence parameters. Therefore, the only parameter that is changed in ~60 % of the birds as a reflection of adult-phase song plasticity is the mean number of repetitions per motif. Consequently, there appears to be no substantial reorganization of song in tutored repeaters due to exposure to normal singers or simply due to passage of time.

The above findings enable us to conclude the following: (1) The temporal frame of song in repeaters may be learned independent of its spectral content. (2) Learning of variant song output involving repeated utterances might be determined by innate constraints, considering that a significant number of birds do not learn to produce such an output despite being tutored to do so. (3) A form of adult-phase plasticity involving the restoration to normalcy of variant song output can occur under the influence of a normal song environment. These conclusions might have significant implications for understanding the development of and recovery from the tendency of stutterers to repeat word fragments, providing a possible animal model for some important motor control aspects of stuttering and other human dysfluencies.

To account for the emergence of spontaneous repeaters and the observed bidirectional outcome in normal juvenile birds of deliberate tutoring by repeaters, we have proposed that the three different categories of birds in our experiments might be regarded as three different points on a scale with tutored non-repeaters and spontaneous repeaters at opposite ends and tutored repeaters lying in the middle. The assumption here is that there is a common mechanism underlying syllable repetitions in the two types of repeaters, involving a motor, sensory, or sensorimotor process. Since 40–50 % of juvenile zebra finches resist learning of repetitions despite being exposed to them in their tutor, this common mechanism could be a variable sensitivity or susceptibility mechanism with tutored non-repeaters, tutored repeaters, and spontaneous repeaters located on a sliding scale of increasing susceptibility [16]. Possible causal mechanisms for the emergence of syllable repetitions could involve some deficiency in plasticity mechanisms underlying learning, in auditory perception, or in the auditory feedback resulting from the song output—a mechanism that has been implicated in the mediation of adult-phase song plasticity associated with bilateral cochlear ablation [18].

Significance of Syllable Repetitions

Studies on zebra finches producing songs with abnormal song syllable repetitions are of significance from three important points of view. First, they offer us a model system to study the interactions between developmental or genetic predispositions and tutor-dependent influences on the learning of vocal communication. Second, they could provide us with zebra finch lineages with possible genetic predisposition for producing variant song patterns consisting of song syllable repetitions, a potential songbird homologue of part-word repetitions of human stuttering. These studies would enable us in the long run to explore cellular and molecular mechanisms underlying the learning and production of these vocal patterns and the mechanisms governing the adult-phase vocal neuromotor plasticity associated with them. Finally, the overall investigative approach that these studies embody might contribute toward the development of a useful animal model of speech motor control disorders. Such an animal model might hold the promise of suggesting a rational basis for prospective therapies for some forms of human dysfluency, such as stuttering. Moreover, studies on long-term adult-phase changes in song and the neuroplastic changes giving rise to them might be relevant to understanding the pathogenesis as well as rehabilitative recovery of human dysfluencies and other speech impairments. With regard to the issue of similarity between part-word repetitions in stuttering and song syllable repetitions, it might be noted that part-word repetitions involve the initial fragment of a word. However, this word does not necessarily occur at the beginning of an uttered sentence. It could occur in the middle or at the terminal portion of a sentence. Additionally, within the word itself, if it is a two-part compound word, such as "airport," the repetitive element could occur in its terminal half.

Song syllable repetitions are analogous to part-word repetitions in three important respects: (a) They involve an involuntary repetitive vocal motor output. (b) Each of them is a repetition of a subcomponent of a larger utterance, which, if repeated, does not constitute an abnormality. Repetition of words is part of normal speech and is quite common in children. Repetition of song motifs is a feature of normal zebra finch song. (c) Both part-word repetitions and song syllable repetitions can result from disruption of auditory feedback. One important difference between persistent song syllable repetitions and part-word repetitions in human stuttering, nevertheless, is that the latter are thought to be a response to a speech motor control deficit [5] where the role of learning is not clear, while the former are a consequence of tutor-based learning, presumably combined with an inability to revert to normal song.

We have speculated above that reduction in song syllable repetitions in a large fraction of tutored repeaters as adults might be related to the adult-phase plasticity seen due to alteration of auditory feedback and during its restoration to normalcy [7]. Can this amelioration of the tendency to repeat syllables be treated as analogous to the recovery seen in the ~3 % of children who outgrow stuttering? Are the remaining birds that do not show such recovery analogous to persistent stutterers? Answers

to these questions would mean that neural mechanisms underlying the development of variant song might lead to insights into the neurophysiology of stuttering. The quantitative measurement of percent similarity between pupil songs and tutor songs has shown that there is no significant difference between repeaters and non-repeaters. This finding indicates that there may not be any measurable difference in song learning or imitation of the tutor song motif between these two types of birds.

Mechanisms Underlying Syllable Repetitions

In our current experiments in persistent tutored repeater birds, we have focused on three possible alternative or mutually complementary hypotheses to account for the emergence of variant song containing stuttering-like repetitive vocal patterns. They are as follows: (1) that the aberrant vocal output might be due to deficient formation of the sensory template of tutor song, (2) that it might reflect an inability to maintain a lasting memory of the learned tutor song template, and (3) that an alteration of the sensory habituation mechanism might be responsible for it. We recognize that many other hypothetical mechanisms need to be investigated, especially those that involve defects in motor control, but we find that the above three possibilities are particularly amenable to the noninvasive imaging and in vitro electrophysiological approaches that have been available to us in our laboratory. We describe the relevant experiments using these approaches below.

Brain Functional Magnetic Resonance Imaging in Songbirds

The most powerful noninvasive imaging modality to study the function of the brain during speech perception, learning, and production is to image the spatiotemporal hemodynamic brain responses to auditory stimulation or speech production by functional MRI (fMRI) in children and adults [19]. In songbirds, analogous fMRI experiments can be performed using either the same imaging equipment as in humans [20] or, preferably, dedicated small-animal MRI scanners [21, 22]. This can be done at fixed time points in adult birds, or in longitudinal studies during development, or even over generations [23, 24]. Whereas in humans fMRI is presently the only viable imaging modality able to capture spatially resolved hemodynamic response over the whole brain noninvasively, in songbirds we have more options. In vivo imaging modalities used for functional imaging of hemodynamic response in the songbird include fMRI and optical imaging, both pioneered by Van der Linden and coworkers [21, 25, 26]. While optical methods are highly promising for certain applications in songbirds [27, 28], even with intact skull [29, 30], almost all in vivo functional imaging studies related to song perception have been performed using nonoptical methods so far [20–22, 24, 25, 31–38].

Auditory Responses in Repeaters vs. Non-repeaters

We studied the mechanisms underlying syllable repetitions by using fMRI in awake sedated zebra finches in order to be able to compare possible functional differences in sensory representation and/or processing of songs between repeaters and non-repeaters at multiple developmental time points. fMRI scans during auditory stimulation show a strong blood-oxygen-level-dependent (BOLD) response in field L, caudal medial nidopallium (NCM), and caudal mesopallium (CM) [20, 22]. Our initial experiments utilized a specially designed radiofrequency coil in a 3 T human whole body scanner [20]. In awake zebra finches, we could record spatial patterns of BOLD responses to female-directed songs that differed in familiarity and significance as a function of auditory experience and song learning history. We observed a difference in fMRI activation of the sensory structures NCM, CM, and field L in response to auditory stimulation between tutored repeaters and non-repeaters [32]. The BOLD response to tutor song is significantly reduced in repeaters compared to non-repeaters (Fig. 7.3). This reduction is unlikely to be due to a general decrease in auditory responsiveness because there is no such decrease in response to pure tone and bird's own song, and the response to an unfamiliar conspecific song is significantly enhanced in repeaters. Since the stimuli were presented in a randomized order, cross-stimulus short-term plasticity effects and systematic changes in sedation or stress levels in the mildly sedated awake birds can also be ruled out as explanations for the weaker tutor song response in repeaters. The selective attenuation of the tutor song response appears to be more pronounced on the right side compared to the left, suggesting a tendency toward lateralization, superficially reminiscent of that seen in fMRI studies in stutterers [39, 40]. However, other results on lateralization of perceptual responses in songbirds provide a complex picture, and lateralization might differ for different brain regions involved in perception [35, 41–46]. The finding of reduced tutor song response in repeaters suggests that the tutor song might be less salient to repeaters or that these birds have a deficient sensory template of tutor song in the auditory structures. It would be important to find out if this difference in BOLD activation developmentally predates the learning of song and to also study the developmental time course of this alteration.

Song learning is thought to depend on the formation of a sensory template of the tutor song. Based on findings of immediate early gene expression studies and multiunit recordings, the higher auditory area NCM has been proposed as a possible site of formation of this template [47–49]. The reduction in response to tutor song in field L suggests that this auditory area might be an additional candidate site for the sensory template. We also found that there is a significant correlation between tutor song response and the degree of similarity of the pupil's song with that of its tutor when both acoustic structure and sequence of syllables in the song motif were taken into account. The bidirectional changes in responsiveness to tutor and conspecific songs suggest that neuroplastic changes in excitatory transmission might be the underlying mechanism that is altered in repeaters. Given that almost half the neurons in NCM are inhibitory [50], synaptic inhibition might also play a role. In this

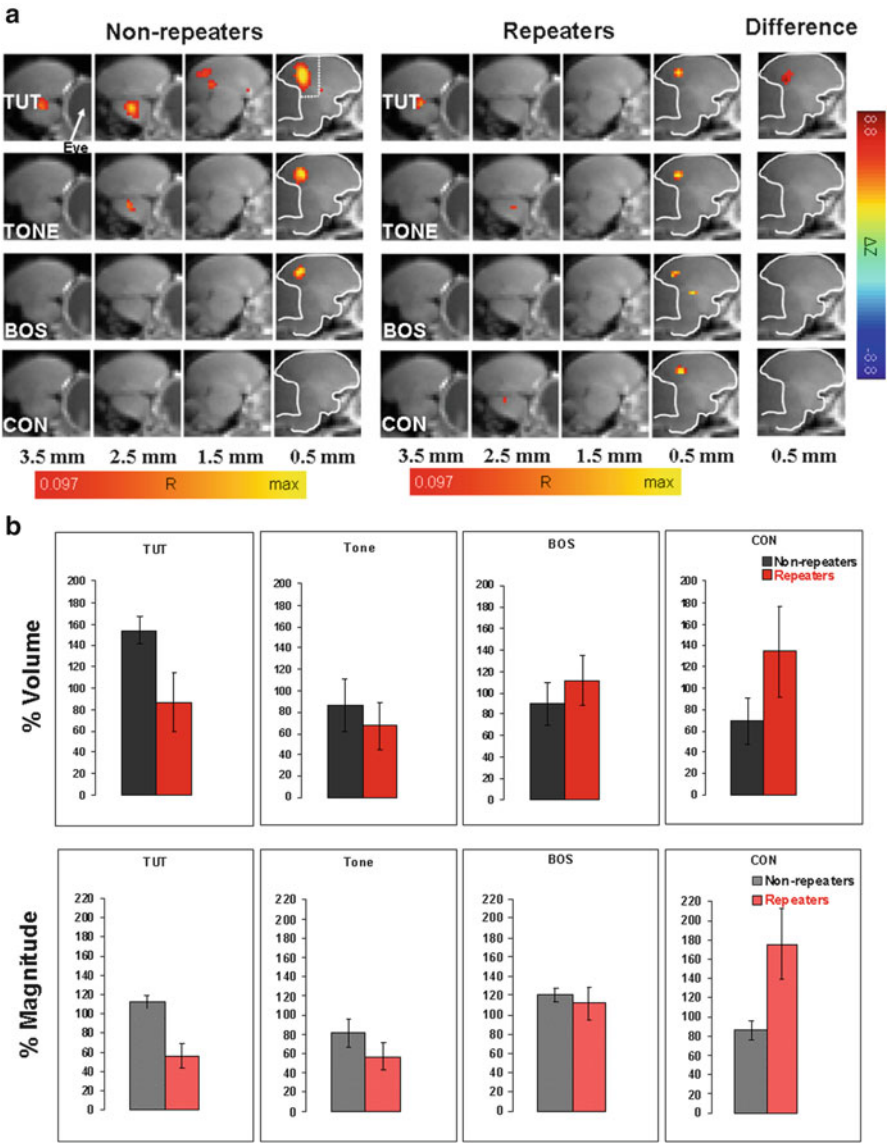


Fig. 7.3 fMRI findings in repeaters and non-repeaters. (**a**) Mean BOLD activation for four different auditory stimuli—tutor song (TUT), 2 kHz pure tone (TONE), bird's own song (BOS), and conspecific song (CON)—for eight non-repeater control birds (non-repeaters) and eight repeater birds (repeaters). Colors denote correlation coefficients *R*, individually scaled in each plot, overlaid to averaged EPI images (*grey*). *White border* depicts the brain template outline. The main activated area that is consistently activated in all images corresponds to NCM, CM, and field L regions. Difference images represent the difference in *z*-values between non-repeaters and repeaters. (**b**) *Plots* showing normalized percent volume and magnitude of the BOLD response to the above four stimuli in repeaters and non-repeaters. The decrease in response to TUT in repeaters was significant for both volume ($p=0.05$) and magnitude ($p=0.003$). The increase in response to CON in repeaters is significant only for magnitude ($p=0.048$). *From Voss HU, Salgado-Commissariat D, Helekar SA. Altered auditory BOLD response to conspecific birdsong in zebra finches with stuttered syllables. PLoS One 2010 Dec 23;5(12):e14415 (32) Creative common license*

context, it should be noted that the hemodynamic response is now known to reflect excitatory and inhibitory postsynaptic potentials, rather than neuronal action potentials [51–53].

Parallels with fMRI Results in Human Stuttering

An important difference between human stuttering and syllable repetitions in our repeater birds is that typically the former occurs at the beginning of words, sentences, or second half of two-part words, whereas the latter in the majority of cases occur at the end of song motifs. However, in acquired stuttering, repetitions can occur at the end of words [52–56]. In terms of neural mechanisms implicated in stuttering in humans and repetitions in zebra finches, the changes that we have observed in repeaters are localized to field L, and possibly to NCM and CM, areas that are analogous to Heschl's gyrus and adjacent auditory association areas, respectively, in the temporal lobe of the human brain. Electrophysiology [54] and functional imaging [55] have shown greater activation of the left superior and middle temporal gyri in stutterers compared to normal controls in a passive listening task [55]. The activity of the left Heschl's gyrus is reduced in stutterers during the production of speech [56]. But in another study, bilateral increase in activation has been observed in this area during speech production, and a decrease related to speech and nonspeech auditory perception is seen in this and other auditory areas [57]. Overall, studies have indicated that stutterers compared to controls show weaker responses in auditory areas to their own speech [58]. Positron emission tomography (PET) studies are also indicative of similar results in auditory and adjacent association areas in the temporal lobe [59]. Speaking tasks causing stuttering events, choral reading or paced speech, and silent reading reveal significant alterations in PET activity in stutterers compared to controls [60–62]. Treatments of stuttering, such as fluency shaping therapy, have shown restorative functional changes in temporal areas of both hemispheres, as well as other areas, but more pronounced on the left side than on the right [40, 63]. In general, the differences in activation in primary auditory and adjacent auditory association areas point to possible deficiencies in self-monitoring of speech, auditory processing of speech, and/or auditory feedback. These processes might bear some relationship to perceptual matching to the auditory template of the tutor song in our repeater birds.

Changes in BOLD Response to Repeated Song Stimulation

In experiments currently underway, we are testing the possibility that syllable repetitions might be due to an alteration in the mechanism of sensory habituation by studying the changes in the BOLD response as a function of familiarization of a song stimulus by its repeated presentation. Figure 7.4 shows the results of our first

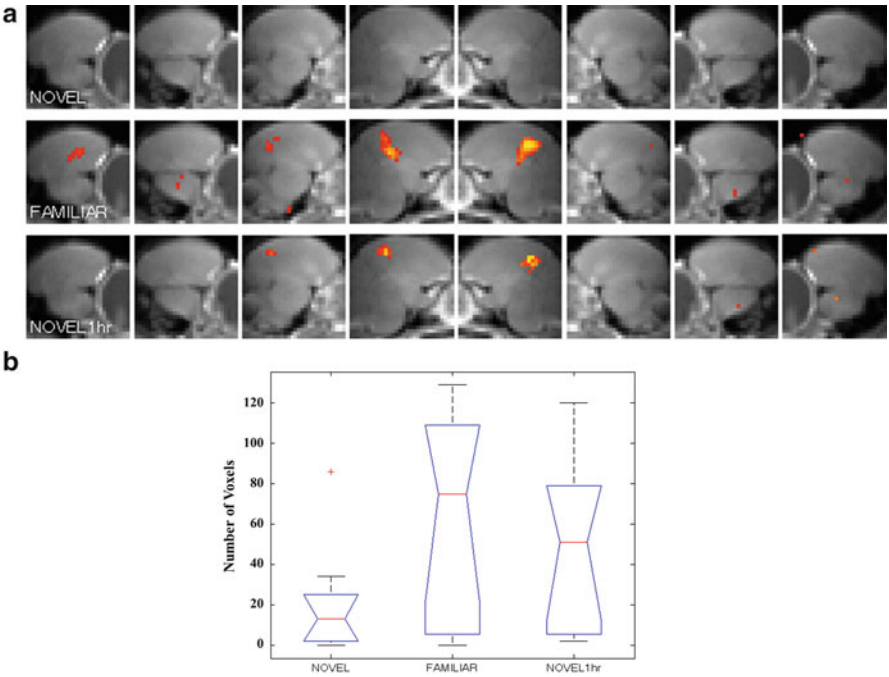


Fig. 7.4 Plasticity of the BOLD response. (**a**) Activations averaged over all nine birds for all slices, going from the left to the right hemisphere (*left to right*), lateral to medial and medial to lateral (*left to right*). *Top* panels represent the first presentation of the novel stimulus (NOVEL); the *middle* panels, a song stimulus that the bird has heard before (FAMILIAR); and the *bottom* panels, the same novel stimulus presented again after 1 hour (NOVEL1hr). The individual activations of initial NOVEL stimulus are averaging out and are not visible here for the chosen significance threshold of $p<0.005$ (multiple tests corrected). (**b**) Number of activated voxels over the medial slices 4 and 5 for the first presentation of a novel song stimulus, presentation of a familiar song, and the second presentation of the novel song 1 h after the first novel song stimulation (*Boxes* are defined as median ±1 quartile; whiskers show data range without outliers, which are denoted by *crosses*; $n=9$)

experiment on nine male non-repeater birds, performed on a 3 T human MRI scanner [64]. A novel stimulus produces a smaller and less consistent BOLD response in the auditory regions of the zebra finch brain. A familiar song stimulus or one that was made familiar by its repetition after 1 h increases its strength and consistency. Specifically, in eight birds, the activation volume in response to novel song presentation 1 h (NOVEL1hr) after its first presentation was significantly ($p=0.04$, Wilcoxon signed-rank test) larger than that in response to the first presentation (NOVEL) itself. NOVEL caused significantly less activation than the first presentation of a familiar song (FAMILIAR). However, the response to NOVEL1hr was not significantly different from the FAMILIAR response ($p=0.21$), reflecting a perceptual familiarization of the novel song. In contrast to the volume of BOLD activation,

its mean relative magnitude did not show a consistent change from NOVEL to NOVEL1hr.

We have now replicated this experiment in anesthetized (1.5–2 % isoflurane) young zebra finches on an animal MRI scanner with 7 T field strength (Avance III BioSpec 70/30 USR, Bruker Biospin MRI, Inc., Billerica, MA). On this type of scanner, songbird fMRI provides a stronger BOLD signal than on human MRI machines [22]. We briefly describe below results from six young (mean age 78 days) birds who were pupils of two different repeater birds. Of these, five birds were non-repeaters, and one was a repeater. Six different stimuli were selected for the first stimulus presentation and then repeated 1 h later. Five of these stimuli were (1) NOV-Rmtf, an unfamiliar (novel) repeater motif, excluding syllable repetitions (same for all birds); (2) NOV-R, the same unfamiliar repeater song motif, but including syllable repetitions (same for all birds); (3) R-TUT, repeater tutor song motif, including syllable repetitions (two different motifs for the two separate groups of pupils, consisting of two and four birds); (4) NOV, an unfamiliar non-repeater song (same for all birds); and (5) TONE, a pure tone with 2 kHz frequency (same for all birds). The meaning of the sixth stimulus changed between birds and, therefore, was not used in across bird comparisons presented here. The six stimuli were played to the birds back to back and always in the same order for each bird, repeated eight times.

Data were analyzed in BrainVoyager QX 2.3 [65]; first, the two initial scans were discarded and the data combined into a volume by applying a 3D co-registration algorithm with trilinear interpolation for detection and a windowed sinc function interpolation for correction, mild spatial smoothing of 0.2 mm in all directions, a temporal high pass of two cycles over the data set for trend removal, and interpolation to an isotropic voxel with 0.3 mm edge size. Echo-planar imaging (EPI) images were manually aligned with an anatomical scan, and the anatomical scan was manually co-registered to an MRI zebra finch atlas [66], which was used as the background for the statistical parametric maps (SPMs). General linear model (GLM) coefficients [67] were estimated for each bird separately by using all six stimulus predictors and a baseline, and the six motion correction parameters were included as nuisance variables. SPMs consisting of t-values were generated for voxels with a false-discovery-rate [68, 69] adjusted significance threshold of $p = 0.05$. SPMs contained response to each of the six stimuli, and individual stimuli were selected for display by choice of the corresponding contrast. Two fixed-effects group analyses were performed by pooling the GLMs of all birds, for the first and second time points each, and displayed for each of the five stimuli by defining the contrasts accordingly. Two random effects analyses were performed by an individual analysis of all positive activation clusters that were found to include parts of field L, using paired *t*-test and analysis of variance (ANOVA). The first random effects analysis was based on the average *t*-value per cluster (or magnitude of activation). The second one was based on the number of voxels included in those clusters (or volume of activation). For each bird separately, these numbers were then normalized by the corresponding average numbers over all stimuli and the two time points, to account for individual differences (in signal to noise ratio) between birds, caused by motion,

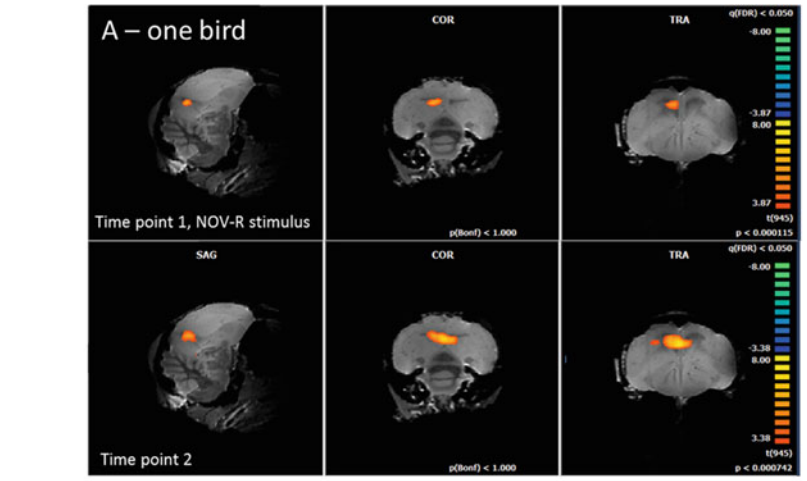
depth of anesthesia, and exact positioning in the scanner and with respect to the radiofrequency coils, among other potential influences.

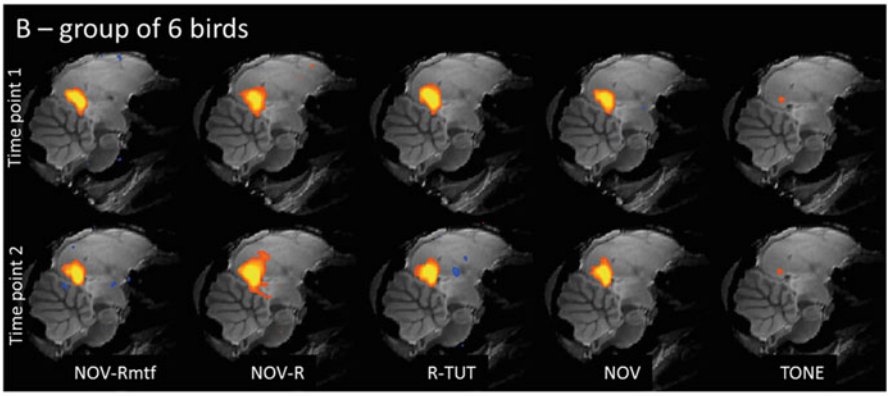
The results of this experiment are summarized in Fig. 7.5. The observed BOLD response in all cases included one main contiguous cluster centered around field L, sometimes partially extending into adjacent areas (see Fig. 7.5a). The fixed-effects group study (Fig. 7.5b) confirmed this finding. Compared to the first scanning session in the second session, 1 h later, the SPMs of individual birds often exhibited larger clusters. A two-way ANOVA with repeated measurement over all stimuli and the two time points resulted in the following: For the BOLD magnitude, p(time) was 0.08, i.e., suggestive of an effect but not significant, p(stimulus) was 0.0000017, and p(interaction) was not significant. For the BOLD volume, p(time) was 0.000014, p(stimulus) was 0.002, and p(interaction) was 0.02. The results of post hoc paired t-tests over time are indicated by asterisks in Fig. 7.5c, d. In summary, an increase in BOLD activation volume is observed over time, an increase in BOLD magnitude is suggestive over time, and there is an interaction between volume increase and type of stimulus. The post hoc tests suggest that the main increase is due to the novel stimuli NOV-Rmtf and NOV-R, although NOV and R-TUT also contribute to the volume increase. This is corroborated by individual volume changes (Fig. 7.5e) for each bird. For NOV-Rmtf and NOV-R all volumes increase, whereas for the other two song stimuli 2–3 volumes stay unchanged or decrease. Note that all stimuli were played to the birds cyclically during the same experiment, so these normalized bird-by-bird comparisons should be relatively independent of the level of anesthesia and other time-dependent factors. Thus, they most likely reflect experience-dependent changes in response to the stimuli. While it is possible that factors unrelated to neuronal activity and plasticity, such as blood pH and pCO_2 (which were not directly measured, unlike the respiratory rate), contribute to the changes observed, the fact that the changes were stimulus specific diminishes the likelihood of this possibility.

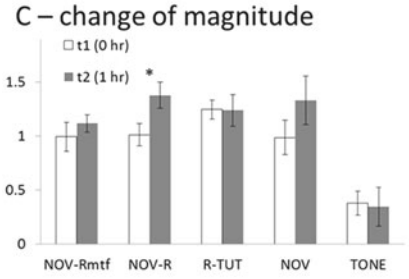
A salient finding of this experiment from the standpoint of understanding the mechanisms related to perception of repeater songs is that the increase in the volume of activation is much greater with respect to repeater song motifs containing syllable repetitions than any other kind of stimulus. Subsequent experiments and analysis of data in repeaters will indicate whether there is a significant difference between repeaters and non-repeaters in terms of these short-term experience-dependent changes.

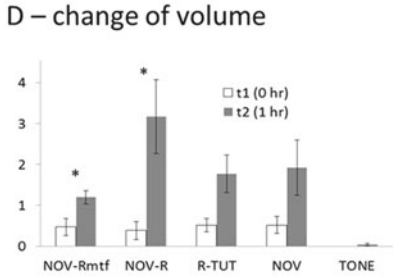
Neuromodulatory Mechanisms in Song Syllable Repetitions and Plasticity

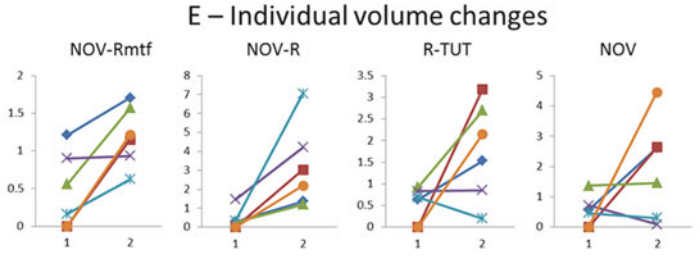
In order to understand mechanisms giving rise to syllable repetitions and adult-phase plasticity, it is necessary to study the physiology of synaptic plasticity in the songbird brain. To date there are only a small number of published studies in this


A – one bird
COR
TRA
q(FDR) < 0.050
-8.00
-3.87
8.00
3.87
t(945)
p < 0.000115
Time point 1, NOV-R stimulus
p(Bonf) < 1.000
SAG
COR
TRA
q(FDR) < 0.050
-8.00
-3.38
8.00
3.38
t(945)
p < 0.000742
Time point 2
p(Bonf) < 1.000


B – group of 6 birds
Time point 1
Time point 2
NOV-Rmtf
NOV-R
R-TUT
NOV
TONE


C – change of magnitude
t1 (0 hr)
t2 (1 hr)
1.5
1
0.5
0
NOV-Rmtf
NOV-R
R-TUT
NOV
TONE


D – change of volume
t1 (0 hr)
t2 (1 hr)
4
3
2
1
0
NOV-Rmtf
NOV-R
R-TUT
NOV
TONE


E – Individual volume changes
NOV-Rmtf
NOV-R
R-TUT
NOV

area of songbird neurobiology. *N*-methyl D-aspartate receptor (NMDA)-dependent long-term potentiation (LTP) of the recurrent collaterals due to conjunction of pre- and postsynaptic action potential activity has been detected in the song nucleus, lateral magnocellular nucleus of the anterior nidopallium (LMAN), during song development. During the same period, the thalamic afferents to this structure have been shown to undergo long-term depression (LTD) as a result of out-of-phase pre- and postsynaptic activity [70]. LTP dependent on activation of NMDA and D1-like dopamine receptors also occurs in area X in adult zebra finches and juveniles older than 47 days [71]. The song control nucleus, robust nucleus of arcopallium (RA), shows developmentally restricted and androgen-regulated LTD that is calcium and NMDA receptor mediated and is reversed by high-frequency stimulation [72].

In our own laboratory, using the same type of in vitro brain slice approach as in the above studies, we have explored the modulation of synaptic plasticity in RA by the neurotransmitter/neuromodulator acetylcholine (ACh), primarily through its action on the neuronal nicotinic receptors (nAChRs). Evidence for the involvement of cholinergic mechanisms in the song control system has been established by a large body of prior literature [73–81]. In terms of physiology, cholinergic basal forebrain has been shown to regulate auditory feedback in the song system [82].

LTP Dependence on Activation of nACHRs

We observed that bath application of nicotine to adult zebra finch brain slices produces significant effects on long-term synaptic plasticity in RA [83]. In these experiments, population excitatory postsynaptic potentials (EPSPs) were recorded extracellularly in RA upon stimulation of afferent fibers from LMAN. Intracellular and patch clamp recordings have shown that LMAN-RA connections are glutamatergic in nature [84–86]. Accordingly, we find that glutamate receptor blockers, APV (2-amino-5-phosphonopentanoic acid) and CNQX (6-cyano-7-nitroquinoxaline-2, 3-dione), reduce the amplitude of the population EPSP in RA, but do not completely eliminate it. Tetanic stimulation (20 pulses at 100 Hz)

Fig. 7.5 Plasticity of the BOLD response to repeater song. (**a**) Representative SPMs of a single bird stimulated with stimulus NOV-R, an unfamiliar repeater song, at an initial time point (*upper row*) and 1 h later (*lower row*). The activation cluster is defined by a false-discovery-rate adjusted significance threshold of $p<0.05$, and colors represent *t*-values as indicated by the scales on the right. (**b**) Fixed-effects group analysis including all six birds, for each of the two time points and the five stimuli as indicated at the *bottom*. (**c**) Random effects analysis of the averaged magnitude of the BOLD response, as described by *t*-values in the general linear model, in the individual clusters with $p<0.05$. Significant differences between the two time points (pairwise *t*-test, $p<0.05$ two-tailed) are marked with an *asterisk*. Mean and standard error of the mean are shown. The units are normalized. (**d**) Same for the volume of activation, defined by the volume of individual clusters with $p<0.05$. (**e**) Individual changes in normalized volume over time for each bird separately. Same *color lines* in all *plots* denote same bird

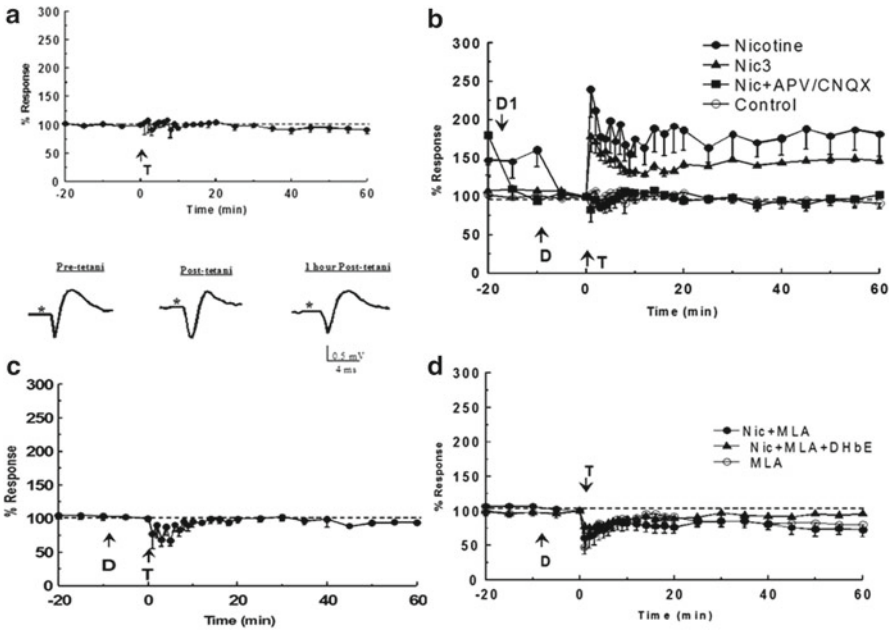


Fig. 7.6 The effects of tetanic stimulation on the extracellular response in zebra finch RA. (**a**) Baseline recordings were conducted for at least 20 min prior to applying a tetanic stimulus (T). The *bottom* panel shows representative traces from a single experiment of evoked extracellular potentials immediately prior to the tetanic stimulation, immediately after the tetanic stimulation, and 1 h after the tetanic stimulation. The peak amplitude of the response is plotted as the percent of the pre-tetanic response immediately preceding the tetani (0 min). (**b**) The effects of a tetanic stimulation on the extracellular response in zebra finch RA in the presence of nicotine (10 μM) or the combination of nicotine and APV/CNQX. In the data sets "Nicotine," "Nic3," and "Nic + APV/CNQX," "D" indicates where nicotine was added to the superfusate, in the absence (Nicotine, Nic3) or presence (Nic + APV/CNQX) of APV/CNQX. "D1" applies only to the data set "Nic + APV/CNQX" and indicates where APV/CNQX was added to the superfusate. "Nic3" represents a subset of experiments from the "Nicotine" group that have been plotted separately to demonstrate that even in the absence of a nicotine-mediated depression of basal response, LTP is elicited. "Control" is the same data from (**a**), replotted for comparison. "T" indicates tetanic stimulation. (**c**) Inhibition of the nicotine-induced long-term potentiation with a nAChR antagonist. In these experiments, a combination (D) of nicotine and nAChR antagonist DH**β**E (1 μM) was added to the superfusate 10 min before the tetanus (T). (**d**) The effects of a tetanic stimulation on the extracellular response in zebra finch RA in the presence of another nAChR antagonist, MLA (10 nM). The combination of nicotine and MLA (Nic + MLA), nicotine, MLA and DHβE (Nic + MLA + DHβE), or MLA alone (MLA) was added (D) to the superfusate 10 min before tetanic stimulation (T). *With permission from Salgado-Commissariat D, Rosenfield DB, Helekar SA. Nicotine-mediated plasticity in robust nucleus of the archistriatum of the adult zebra finch. Brain Res 2004 Aug 20;1018(1):97–105 (84)*

known to produce LTP in rodent hippocampal slices [87, 88] did not potentiate this response (Fig. 7.6a). In the presence of superfused nicotine (10 μM), however, the same type of tetanic stimulation induced LTP with statistically significant 82 % (±28.8 %) increase in the response up to 1 h after stimulation (Fig. 7.6b). LTP

induced in the presence of nicotine reflects the LTP of glutamate receptor-mediated synaptic response and possibly also the involvement of these receptors in induction of LTP because it is not seen when glutamate receptor blockers are added to the superfusate (Nic + APV/CNQX, Fig. 7.6b).

The nAChR antagonist dihydro-β-erythroidine (DHβE), which is a more potent blocker of receptors containing the α4-subunit (Fig. 7.6c), blocks the induction of LTP when co-applied with nicotine. While 1 μM DHβE in our experiments should predominantly block α4-subunit-containing nAChRs, its blockade of α3-subunit-containing nAChRs cannot be ruled out because at this concentration it reduces α3-nAChR-mediated current by 30 % in rodent hippocampal neurons [89]. Methyllycaconitine (MLA, 10 nM), known to more effectively block α7-subunit-containing nAChRs [89], also blocks the nicotinic receptor-dependent LTP (Fig. 7.6d). Moreover, in the presence of MLA, a LTD appears to be unmasked. Taken together these results point to the presence in the song nucleus RA of a bidirectional plasticity that is under the modulatory control of two or more types of nAChRs.

Apart from the effects on synaptic plasticity, intracellular recordings in single neurons within RA indicate that nAChR activation affects the excitability of neurons. Nicotine increases the number of action potentials induced by a depolarizing stimulus. There is also a 58 % increase in the frequency of spontaneous action potentials. The amplitude of the after hyperpolarization is significantly reduced in the presence of nicotine. LTP induction in the presence of nicotine could be partially accounted for by this overall increase in neuronal excitability in response to nicotine, but in addition it might also be mediated by calcium influx through pre- or postsynaptic α7 nAChRs.

Effects of Nicotinic Receptor Modulation on Song Learning

As far as effects of nicotine in vivo are concerned, the most significant results were obtained in the early sensory phase of song learning. We performed these experiments in 15 male birds. We injected nicotine (1 mg/kg body weight intramuscularly) in juvenile birds. Twenty-day-old birds were divided into four groups: two nicotine-treated and two matched control saline-treated groups. In both nicotine and saline groups, daily injections were carried out from 20 through 35 days post-hatch, after an initial 19-day isolation of birds with females. In one test/control group pair song tutoring was done from 20 through 35 days post-hatch, followed by song isolation until 100 days post-hatch. In the other test/control group pair tutoring was done from 20 through 65 days post-hatch, followed by song isolation until 100 days post-hatch. Birds were tutored in each case by a normal adult tutor in an adjacent cage.

The results of these experiments are as follows. In the 20–35-day-tutored nicotine group, two of four birds developed songs containing syllable repetitions with repeater motif frequency of 80 and 35 %. Figure 7.7 shows examples of their song motifs. All three of the matched control saline-treated birds produced normal

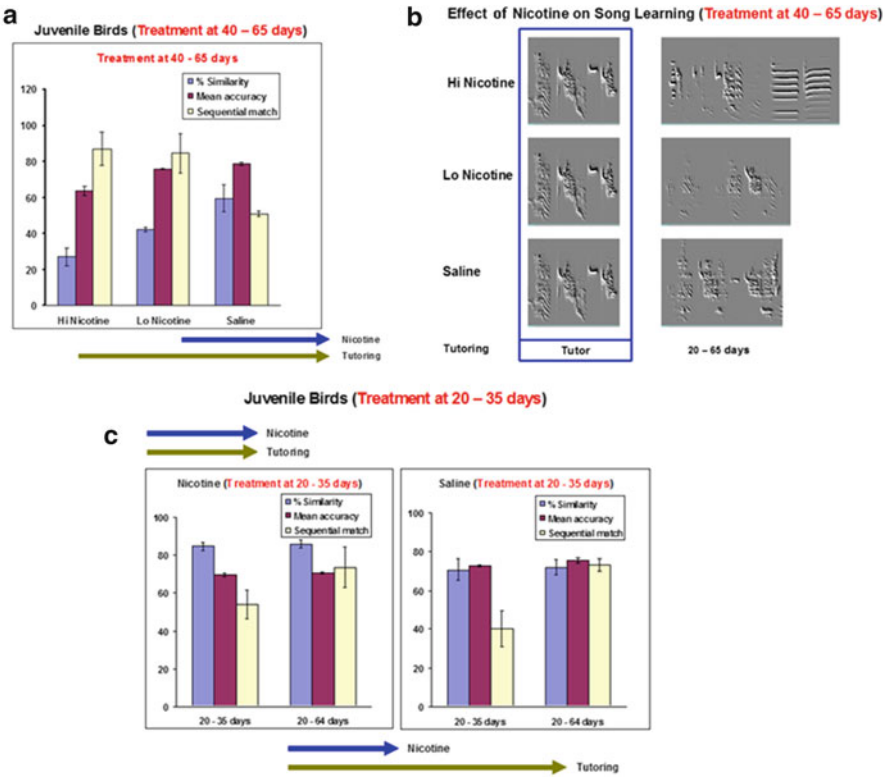


Fig. 7.7 The effect of nicotine treatment on song learning. (**a**) Juvenile male birds, approximately 40 days old, were given daily injections of either nicotine (0.1 mg/kg—Lo Nicotine, or 1 mg/kg—Hi Nicotine) or vehicle (control group) for 25 days. Experimental birds were housed in adjacent cages to the tutor. All injections were administered 30 min before the lights were turned on. Bar graph shows dose-dependent reduction in percent similarity of pupil songs to the tutor's song. (**b**) Spectrograms show motifs of pupils recorded at post-hatch day 100 in comparison with the tutor motif. The percent similarity, mean accuracy, and sequential match values were obtained using Sound Analysis 3. (**c**) Treatment with nicotine from 20 to 35 days post-hatch causes the development of variant song motifs with syllable repetitions. Bar graphs show a small increase in percent similarity of pupil song motifs to the tutor motif due to treatment with nicotine compared to saline control, irrespective of whether tutoring was stopped at 35 days or carried on until 64 days

non-repeater songs. There was no significant difference between the nicotine and saline groups in the percent similarity between tutor and pupil song motifs (nicotine 59.7 ± 3 %, $n = 4$; saline 45.4 ± 2.7 %, $n = 3$), indicating that nicotine did not produce any impairment in song imitation per se during this early phase of learning. In the 20–65-day-tutored groups, none of the birds, i.e., neither the nicotine-treated nor the saline-treated group, produced songs with syllable repetitions. There was also no significant nicotine-induced change in song imitation in the nicotine group (percent similarity 58.7 ± 2.8 %, $n = 4$) compared to the matched saline group

(percent similarity 61.1 ± 2.4 %, $n = 4$). While these preliminary findings need to be replicated in a larger set of birds, they suggest that brain nAChR activation or desensitization during the sensory phase of song learning can induce zebra finches to become syllable repeaters in spite of being tutored with a normal non-repeater song. Continued tutoring for additional 30 days (36–65 days) after cessation of nicotine treatment prevents the development of repeater song, suggesting further that restorative changes during the sensorimotor phase of learning might reverse or correct nAChR-dependent alterations in the sensory template or learning mechanisms. However, as stated earlier, treatment with nicotine (1 mg/kg) during tutoring at 40–65 days does not lead to the production of repeater song. This finding suggests that the nAChR-sensitive mechanism critical to induction of repetitive syllable output within song motifs operates during the early sensory phase of song learning, and not during the late sensorimotor phase.

Treatment with nicotine in adults (>120 days post-hatch) in a manner similar to that in juveniles produced a modest but significant reduction in percent similarity between pre- and posttreatment song with high dose of nicotine compared to control birds (data not shown). Such an alteration was not seen with the low dose of nicotine. Again, there does not seem to be a significant alteration in the composition of syllables in the song motif. From the results of the higher dose of nicotine, we infer that excessive nAChR activation or desensitization might interfere with the mechanism underlying the maintenance of the integrity of song in adults. Because auditory feedback is critical to this process, it is possible that the effect of nicotine is on cochlear nAChRs leading to the impairment of hearing itself.

Conclusions

The parallels between birdsong learning in songbirds and acquisition of speech in humans offer us the opportunity to treat aberrations of birdsong as simple models of dysfluencies such as stuttering and to study their neurobiological underpinnings. In this review, we have attempted to demonstrate that the songbird zebra finch allows us to investigate, at multiple levels, a possible minimal model of stuttering involving involuntary abnormal repetition of song syllables, in order to address causal mechanisms. At the cellular level, we can examine synaptic plasticity, and its neuromodulation, and lay the groundwork for future pharmacotherapeutic approaches against speech problems. At the behavioral level, we can manipulate the learning environment during development to induce changes in song and assess innate tendencies and vulnerabilities that may be relevant from the translational standpoint. At the systems level with fMRI, we can compare and contrast brain activation under relatively similar paradigms of speech and song perception, using a common experimental platform. All these approaches, combined with the genetic information that is now available through the zebra finch genome initiative, would in the near future facilitate a deeper understanding of the neurodevelopmental basis of stuttering and other speech motor control disorders.

Acknowledgments This work was funded by NIH grants DC04778-01A1 (SAH) and MH073900-01 (SAH), NSF grants IOS 0956306 (HUV) and IOS 1065678 (SAH), grants from M. R. Bauer Foundation (DBR) and Lowin Medical Research Foundation (DBR), and Weill Cornell Medical College—The Methodist Hospital Research Institute collaboration grants (HUV and SAH).

References

1. Porfert A, Rosenfield D (1978) Prevalence of stuttering. J Neurol Neurosurg Psychiatry 41(10):954–956
2. Bloodstein O (1995) A handbook on stuttering. Singular, San Diego
3. Nudelman HB, Herberich KE, Hoyt BD, Rosenfield DB (1989) A neuroscience model of stuttering. J Fluency Disord 14:399–427
4. Nudelman H, Herbrich K, Hess K, Hoyt B, Rosenfield D (1992) A model of the phonatory response time of stutterers and fluent speakers to frequency-modulated tones. J Acoust Soc Am 92(4 Pt 1):1882–1888
5. Rosenfield DB (2013) Disorders of fluency and voice. In: Whitaker H (ed) International encyclopedia of the social and behavioral sciences, 2nd edn. American Elsevier, New York
6. Timmons B (1982) Physiological factors related to delayed auditory feedback and stuttering: a review. Percept Mot Skills 55(3 Pt 2):1179–1189
7. Leonardo A, Konishi M (1999) Decrystallization of adult birdsong by perturbation of auditory feedback. Nature 399(6735):466–470
8. Immelmann K (1969) Song development in the zebra finch and other estrildid finches. In: Hinde RA (ed) Bird vocalizations. Cambridge University Press, Cambridge, pp 61–77
9. Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. Annu Rev Neurosci 22:567–631
10. Sossinka R, Bohner J (1980) Song types in the zebra finch *Poephilia guttata castanotis*. Z Tierpsychol 53:123–132
11. Sturdy CB, Phillmore LS, Weisman RG (1999) Note types, harmonic structure, and note order in the songs of zebra finches (Taeniopygia guttata). J Comp Psychol 113:194–203
12. Helekar SA, Marsh S, Viswanath NS, Rosenfield DB (2000) Acoustic pattern variations in the female-directed birdsongs of a colony of laboratory-bred zebra finches. Behav Processes 49(2):99–110
13. Tchernichovski O, Lints T, Mitra PP, Nottebohm F (1999) Vocal imitation in zebra finches is inversely related to model abundance. Proc Natl Acad Sci U S A 96(22):12901–12904
14. Cooper BG, Goller F (2004) Partial muting leads to age-dependent modification of motor patterns underlying crystallized zebra finch song. J Neurobiol 61(3):317–332
15. Horita H, Wada K, Jarvis ED (2008) Early onset of deafening-induced song deterioration and differential requirements of the pallial-basal ganglia vocal pathway. Eur J Neurosci 28(12): 2519–2532
16. Helekar SA, Espino GG, Botas A, Rosenfield DB (2003) Development and adult phase plasticity of syllable repetitions in the birdsong of captive zebra finches (Taeniopygia guttata). Behav Neurosci 117(5):939–951
17. Tchernichovski O, Nottebohm F, Ho CE, Pesaran B, Mitra PP (2000) A procedure for an automated measurement of song similarity. Anim Behav 59(6):1167–1176
18. Brainard M, Doupe A (2000) Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. Nature 404(6779):762–766
19. Huettel SA, Song AW, McCarthy G (2008) Functional magnetic resonance imaging, 2nd edn. Sinauer, Sunderland
20. Voss HU, Tabelow K, Polzehl J, Tchernichovski O, Maul KK, Salgado-Commissariat D et al (2007) Functional MRI of the zebra finch brain during song stimulation suggests a lateralized response topography. Proc Natl Acad Sci U S A 104(25):10667–10672

21. Van Meir V, Boumans T, De Groof G, Verhoye M, Van Audekerke J, Van der Linden A (2003) Functional magnetic resonance imaging of the songbird brain when listening to songs. In: 2003 Neuroscience Meeting Planner Online ed. Society for Neuroscience, Washington, DC
22. Van Meir V, Boumans T, De Groof G, Van Audekerke J, Smolders A, Scheunders P et al (2005) Spatiotemporal properties of the BOLD response in the songbirds' auditory circuit during a variety of listening tasks. Neuroimage 25(4):1242–1255
23. Lipkind D, Tchernichovski O (2011) Quantification of developmental birdsong learning from the subsyllabic scale to cultural evolution. Proc Natl Acad Sci U S A 108(suppl 3):15572–15579
24. Maul KK, Voss HU, Parra LC, Salgado-Commissariat D, Ballon D, Tchernichovski O et al (2010) The development of stimulus-specific auditory responses requires song exposure in male but not female zebra finches. Dev Neurobiol 70(1):28–40
25. Boumans T, Vignal C, Ramstein S, Verhoye M, Van Audekerke J, Mottin S et al (2005) Detection of haemodynamic changes in zebra finch brain by optical and functional magnetic resonance imaging. J Cereb Blood Flow Metab 25:S388
26. Vignal C, Boumans T, Montcel B, Ramstein S, Verhoye M, Van Audekerke J et al (2008) Measuring brain hemodynamic changes in a songbird: responses to hypercapnia measured with functional MRI and near-infrared spectroscopy. Phys Med Biol 53(10):2457–2470
27. Keary N, Voss J, Lehmann K, Bischof HJ, Lowel S (2010) Optical imaging of retinotopic maps in a small songbird, the zebra finch. PLoS One 5(8):e11912
28. Mottin S, Montcel B, de Chatellus HG, Ramstein S (2011) Functional white-laser imaging to study brain oxygen uncoupling/recoupling in songbirds. J Cereb Blood Flow Metab 31(2):393–400
29. Abi-Haidar D, Olivier T (2009) Confocal reflectance and two-photon microscopy studies of a songbird skull for preparation of transcranial imaging. J Biomed Opt 14(3):034038
30. Ramstein S, Vignal C, Mathevon N, Mottin S (2005) In vivo and noninvasive measurement of a songbird head's optical properties. Appl Optics 44(29):6197–6204
31. Boumans T, Vignal C, Smolders A, Sijbers J, Verhoye M, Van Audekerke J et al (2008) Functional magnetic resonance imaging in zebra finch discerns the neural substrate involved in segregation of conspecific song from background noise. J Neurophysiol 99(2):931–938
32. Voss HU, Salgado-Commissariat D, Helekar SA (2010) Altered auditory BOLD response to conspecific birdsong in zebra finches with stuttered syllables. PLoS One 5(12):e14415
33. Boumans T, Gobes SM, Poirier C, Theunissen FE, Vandersmissen L, Pintjens W et al (2008) Functional MRI of auditory responses in the zebra finch forebrain reveals a hierarchical organisation based on signal strength but not selectivity. PLoS One 3(9):e3184
34. Boumans T, Theunissen FE, Poirier C, Van der Linden A (2007) Neural representation of spectral and temporal features of song in the auditory forebrain of zebra finches as revealed by functional MRI. Eur J Neurosci 26(9):2613–2626
35. Poirier C, Boumans T, Verhoye M, Balthazart J, Van der Linden A (2009) Own-song recognition in the songbird auditory pathway: selectivity and lateralization. J Neurosci 29(7): 2252–2258
36. Tindemans I, Verhoye M, Balthazart J, Van Der Linden A (2003) In vivo dynamic ME-MRI reveals differential functional responses of RA- and area X-projecting neurons in the HVC of canaries exposed to conspecific song. Eur J Neurosci 18(12):3352–3360
37. Van der Linden A, Verhoye M, Van Meir V, Tindemans I, Eens M, Absil P et al (2002) In vivo manganese-enhanced magnetic resonance imaging reveals connections and functional properties of the songbird vocal control system. Neuroscience 112(2):467–474
38. Poirier C, Boumans T, Vellema M, De Groof G, Charlier TD, Verhoye M et al (2011) Own song selectivity in the songbird auditory pathway: suppression by norepinephrine. PLoS One 6(5):e20131
39. De Nil LF, Kroll RM, Lafaille SJ, Houle S (2003) A positron emission tomography study of short- and long-term treatment effects on functional brain activation in adults who stutter. J Fluency Disord 28(4):357–379; quiz 79–80

40. Neumann K, Euler HA, von Gudenberg AW, Giraud AL, Lanfermann H, Gall V et al (2003) The nature and treatment of stuttering as revealed by fMRI A within- and between-group comparison. J Fluency Disord 28(4):381–409; quiz 409–410
41. Cynx J, Williams H, Nottebohm F (1992) Hemispheric differences in avian song discrimination. Proc Natl Acad Sci U S A 89(4):1372–1375
42. George I, Vernier B, Richard JP, Hausberger M, Cousillas H (2004) Hemispheric specialization in the primary auditory area of awake and anesthetized starlings (Sturnus vulgaris). Behav Neurosci 118(3):597–610
43. George I, Cousillas H, Richard JP, Hausberger M (2005) State-dependent hemispheric specialization in the songbird brain. J Comp Neurol 488(1):48–60
44. Phan ML, Vicario DS (2010) Hemispheric differences in processing of vocalizations depend on early experience. Proc Natl Acad Sci U S A 107(5):2301–2306
45. 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
46. Espino GG, Lewis C, Rosenfield DB, Helekar SA (2003) Modulation of theta/alpha frequency profiles of slow auditory-evoked responses in the songbird zebra finch. Neuroscience 122(2):521–529
47. Bolhuis JJ, Zijlstra GG, den Boer-Visser AM, Van Der Zee EA (2000) Localized neuronal activation in the zebra finch brain is related to the strength of song learning. Proc Natl Acad Sci U S A 97(5):2282–2285
48. Bolhuis JJ, Hetebrij E, Den Boer-Visser AM, De Groot JH, Zijlstra GG (2001) Localized immediate early gene expression related to the strength of song learning in socially reared zebra finches. Eur J Neurosci 13(11):2165–2170
49. 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 U S A 103(4):1088–1093
50. Pinaud R, Terleph TA, Tremere LA, Phan ML, Dagostin AA, Leao RM et al (2008) Inhibitory network interactions shape the auditory processing of natural communication signals in the songbird auditory forebrain. J Neurophysiol 100(1):441–455
51. Northoff G, Walter M, Schulte RF, Beck J, Dydak U, Henning A et al (2007) GABA concentrations in the human anterior cingulate cortex predict negative BOLD responses in fMRI. Nat Neurosci 10(12):1515–1517
52. Logothetis NK (2003) The underpinnings of the BOLD functional magnetic resonance imaging signal. J Neurosci 23(10):3963–3971
53. Lee JH, Durand R, Gradinaru V, Zhang F, Goshen I, Kim DS et al (2010) Global and local fMRI signals driven by neurons defined optogenetically by type and wiring. Nature 465(7299):788–792
54. Liotti M, Ingham JC, Takai O, Paskos DK, Perez R, Ingham RJ (2010) Spatiotemporal dynamics of speech sound perception in chronic developmental stuttering. Brain Lang 115(2): 141–147
55. De Nil LF, Beal DS, Lafaille SJ, Kroll RM, Crawley AP, Gracco VL (2008) The effects of simulated stuttering and prolonged speech on the neural activation patterns of stuttering and nonstuttering adults. Brain Lang 107(2):114–123
56. Watkins KE, Smith SM, Davis S, Howell P (2008) Structural and functional abnormalities of the motor system in developmental stuttering. Brain 131(Pt 1):50–59
57. Chang SE, Kenney MK, Loucks TM, Ludlow CL (2009) Brain activation abnormalities during speech and non-speech in stuttering speakers. Neuroimage 46(1):201–212
58. Brown S, Ingham RJ, Ingham JC, Laird AR, Fox PT (2005) Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. Hum Brain Mapp 25(1):105–117
59. Fox PT, Ingham RJ, Ingham JC, Zamarripa F, Xiong JH, Lancaster JL (2000) Brain correlates of stuttering and syllable production. A PET performance-correlation analysis. Brain 123(Pt 10):1985–2004

60. Braun AR, Varga M, Stager S, Schulz G, Selbie S, Maisog JM et al (1997) Altered patterns of cerebral activity during speech and language production in developmental stuttering. An H2(15)O positron emission tomography study. Brain 120(Pt 5):761–784
61. De Nil LF, Kroll RM, Kapur S, Houle S (2000) A positron emission tomography study of silent and oral single word reading in stuttering and nonstuttering adults. J Speech Lang Hear Res 43(4):1038–1053
62. Fox PT, Ingham RJ, Ingham JC, Hirsch TB, Downs JH, Martin C et al (1996) A PET study of the neural systems of stuttering. Nature 382(6587):158–161
63. Neumann K, Preibisch C, Euler HA, von Gudenberg AW, Lanfermann H, Gall V et al (2005) Cortical plasticity associated with stuttering therapy. J Fluency Disord 30(1):23–39
64. Voss HU, Salgado-Commissariat D, Ballon D, Helekar SA (2007). Functional neuroimaging in the songbird zebra finch reveals plasticity of the auditory response based on song familiarity. In: IBRO World Congress of Neuroscience, Melbourne, Australia 2007
65. Goebel HH, Wisniewski KE (2004) Current state of clinical and morphological features in human NCL. Brain Pathol 14(1):61–69
66. Poirier C, Vellema M, Verhoye M, Van Meir V, Wild JM, Balthazart J et al (2008) A three-dimensional MRI atlas of the zebra finch brain in stereotaxic coordinates. Neuroimage 41(1):1–6
67. Friston KJ, Holmes AP, Worsley K, Poline JB, Frith CD, Frackowiak RSJ (1995) Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp 2:189–210
68. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate—a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 57:289–300
69. Genovese CR, Lazar NA, Nichols T (2002) Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 15(4):870–878
70. Boettiger CA, Doupe AJ (2001) Developmentally restricted synaptic plasticity in a songbird nucleus required for song learning. Neuron 31(5):809–818
71. Ding L, Perkel DJ (2004) Long-term potentiation in an avian basal ganglia nucleus essential for vocal learning. J Neurosci 24(2):488–494
72. Sizemore M, Perkel DJ (2011) Premotor synaptic plasticity limited to the critical period for song learning. Proc Natl Acad Sci U S A 108(42):17492–17497
73. Ryan S, Arnold A (1981) Evidence for cholinergic participation in the control of bird song; acetylcholinesterase distribution and muscarinic receptor autoradiography in the zebra finch brain. J Comp Neurol 202(2):211–219
74. Zuschratter W, Scheich H (1990) Distribution of choline acetyltransferase and acetylcholinesterase in the vocal motor system of zebra finches. Brain Res 513(2):193–201
75. Li R, Zuo M, Sakaguchi H (1999) Auditory-vocal cholinergic pathway in zebra finch brain. Neuroreport 10(1):165–169
76. 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(2):239–246
77. Sakaguchi H, Saito N (1991) Developmental change of cholinergic activity in the forebrain of the zebra finch during song learning. Brain Res Dev Brain Res 62(2):223–228
78. Sakaguchi H, Saito N (1989) The acetylcholine and catecholamine contents in song control nuclei of zebra finch during song ontogeny. Brain Res Dev Brain Res 47(2):313–317
79. Watson J, Adkins RE, Whiting P, Lindstrom J, Podleski T (1988) Autoradiographic localization of nicotinic acetylcholine receptors in the brain of the zebra finch (Poephila guttata). J Comp Neurol 274(2):255–264
80. Sakaguchi H, Li R, Taniguchi I (2000) Sex differences in the ventral paleostriatum of the zebra finch: origin of the cholinergic innervation of the song control nuclei. Neuroreport 11(12): 2727–2731
81. Reiner A, Laverghetta AV, Meade CA, Cuthbertson SL, Bottjer SW (2004) An immunohistochemical and pathway tracing study of the striatopallidal organization of area X in the male zebra finch. J Comp Neurol 469(2):239–261

82. Shea SD, Margoliash D (2003) Basal forebrain cholinergic modulation of auditory activity in the zebra finch song system. Neuron 40(6):1213–1226
83. Salgado-Commissariat D, Rosenfield DB, Helekar SA (2004) Nicotine-mediated plasticity in robust nucleus of the archistriatum of the adult zebra finch. Brain Res 1018(1):97–105
84. Mooney R (1992) Synaptic basis for developmental plasticity in a birdsong nucleus. J Neurosci 12(7):2464–2477
85. 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
86. 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(20):9107–9116
87. Fujii S, Ji Z, Morita N, Sumikawa K (1999) Acute and chronic nicotine exposure differentially facilitate the induction of LTP. Brain Res 846(1):137–143
88. Fujii S, Sumikawa K (2001) Nicotine accelerates reversal of long-term potentiation and enhances long-term depression in the rat hippocampal CA1 region. Brain Res 894(2): 340–346
89. Alkondon M, Albuquerque EX (1993) Diversity of nicotinic acetylcholine receptors in rat hippocampal neurons. I. Pharmacological and functional evidence for distinct structural subtypes. J Pharmacol Exp Ther 265(3):1455–1473

Part III
Mammalian Models of Vocal Communication

Chapter 8
The Repertoire of Communication Calls Emitted by Bats and the Ways the Calls Are Processed in the Inferior Colliculus

George D. Pollak, Sari Andoni, Kirsten Bohn, and Joshua X. Gittelman

Abstract Bats have among the richest and most sophisticated repertoire of vocal communication calls of any mammalian group. In this review, we first describe the range of calls bats emit and the acoustic features that comprise their calls. Of particular importance are frequency modulations (FMs), as these are components in the vast majority of bats' communication calls as well as the calls they emit for echolocation. We then consider the processing of communication calls in the inferior colliculus (IC). We show that neurons in the IC are selective for the various calls the bats emit and that this selectivity is shaped by inhibition. Computational studies showed that some neurons had one feature or filter characterized by its spectrotemporal receptive field (STRF) generated by spike-triggered averaging. In these cells, convolving conspecific calls with the STRF provides an accurate prediction of their responses to conspecific calls. Moreover a single linear combination of the excitatory and inhibitory fields explains their responses to the direction and velocity of FM sweeps. Most IC cells, however, had several spectrotemporal filters. In these cells, the nonlinear combination of two or more filters predicted the cell's selectivity for FM sweeps and its responses to calls. The ways in which excitation and inhibition interacted to generate FM selectivity were also evaluated with in vivo whole-cell recordings. Those studies showed that the relative timing of excitation and inhibition had only a small influence on the amplitudes of the excitatory postsynaptic potentials (EPSPs) evoked by an FM signal. How the change in EPSP amplitude influenced discharge probability depended in large part on how close the EPSP was

G.D. Pollak, Ph.D. (✉) • S. Andoni, Ph.D.
Section of Neurobiology, Center for Perceptual Systems, The University of Texas at Austin, Austin, TX, USA
e-mail: gpollak@mail.utexas.edu

K. Bohn, M.S., Ph.D.
Florida International University, Miami, FL, USA

J.X. Gittelman, Ph.D.
Department of Science, Washington State University, Vancouver, WA, USA

S.A. Helekar (ed.), *Animal Models of Speech and Language Disorders*,
DOI 10.1007/978-1-4614-8400-4_8,

to spike threshold. If the EPSP amplitude is far from threshold, even timing changes of several ms would have little or no effect on spike probability. Conversely, if the EPSP amplitude is near threshold, then even a change in EPSP amplitude as small as a fraction of a millivolt could affect discharge probability and thus modulate the cell's spiking directional selectivity. Taken together, these studies showed that neurons in the auditory midbrain encode specific spectrotemporal features of natural communication sounds by means of their selectivity to FM features present in their conspecific calls.

Keywords Echolocating bats • Communication calls • Inhibition • Spectrotemporal receptive fields • In vivo whole-cell recordings • Spike timing

Introduction

Natural sounds, such as conspecific vocalizations and human speech, are vital for social communication, foraging, mating, and therefore survival. Bats, perhaps more than any other mammal, depend on their hearing for survival. Not only do they rely on hearing for orientation and hunting through echolocation, but hearing is also critically important for social communication. Many bats live in large colonies where they engage in a myriad of social interactions, which are accomplished largely with sound since they live in dark environments where visual displays are of no use [1]. Their communication signals can be quite elaborate [2, 3], and some species are capable of vocal learning [4–7]. Indeed, the repertoire of signals bats use for vocal communication are remarkably rich and sophisticated [8–10].

Vocal communication was presumably used by their ancestors before bats took to the night sky to exploit a food supply for which there was little competition. Thus, we view the processing of communication signals as one of the primary tasks for which their auditory systems were designed, and adaptations required for echolocation were subsequently added to enable the various species of bats to compete successfully for food resources in a wide range of different habitats. Two noteworthy features are consistent with this idea. The first is that the acoustic features of echolocation calls, which are largely composed of brief, downward sweeping frequency modulations (FMs), are remarkably similar to the FMs in many of the communication calls [11]. The second feature is that the auditory systems of bats are similar to the auditory systems of all other mammals, with the same structures, wiring, and mechanisms for processing information that are possessed by all other mammals [12–16]. What distinguishes the auditory system of bats are not novel mechanisms, but rather that some common mechanisms and features are far more pronounced in their auditory systems than in other mammals.

In the following sections, we first present some of the communication signals bats emit together with the behaviors in which the signals are emitted to illustrate the richness and variety of their vocal repertoires. The subsequent sections then deal with the neural processing of communication calls in the inferior colliculus (IC),

the midbrain auditory nucleus. We emphasize the IC because it receives the convergent projections from almost all of the lower auditory nuclei and their interaction in the IC produces an output that synthesizes the convergent inputs [14, 17, 18]. The net result of those syntheses is that a variety of new response properties are either formed de novo in the IC or response properties that have been formed in lower nuclei are sharpened or further modified in the IC [19–22].

The sections on neural processing have three themes. The first theme is that IC cells are tuned to respond to the direction and sweep velocity of the various FMs present in the bats' conspecific communication calls. Additionally, the tuning for FM selectivity is shaped by the interactions of the excitatory and inhibitory innervation that plays upon IC neurons. The second theme is that the IC population is heterogeneous, where some cells form their selective response properties with linear processing, whereas others form similar response properties through nonlinear processing. Following from the above, the third theme is that the various computations employed by the IC endow these neurons with selectivities for features of FM sweeps, and those selectivities, in turn, largely create the response selectivities for the various conspecific communication calls these animals hear in their daily lives.

The Vocal Repertoire of Bats

The variety and complexity of communication calls bats use are well illustrated by Mexican free-tailed bats (*Tadarida brasiliensis*), members of the family Molossidae. These bats are common in the Southwestern United States where they live in caves with populations that often number in the millions. Here males use vocal signals to establish dominance hierarchies, maintain territories, garner females into harems, and defend their harems against intruding males, whereas females use vocal signals for recognition of and bonding with their pups among other behaviors [3, 11, 23–26].

To give a flavor of the variety of calls emitted by these animals, a sample is shown in Fig. 8.1, together with the behaviors the bats displayed during the emission of each call type. Each call is composed of one or more repetitions of a syllable or note. A syllable is an individual or discrete acoustic element, where each syllable is composed of multiple harmonics with spectral components that change in amplitude and often in frequency throughout its duration. For example, there are five syllables (discrete acoustic elements) shown for the directive call in Fig. 8.1e. Syllables range not only in duration (from 2–3 ms to over 100 ms) and but also in their spectral structures. For example, some syllables are simply brief downward sweeping FMs (e.g., the individual syllables in the irritation call in Fig. 8.1f), whereas other are more complex and have both upward and downward FMs (e.g., directive, 8.1e, and herding calls, 8.1a), and yet others have only harmonic stacks of constant frequencies (marking, 8.1b, and mounting calls, 8.1c) or harmonic stacks of very shallow frequency modulations (alarm calls, 8.1k). The temporal sequence in which the syllables are emitted is also an important feature that varies with behavioral context [10]. The syllables produced in several different calls

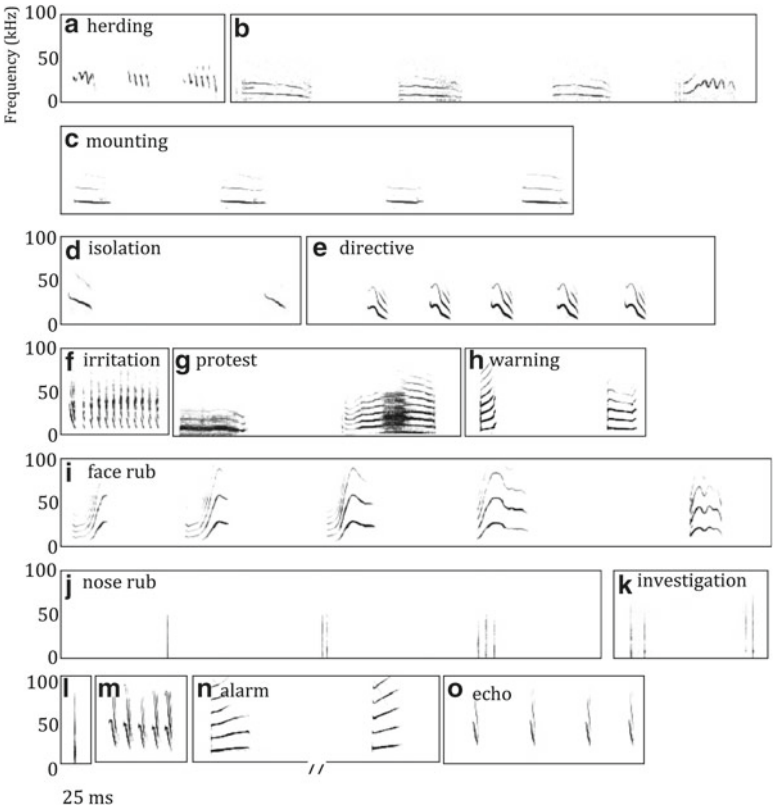


Fig. 8.1 Spectrograms that show the various communication calls emitted by Mexican free-tailed bats. (**a**) *Herding Calls* were emitted while a male forcefully pushed one or more females with his muzzle or wing into his territory. (**b**) *Marking Calls* were emitted by a dominant male while rubbing his face and gular gland on the surface of his territories. (**c**) *Mounting Calls* were emitted by males to convey dominance, when males would mount females and forcefully push their muzzles repeatedly between their shoulders. (**d**) *Isolation Calls* were emitted by pups immediately after birth and throughout development. Pups called when they were isolated or hungry. (**e**) *Directive Calls* were emitted by females while giving birth and throughout pup development when females approached pups or in response to their pups' isolation calls. (**f**) *Irritation Calls* were emitted when bats were jostled by other bats. (**g**) *Protest Calls* were emitted in response to aggressive behaviors by other bats. (**h**) *Warning Calls* were emitted prior to aggressive encounters. (**i**) *Face-Rubbing Calls* were used for social bonding. They were emitted in roost sites while approaching another bat and rubbing their muzzles across the body of the other bat. (**j**) *Food Solicitation Calls* were emitted during or immediately prior to feeding. (**k**) *Alarm Calls* were emitted during periods of high levels of aggression. (**l**) *Echolocation Calls* were emitted for orientation while the bats were flying (adapted from Bohn KM, Schmidt-French B, Schwartz C, Smotherman M, Pollak GD. Versatility and stereotypy of free-tailed bat songs. PLoS One. 2009;4(8):e6746 [3])

associated with completely different contexts are indistinguishable except for differences in temporal intervals or repetition rate.

Bats not only emit the simpler types of calls illustrated in Fig. 8.1, they also sing elaborate songs. During the breeding season, many animals emit simple repetitions

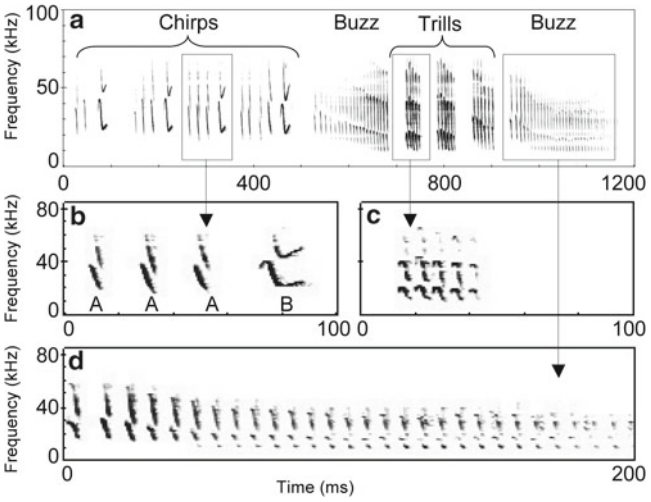


Fig. 8.2 The courtship song of a Mexican free-tailed bat. (**a**) One complete song showing the three types of phrases: chirps, buzzes, and trills. (**b**) Expanded section of a chirp phrase showing the A and B syllables. (**c**) Expanded section of a trill. (**d**) Expanded section of a buzz (adapted from Bohn KM, Schmidt-French B, Schwartz C, Smotherman M, Pollak GD. Versatility and stereotypy of free-tailed bat songs. PLoS One. 2009;4(8):e6746 [3])

of one or a few syllables that are generally referred to as mating or advertisement "calls." In a few exceptional animal groups, such as songbirds [27] and whales [28], these advertisement signals can be more complex vocalizations termed "songs." The major difference between mating "calls" and "songs" is that songs are longer and more complex and contain multiple types of elements (e.g., syllables or notes) that are combined in a stereotypical manner [27, 29]. Therefore, songs have an added dimension of complexity in the form of "syntax"—the patterns by which elements are ordered and combined. Indeed, in most songs, element ordering is not random, but is instead highly structured, with individual, regional, and/or species-specific patterns [30].

A remarkable feature of Mexican free-tails is that the ways in which phrases are combined to form songs follow broad syntactical rules, yet males dynamically vary phrase order from one rendition to the next. During the breeding season, dominant males sing their courtship songs when females approach their territories [3, 11]. Their courtship songs are composed of three types of phrases. A phrase is composed of one or more syllables that form a distinct and reproducible unit, and the phrases are combined to form songs. The three phrases are chirps, trills, and buzzes (Fig. 8.2). Chirps are phrases composed of two types of syllables: "A" and "B" syllables. "A" syllables are short (~5 ms) downward sweeping FMs (Fig. 8.2b). B syllables are longer (~15 ms) and more complex than A syllables. B syllables often begin with an upward FM followed by a longer downward FM, and some signals end with a second upward FM. Thus, their spectral contours often have multiple inflection points. Several A syllables always precede each B syllable, and the sequence of several A syllables followed by a B syllable is then repeated to form the chirp phrase.

The second type of phrase is the trill. Trills are composed of short (3–4 ms), downward FM syllables that are sometimes connected, resulting in sinusoidal patterns (Fig. 8.2c). Trill syllables, whether discrete or connected, are produced as a distinct phrase or burst with durations of approximately 25 ms. Sequential trill phrases are often emitted in songs, but are highly distinctive since each phrase is separated from the next by a silent interval that is much greater than the interval between syllables within each trill phrase.

The third phrase in song is the buzz (Fig. 8.2). Buzzes are also composed of short (3 ms) downward FM syllables that are always separated by a few ms. Although the acoustical structure of trill and buzz syllables are similar, the phrases are distinguished by the number of syllables they contain, where trills have only 3–4 four syllables, whereas buzzes have on average 35 syllables. They are also distinguished by the spectral structure of the syllables. The initial FM syllables in each buzz have relatively high beginning and end frequencies and are followed by 5–10 syllables with progressively lower beginning and end frequencies (Fig. 8.2d).

Given this acoustic complexity and variety of their vocal communication calls, the question naturally arises as to how the auditory system of bats' processes and represent the various calls and songs they emit. As was shown above, the vast majority of calls contain FMs, and as we show in the following sections, IC cells are tuned to respond to the direction and sweep velocity of the various FMs present in these signals. Additionally, the tuning for FM selectivity is shaped by the interactions of the excitatory and inhibitory inputs that play upon IC neurons. It follows, therefore, that the computations employed by the IC endow IC neurons with selectivity for FM features and that selectivity, in turn, determines, in large part, how IC neurons respond to conspecific communication calls.

Responses to Vocal Communication Calls Are Selective

When a series of conspecific communication calls is presented to a group of isofrequency IC neurons (i.e., neurons tuned to the same frequency), most neurons respond to only a subset of the calls and not to other calls; thus each neuron expresses response selectivity. Selectivity of this sort is seen in the IC of all mammals that have been studied [31–34]. Selectivity is illustrated in Fig. 8.3, which shows a suite of ten species-specific communication and echolocation calls and the responses that were evoked from four IC cells in a Mexican free-tailed bat that were all tuned to about the same frequency. Each call had a different and unique spectrotemporal structure (Fig. 8.3), was broadband, and had multiple harmonics. Each was presented at an intensity that was at least 20 dB above the neuron's threshold at the frequency to which the neuron was most sensitive, its best frequency (BF). Each call had suprathreshold energy that encroached upon each neuron's excitatory tuning curve. The differential responses to each of the calls showed that IC cells are not only selective, but their selectivities are diverse, in that the particular subset of calls

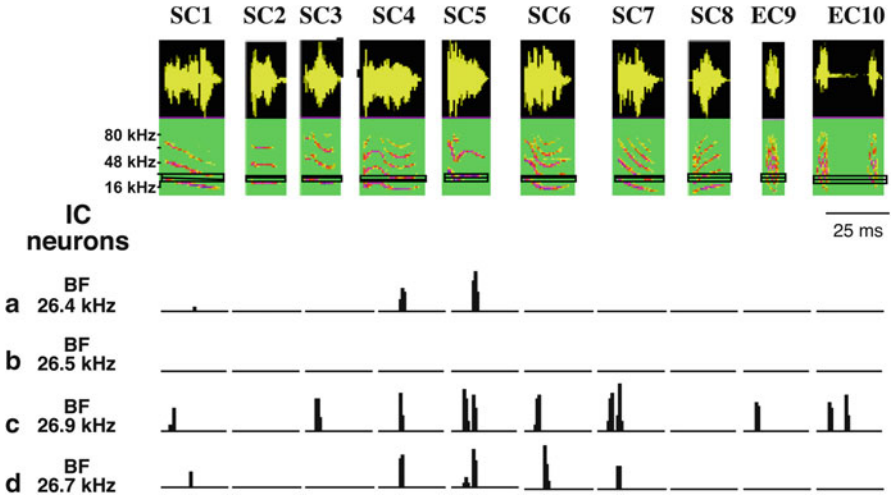


Fig. 8.3 Responses of four IC neurons to ten species-specific calls. Eight of the calls (SC1–SC8) are social communication calls and two others (EC9–EC10) are echolocation calls. The four IC cells are isofrequency and all tuned to about 26 kHz. The IC cells were selective in that each fired to only a subset of the ten calls although each of the calls had suprathreshold energy that swept through each neuron's excitatory tuning curve. The selectivity was also heterogeneous in that each cell fired to a particular subset of calls that was different from the subset to which the other cells fired. One cell failed to fire to any of the calls (adapted from Klug A, Bauer EE, Hanson JT, Hurley L, Meitzen J, Pollak GD. Response selectivity for species-specific calls in the inferior colliculus of Mexican free-tailed bats is generated by inhibition. J Neurophysiol. 2002 Oct;88(4):1941–54 [35])

that evoke discharges varied from neuron to neuron, even though the neurons are all tuned to the same frequency and all the signals have suprathreshold energy that stimulate the neuron's excitatory tuning curves.

The selectivity for calls is shaped in the IC by the inhibitory innervation that plays upon IC cells [35, 36]. The profound impact of inhibition on the coding of communication calls is illustrated in Fig. 8.4 which shows the responses of nine IC neurons to two different calls, social communication call 4 (SC4) and social communication call 6 (SC6). The responses evoked by the two calls were recorded before and while inhibition was blocked by the iontophoretic application of bicuculline and/or strychnine. Each of the nine neurons had a different BF and the cells are arranged from low to high, which corresponds to the tonotopic organization of the IC. Note that the calls had similar spectrotemporal features but evoked different responses among the population. Before inhibition was blocked, the nine neurons expressed different selectivities, since only three of the nine neurons responded to call SC4 and four different neurons responded to call SC6. Blocking inhibition virtually eliminated selectivity and allowed all nine neurons to respond to both calls.

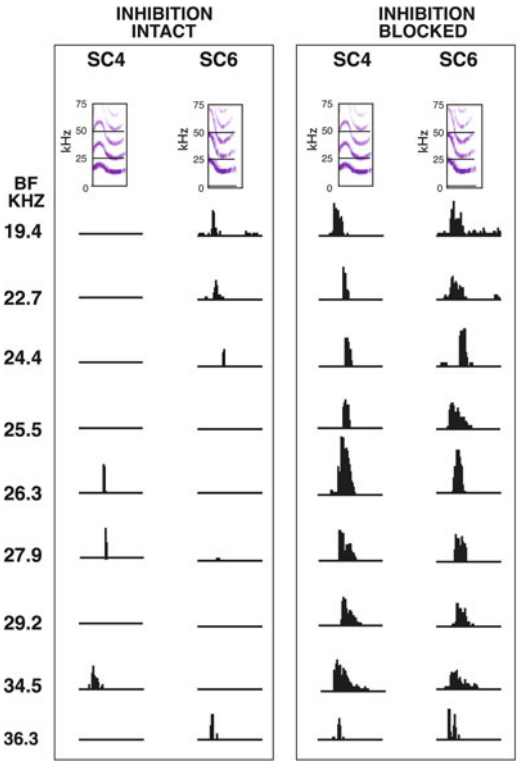


Fig. 8.4 Responses of nine IC neurons to two conspecific communication calls, SC4 and SC6, before and while inhibition was blocked. The BFs of the neurons are arranged from low to high, which corresponds to the tonotopic organization of the IC. Note that the calls had similar spectro-temporal features but evoked different responses among the "population." Before inhibition was blocked, the selectivities of the two neurons were different, the three neurons that responded to call SC2 did not respond to SC4, and the four neurons that responded to SC6 did not respond to SC4. Blocking inhibition eliminated selectivities and all neurons responded to both signals (adapted from Klug A, Bauer EE, Hanson JT, Hurley L, Meitzen J, Pollak GD. Response selectivity for species-specific calls in the inferior colliculus of Mexican free-tailed bats is generated by inhibition. J Neurophysiol. 2002 Oct;88(4):1941–54 [35])

Spectrotemporal Receptive Fields Reveal the Importance of Sideband Inhibition

The inhibitory feature that shaped selectivity was the structure of each neuron's sideband inhibition. Sideband, or surround inhibition as it is sometimes called, is comprised of the frequencies that flank the excitatory frequency region of a neuron's tuning curve and evoke inhibition. As indicated above, when sideband inhibition is eliminated by the iontophoretic application of bicuculline and/or strychnine,

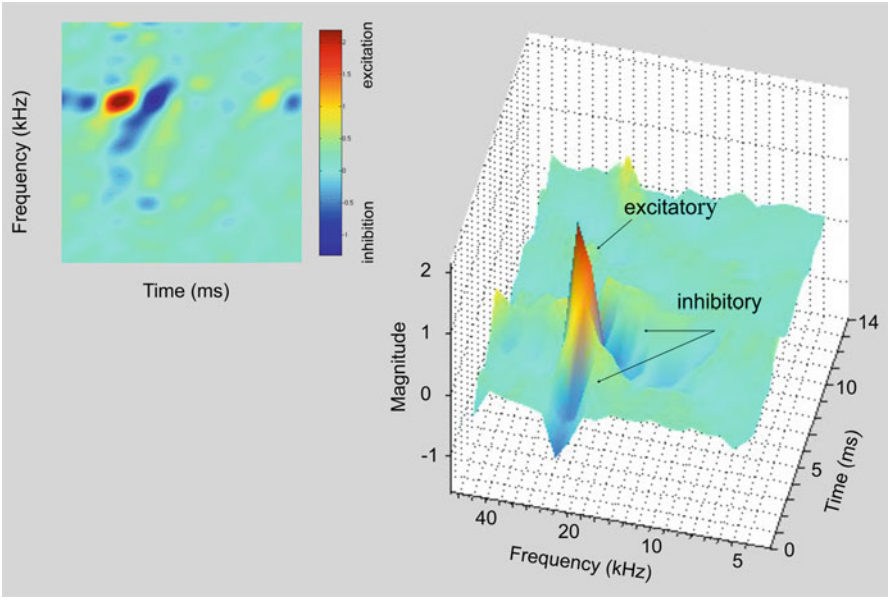


Fig. 8.5 A spectrotemporal receptive field (STRF) recorded from an IC neuron shown in both two-dimensional (*top panel*) and three-dimensional (*bottom panel*) views. See text for further explanation (adapted from Andoni S, Li N, Pollak GD. Spectrotemporal receptive fields in the inferior colliculus revealing selectivity for spectral motion in conspecific vocalizations. J Neurosci. 2007 May 2;27(18):4882–93 [37])

IC neurons respond to many more, or even all of the calls presented, than they did before inhibition was blocked [33, 35–39]. Specifically, it must be the timing and magnitude of inhibition relative to excitation that underlies selectivity, but exactly how those features are expressed in each IC cell and how they differ among IC cells to create the diverse selectivities among isofrequency cells could not be determined from blocking inhibition alone.

To obtain a more detailed picture of both the excitatory and inhibitory fields in IC cells, a large number of complex signals called moving ripples were presented. These are complex signals that contain a broad range of both spectral and temporal modulations that have been used by numerous investigators to generate spectrotemporal receptive fields (STRFs) [40–44]. We also used these signals to generate STRFs by a process analogous to spike-triggered averaging of the signals that preceded each spike [37]. The STRF derived from one IC cell is shown in both 2D and 3D forms in Fig. 8.5. The idea is that each ripple stimulus contains a broad spectrum. Frequencies that are always present prior to a discharge sum and thereby form the red region in the 2D and the peak in the 3D STRF. Frequencies that are rarely or never present prior to a discharge form the blue regions in the 2D and the valleys or nadirs in the 3D STRF. The frequencies represented in the peak and red colors are

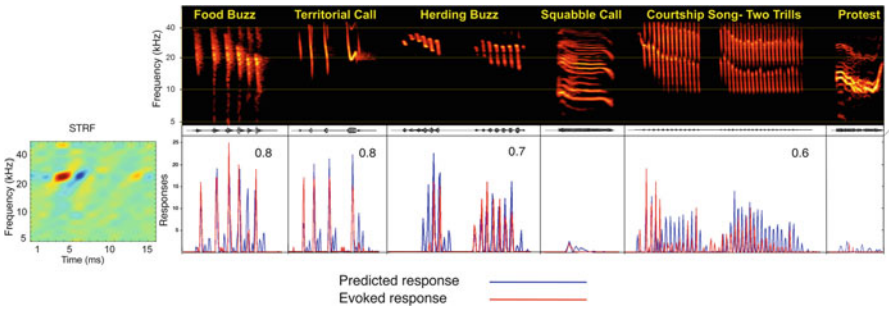


Fig. 8.6 STRF provides accurate predictions of responses to species-specific calls. Spectrograms of each species-specific vocalization are shown in the top, with the evoked responses (*red*) and the responses predicted from the STRF (*blue*) displayed below each call. Convolving the STRF with the spectrogram of each call generated predicted responses. The correlations between the predicted and actual responses are shown in *top right* of each panel. Convolutions predicted the call selectivity of the neuron because they predicted high response magnitudes for those calls that evoked strong responses, but they also predicted very low response magnitudes for the calls that evoked little or virtually no responses (adapted from Andoni S, Li N, Pollak GD. Spectrotemporal receptive fields in the inferior colliculus revealing selectivity for spectral motion in conspecific vocalizations. J Neurosci. 2007 May 2;27(18):4882–93 [37])

presumed to be excitatory, whereas the frequencies in the nadirs and blue colors are presumed to be inhibitory. Moreover, whenever the neuron fires, some frequencies will be present in a random fashion, thereby generating the green background color in the 2D and the green baseline in the 3D STRF in Fig. 8.5. Given these assumptions, the STRF provides a picture of relative magnitudes and temporal relationships of excitation and inhibition.

Convolving the STRF with a suite of communication calls should yield predicted responses that are in close agreement with the responses that are actually evoked by each call. The hypothesis is that the STRF is a linear filter that represents the optimal signal to which the neuron is tuned [42]. Thus, the prediction is that the strongest responses should be evoked by stimuli that are most similar to the spectrotemporal features of the neuron's STRF, and the more the spectrotemporal features of the signal differ from the STRF, the weaker the predicted response. The responses predicted by the convolutions can then be compared to the responses that were actually evoked by the same calls.

In about 25 % of the IC cells, the responses evoked by the calls were accurately predicted by the convolutions [37]. An example is shown in Fig. 8.6. The convolutions not only accurately predicted the calls to which the neurons responded, they also predicted the temporal discharge pattern evoked by each call. Equally important, they also predicted the calls to which the neurons did not respond. In short, the STRF in these cells captured the essential features of the cell and provided a picture of the relative magnitude and timing of excitation and inhibition, which in turn predicted how the cell would respond to any of the communication calls or to any other stimulus.

STRFS Explain FM Directional and Velocity Selectivities

Directional selectivity for FM sweeps is strongly influenced by inhibition, since blocking inhibition greatly reduces directional preferences in IC neurons [37, 45–47]. However, it is not inhibition per se that shapes directional selectivity, but rather the important feature is the tilting of the inhibitory fields along the spectrotemporal axis of the STRF, i.e., the degree to which their receptive field is inseparable [37]. Tilted inhibitory fields enhance directional preferences, or even create them, because signals sweeping in the non-preferred direction simultaneously evoke both excitation and inhibition, thereby suppressing responses to that FM direction, whereas signals sweeping in the preferred direction activate excitation and inhibition at different times, thereby allowing the cell to respond to the preferred direction (Fig. 8.7). This interpretation is supported by results obtained when inhibition was blocked by the iontophoretic application of bicuculline and/or strychnine [37]. Blocking inhibition not only reduced or even eliminated the inhibitory fields in their STRFs (Fig. 8.8) but also reduced both inseparability and direction selectivities in the IC (not shown).

The degree of tilt in the receptive field shapes both the neuron's directional selectivity and the FM velocity that evokes the strongest response [37, 48]. The response strength is determined by the correspondence between the tilt in the excitatory field and the rate of frequency sweep or FM velocity. Thus neurons with strong tilts are most sensitive to high FM velocities, whereas neurons with lesser tilts are most sensitive to lower FM velocities. Based on these features, it was estimated that most IC neurons had best velocities between 5 and 100 octaves/s, with a mean of ~60 octaves/s (Fig. 8.9c).

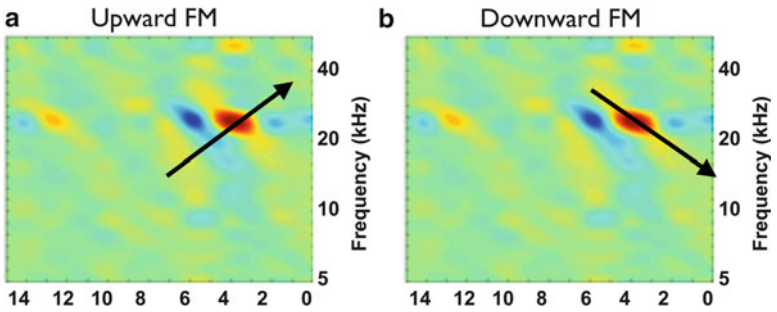


Fig. 8.7 Tilted receptive fields impart directional selectivity for FM sweeps. (**a**) Cell with a tilted (inseparable) receptive field. *Arrows* indicate how a downward (*left panel*) and upward (*right panel*) FM sweep would traverse the STRF at one point in time. The key feature is that at some point in time, the downward FM will only sweep through the excitatory portion of the STRF without encroaching upon the inhibitory portion and thereby excite and drive the cell. In contrast, the upward FM will never encroach only upon the excitatory part of the STRF but rather will sweep through both its excitatory and inhibitory portions, which will suppress excitation thereby preventing the neuron from firing. This is the same STRF shown in Figs. 8.5 and 8.6 but flipped in time (STRF is adapted from Andoni S, Li N, Pollak GD. Spectrotemporal receptive fields in the inferior colliculus revealing selectivity for spectral motion in conspecific vocalizations. J Neurosci. 2007 May 2;27(18):4882–93 [37])

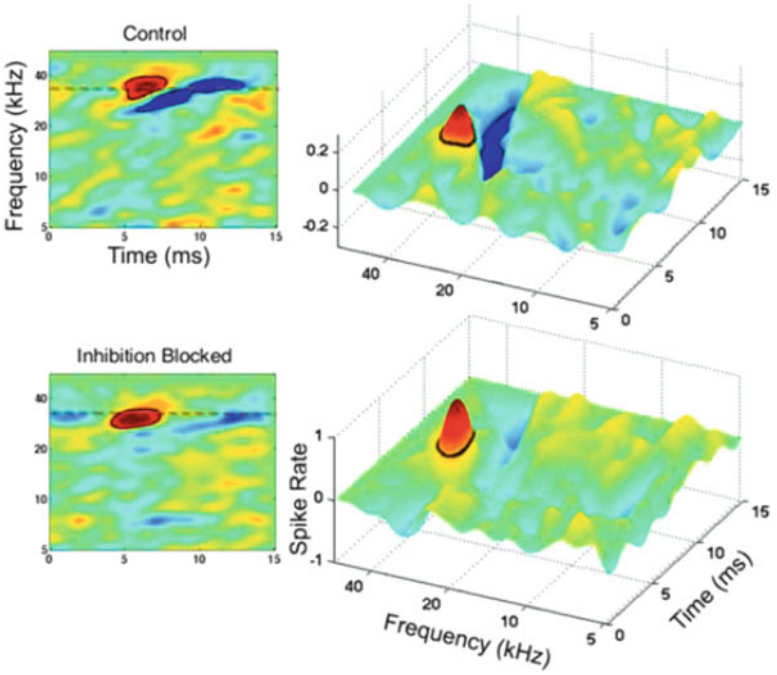


Fig. 8.8 Effects of blocking inhibition on the STRF of an IC neuron. The reduction in surrounding inhibition in the STRF of a neuron is apparent from a comparison of the inhibitory (*blue*) regions of the STRFs before blocking inhibition (control) and while inhibitory receptors were blocked by the iontophoretic application of bicuculline and strychnine (adapted from Andoni S, Li N, Pollak GD. Spectrotemporal receptive fields in the inferior colliculus revealing selectivity for spectral motion in conspecific vocalizations. J Neurosci. 2007 May 2;27(18):4882–93 [37])

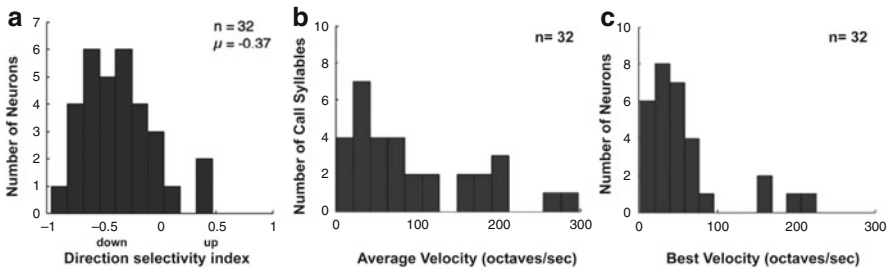


Fig. 8.9 Directional selectivities of IC cells, call velocities, and best velocities of IC cells. (**a**) Distribution of directional selectivity indices in 32 IC cells. Downward preferring cells have negative indices, upward preferring cells have positive indices and nondirectional cells have an index of 0. A small number of cells preferred upward sweeps or were nondirectional. The vast majority, however, preferred downward FMs. (**b**) Distribution of the FM velocities found in 21 calls that contained 32 different syllables. (**c**) Distribution of best velocities to which 32 IC neurons are tuned. The two distributions are well correlated ($r = 0.7$), showing the close correspondence between the FM velocities in their communication calls and the FM velocities to which IC neurons are tuned. The STRFs of all cells shown yielded good predictions for responses to communication calls (adapted from Andoni S, Li N, Pollak GD. Spectrotemporal receptive fields in the inferior colliculus revealing selectivity for spectral motion in conspecific vocalizations. J Neurosci. 2007 May 2;27(18):4882–93 [37])

Of particular importance is the close agreement between the FM features in their conspecific communication sounds and the tuning for those FM features among the IC population [19, 37, 48]. As can be seen in the spectrograms of the various calls in Fig. 8.3, all echolocation and most communication signals emitted by Mexican free-tailed bats contain FMs. At least a portion of the FMs in almost all calls sweep downward at velocities ranging from 0 to 250 octaves/s (Fig. 8.9c). Consistent with these signal features, the IC of all bats have cells selective for both upward and downward FMs, but the majority of cells are selective for the downward direction (Fig. 8.9a) [37, 45–47, 49–51]. Moreover, the range of preferences for sweep velocities corresponds closely to the sweep velocities in the signals these animals emit (Fig. 8.9b, c) [37]. Thus the structure of their excitatory and inhibitory fields biases many IC neurons for downward direction selectivity and shapes their responsiveness to the FM velocities and other features present in their vocalizations.

Predictive STRFs Were Found in Only a Minority of IC Neurons

The STRFs of cells in which the convolutions accurately predicted responses and explained response selectivities present a comprehensive view of the quantitative features of excitation and inhibition in both frequency and time. The cells that yielded predictive STRFs must have linearly added the response of inhibitory and excitatory frequencies evoked by the rippled stimuli. Since STRFs reflect the average signal generated by such linear additions, the average representation of the excitatory and inhibitory fields generated by ripple stimuli was appropriate for predicting responses to other complex stimuli, such as the communication calls.

The neuronal population in the IC, however, is heterogeneous [19, 20, 52], and most IC cells did not behave in the relatively simple way that the IC cells described above did. Specifically, predictive STRFs were found in only 25 % of IC cells; the STRFs in most cells (~75 %) provided poor predictions or were non-interpretable [37]. Those cells apparently had either static or dynamic nonlinear response properties that were stronger than the linear response properties extracted by the STRFs generated by ripples. Stated differently, there was no linear relationship between the magnitudes of the excitation and inhibition in time and frequency that would apply to every complex signal. Therefore, the "STRFs" computed for those cells could not predict the response to a new complex signal, such as the conspecific calls, because the nonlinear interactions of excitation and inhibition would be different than and scale differently for the call than the average derived from the ripples.

Most Neurons Had More Than One Spectrotemporal Filter

Neurons in which the STRF generated by spike-triggered averaging yielded poor predictions for calls had multiple spectrotemporal features of the stimulus that defined the neuron's overall receptive field [48]. In these neurons, the nonlinear

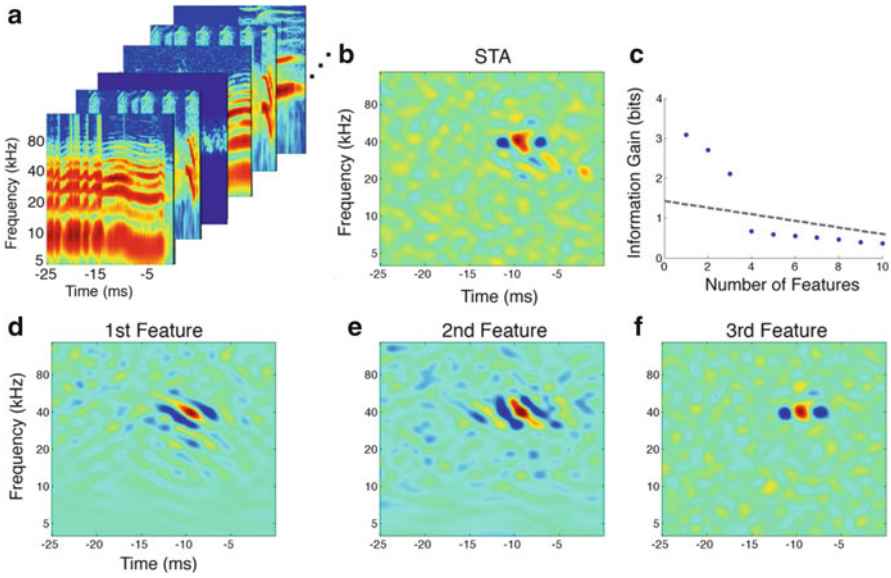


Fig. 8.10 Multiple spectrotemporal features in a nonlinear IC cell. To extract the relevant features encoded by an IC neuron, a large number of communication calls was presented to IC neurons, and each call or stimulus segment that preceded a spike was summed to generate the spike-triggered average, as shown in (**b**). Only a portion of the total calls is shown in (**a**). The ensemble of calls that evoked spiking was then searched for the set of spectrotemporal features that maximized the amount of information preserved between the stimulus and the spiking response. The plot in (**c**) shows the amount of information gained as the number of spectrotemporal features considered is increased. The first three most relevant features are shown in (**d**) (adapted from Andoni S, Pollak GD. Selectivity for spectral motion as a neural computation for encoding natural communication signals in bat inferior colliculus. J Neurosci. 2011 Nov 16;31(46):16529–40 [48])

combination of multiple spectrotemporal features predicted the neuron's spiking responses. The computation used was a spike-triggered covariance procedure somewhat similar to principal component analysis [53, 54]. This method yielded multiple relevant features in most of IC cells, where the first spectrotemporal feature captured the most information on the stimulus–response relationship of each neuron (Fig. 8.10). In this study, the set of relevant spectrotemporal features was not computed from rippled stimuli, but rather was computed from the responses evoked by a large number of conspecific communication calls. We used natural calls because previous studies showed that STRFs derived from natural stimuli in both the IC of songbirds [55] and in the cortex of ferrets [56] were significantly different than the STRFs derived with synthetic stimuli. Most importantly, the receptive fields derived with natural stimuli provided far better predictions of responses to natural calls than did the receptive fields derived with synthetic stimuli [55].

Predicted responses for both electronically generated FMs and conspecific calls were then calculated using either the first most informative spectrotemporal feature alone or the two most informative spectrotemporal features (Fig. 8.11). The most

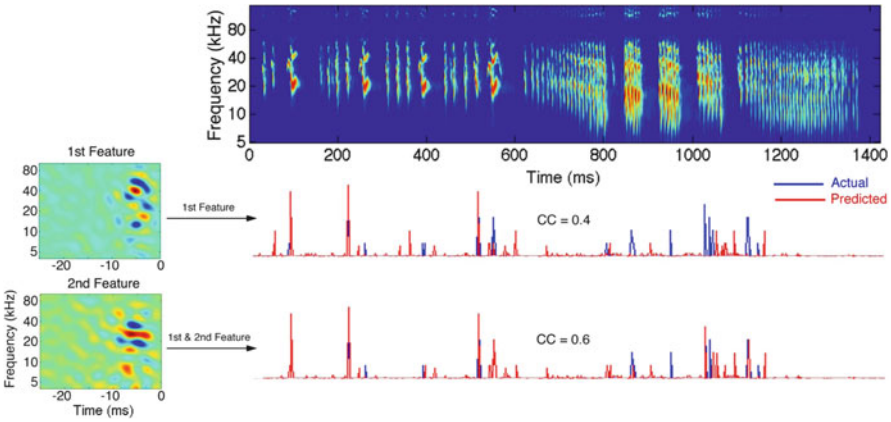


Fig. 8.11 Convolution with first feature only and with both first and second features. Responses predicted from spectrotemporal features improve when multiple stimulus features are considered. The two most informative features of an IC neuron are shown. Convolving the calls shown with only the first feature yielded poor predictions, with an average correlation coefficient between the predicted responses and those evoked by the calls of only 0.4. When both the first and second features were used to calculate the predicted responses, the correlation coefficient increased to 0.6. This shows that this IC neuron is tuned for multiple spectrotemporal features of natural calls (adapted from Andoni S, Pollak GD. Selectivity for spectral motion as a neural computation for encoding natural communication signals in bat inferior colliculus. J Neurosci. 2011 Nov 16;31(46):16529–40 [48])

significant finding was that the predicted responses were poor when only the first feature was used but improved significantly when two features were used. The correlation coefficient between the predicted and the evoked responses for calls had a mean of 0.46 with only one feature but increased to a mean of 0.61 when two features were used. Furthermore, using a two-feature model captured greater mutual information between the calls and their responses than using each either feature independently. This showed that these neurons did indeed have two or more spectrotemporal filters that determined the responses to calls. The relevance of the two filters was further supported by the near perfect agreement between the responses evoked by electronically generated FMs and the responses predicted with the nonlinear combination of the two most relevant features.

Most IC Cells Are Tuned to Nonredundant Spectrotemporal Modulations

The FM velocities to which the first features were tuned agreed closely with the FM velocities in the bats' communication calls (Fig. 8.12b). A more detailed analysis of these features revealed an interesting relationship that further contributes to response

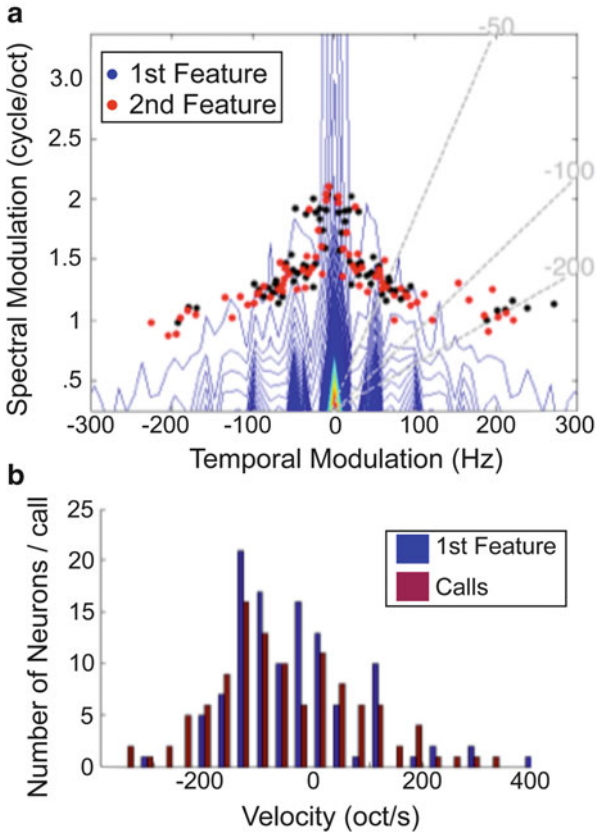


Fig. 8.12 Comparison of spectrotemporal modulations in conspecific calls to neuron tuning. (**a**) The contour plot shows the modulation spectrum of a large repertoire of bat calls. The *black* and *red dots* designate the peak modulations present in the first and second most informative features of IC neurons. Note that the peak tuning in the IC is organized to detect various FM velocities (*dashed lines*) while avoiding redundant energy found in most calls. (**b**) Distribution of FM velocities found in the calls match the velocities of the most informative feature that represents the velocity tuning of IC neurons. This suggests that IC neurons are tuned to detect the range of FM cues present in their social communication calls (adapted from Andoni S, Pollak GD. Selectivity for spectral motion as a neural computation for encoding natural communication signals in bat inferior colliculus. J Neurosci. 2011 Nov 16;31(46):16529–40 [48])

selectivity for calls. When the modulation spectrum of bat vocalizations in their repertoire is overlaid on the modulation tuning of the first and second most informative spectrotemporal features of the nonlinear cells, the best modulation tuning in the majority of these cells fall on either side of the dense areas in the contour plot (Fig. 8.12). This indicates that most neurons are relatively insensitive to the modulations that are most common or redundant across the calls. This property of IC neurons was previously shown in the midbrain of songbirds [57]. The majority of neurons are tuned instead to modulations that represent FM directions and velocities that are

present in some calls but not others. What this suggests is that the tuning in the IC is designed to detect modulations that deviate from the common modulations found across calls, thereby allowing each neuron to be selective for the modulation that represents a different direction and velocity. In this way, each IC neuron responds most strongly to calls that have the FM sweeping direction and velocity to which the neuron is tuned while failing or only responding weakly to calls with modulations outside of its tuning.

Directional Preferences for FMs Measured with In Vivo Whole-Cell Recordings

It was shown that the iontophoresis of drugs that blocked GABAergic and glycinergic receptors eliminated the inhibitory fields in IC cells with a single filter determined by spike-triggered averaging, which showed that the inhibitory fields were generated primarily at the IC (Fig. 8.8) [37]. However, blockers were not used in the studies of cells with multiple filters, and thus the underlying events in these cells could not be determined with extracellular recordings [48]. Thus in most STRF studies, in which neural activity is recorded with extracellular electrodes, inhibition cannot be measured directly, but rather inhibition has to be inferred from the suppressive effects of some stimulus manipulation on the excitation evoked by another signal. Furthermore, with extracellular recordings, there is an uncertainty about whether the observed spike suppression was due to inhibition at the IC or whether suppression was inherited from the inhibition that occurred in a lower nucleus that projects to the IC.

To overcome some of these uncertainties, a more direct and detailed view of sound-evoked inhibition was obtained with in vivo whole-cell recordings from the IC in response to FM sweeps in awake bats [58–60]. With patch recordings, as with extracellular recordings, the discharges evoked in most IC cells exhibited a preference for downward sweeping FMs [58, 61]. With patch recordings, however, both the inputs to the cells, expressed in the amplitudes of postsynaptic potentials (PSPs), and their outputs, their discharges, are obtained (Fig. 8.13). The selectivity differences of the inputs can be quantified by computing a PSP directional index (PSP amplitude evoked by the downward FM minus PSP amplitude evoked by the upward FM divided by the sum of the two amplitudes). Similarly, the selectivity differences of the outputs, discharges, are quantified by computing a discharge directional index based on spike counts rather than PSP amplitudes. Thus, the directional preferences of the inputs can be quantitatively compared to the directional preferences of the outputs.

In most IC cells, the differences in the discharge vigor evoked by upward and downward FMs are substantially greater than the differences in the magnitudes of the excitatory postsynaptic potential (EPSPs) evoked by the same signals [58]. The discharge output of the cell in Fig. 8.13, for example, was perfectly selective for the preferred (downward) FM; it fired to every presentation of the preferred FM and

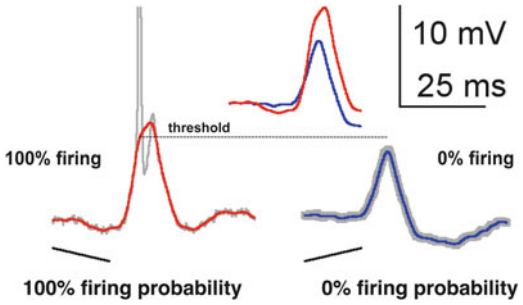


Fig. 8.13 Responses recorded with a patch electrode to upward and downward FMs. Notice that there was only a slight difference in EPSPs evoked by upward and downward FMs yet the neuron fired to every downward FM presented and failed to fire to any of the upward FMs. The nonlinear amplification imparted by spike threshold rendered this neuron 100 % selective for downward FMs

never fired to the null (upward) FM and thus had a discharge directional index of 1.0. In marked contrast, the EPSP amplitudes evoked by the two signals were similar, and the same cell had a PSP directional index of only 0.2. The disparity in the high spike selectivity compared to the low EPSP selectivity is due to the nonlinear influence of spike threshold, where the larger EPSP evoked by the preferred FM in this cell was above threshold, and evoked a discharge on every presentation, whereas the EPSP evoked by the null FM, while only slightly smaller, was a few mV below threshold and failed to evoke spikes. Although the cell in Fig. 8.13 is an extreme example, it illustrates the general finding that the inputs (PSPs) were less selective than the outputs (spikes). On average, the spike-DSI was more than twice as large as the PSP-DSI among the IC population [58].

The Role of Spike Timing for Creating Directional Selectivity

While the intracellular recordings with patch electrodes, like the recordings with extracellular electrodes, showed that most IC cells express directional preferences for FM sweeps, the comparison of PSPs and spikes did not by itself show how the interactions of excitation and inhibition shaped the directional preferences of the cells. Previously, we proposed that the directional preferences of cells in which their STRFs predicted responses to calls is formed by the relative timing of the excitatory compared to the inhibitory inputs evoked by an FM sweep, which is the most widely accepted explanation for the formation of directional preferences. The acceptance of this explanation is based on two principal observations. The first is that neurons selective for downward (or upward) FM sweeps have inhibitory fields that are lower (or higher) in frequency than the frequencies that activate their excitatory fields. These features were confirmed by the excitatory and inhibitory response fields in the linear STRFs, as illustrated by the cell in Figs. 8.5–8.7, and were shown in a large number of previous studies [37, 49, 62–68]. The second observation is that blocking inhibition reduces or eliminates directional preferences, as shown for

cells with linear STRFs (e.g., Fig. 8.8) and by other investigators in previous studies [37, 45, 46, 69]. The timing hypothesis, which follows from the results of those experiments, posits that downward FM signals first sweep through the excitatory field, thereby evoking an initial excitation, and slightly later in time, the signal sweeps through the inhibitory field [45, 63, 64, 66, 68, 70]. With upward sweeping FMs, on the other hand, inhibition is activated first, and the initial inhibition quenches the subsequent excitation. This is exactly the result obtained from the STRFs of IC neurons shown in Fig. 8.7. The same arguments apply for upward preferring cells, but the frequencies of the excitatory and inhibitory fields are reversed.

We point out that there is an additional implicit assumption in this explanation. Specifically, the explanation assumes that the inputs behave in a linear manner, where the excitatory and inhibitory inputs are evoked in synch with the spectrotemporal features of the signals. Thus, the same excitation and inhibition are evoked by downward and upward sweeping FMs, where the timing of excitation and inhibition is reversed because the temporal features of the signals are reversed. This is the assumption used to explain the directional selectivity based on the STRF shown in Fig. 8.8 and for all the other IC neurons that had linear STRFs; the strengths and relative timings of excitation and inhibition should simply be reversed as the FM direction is changed from upward to downward.

In short, there is strong evidence from a variety of different studies in a variety of mammals to support the hypothesis that spectral arrangement of the excitatory and inhibitory fields generates FM directionality, and that hypothesis also explains why blocking inhibition eliminates directionality.

FM Directional Selectivity Formed by Timing Disparities of Excitation and Inhibition Does Not Apply to All IC Cells

The IC, however, is heterogeneous, where the formation of a particular response property is formed in different ways among its neuronal population [19, 20, 71]. With regard to the formation of FM directional preferences, sensitivity for small timing differences between excitation and inhibition should be effective in cells with low input resistances and fast time constants. A recent study of IC cells in bats showed that about half of the cells in the IC do indeed have low input resistances that range from 40 to 100 Ω and fast time constants [52].

Presumably these are the cells whose FM preferences are formed by the relative timing of excitation and inhibition. The other side of the finding is that about half of the IC population has high input resistances and long time constants, features that are inappropriate for sensitivity to small changes in the timing of excitation and inhibition.

In this regard, it is interesting that nonlinear cells with multiple spectrotemporal features had symmetric nonlinearities that indicate their insensitivity to the timing of excitation and inhibition, in contrast to cells with linear STRFs. It may well be that the nonlinear cells had high input resistances and slow time constants, although there is no direct evidence of this correspondence.

The Timing of Excitation and Inhibition Was Explored with In Vivo Whole-Cell Recordings

To evaluate the role of the timing of excitation and inhibition in IC cells with high input resistances, the excitatory and inhibitory conductances that generated the responses to an upward and a downward sweeping FM were computed in a subset of IC neurons [58–60]. The timing of the inhibitory conductances was then advanced or delayed in time relative to the excitatory conductances, and the EPSPs that would be evoked by changes in the timing were computed in a model.

The justification for modeling the responses is that the predicted EPSPs computed from the conductances were in close agreement with the responses actually evoked by the preferred and null FM signals, as illustrated by two cells in Fig. 8.14. The excitatory (ge, red line) and inhibitory (gi, blue line) conductances derived for the preferred and null FMs are shown in Fig. 8.15 and were used to compute the predicted responses shown in Fig. 8.14. Since the conductances predict the EPSPs evoked by acoustic signals, it follows that they should also provide an accurate prediction of the EPSPs that would be evoked if one of the conductances, the inhibitory conductance in this case, was delayed or advanced in time. Moreover, since firing thresholds were determined from the sound-evoked responses, the amplitudes of the modeled EPSPs could be related to the cell's threshold. Given these assumptions, the modeled responses would permit an assessment of the consequences of shifting the timing of the inhibitory inputs, rather than shifting the timing of the spectral components in the acoustic signals, as is universally done to assess the role of input timing.

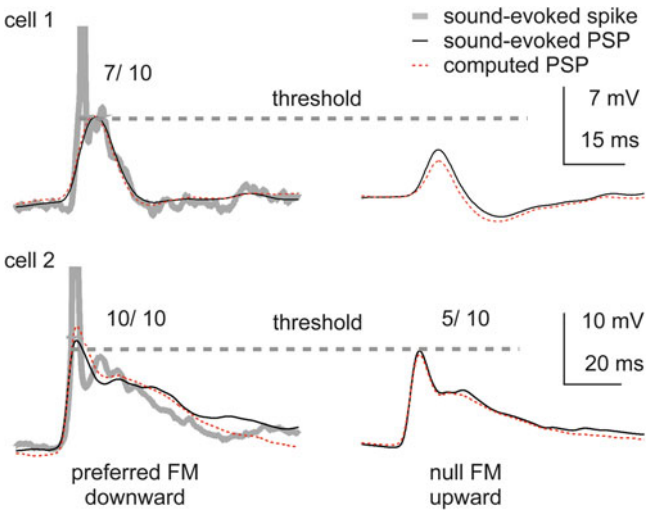


Fig. 8.14 Two directionally selective cells. *Black traces* are the measured PSPs (mean of ten trials, spikes removed by filtering), *red traces* are PSPs computed from derived conductance waveforms, and *gray traces* illustrate spiking with a single sweep response. *Dashed line* is spike threshold

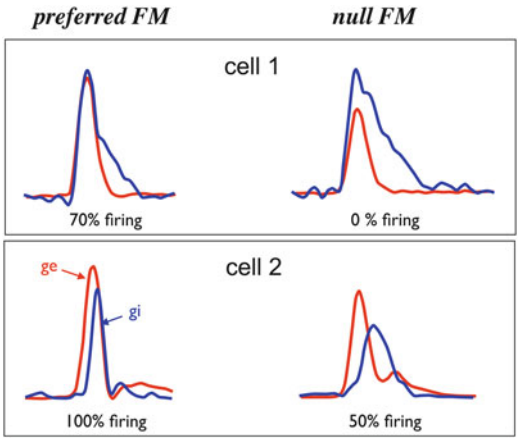


Fig. 8.15 Timing of excitation and inhibition provides no information about directional preferences. The calculated excitatory (ge shown in red) and inhibitory (gi shown in blue) conductances for a downward and upward FM sweep in two IC neurons. These are the same cells whose responses to the FMs are shown in Fig. 8.14. In cell 1 (*top panel*) the timing of the excitatory and inhibitory conductances were virtually simultaneous for both the preferred and null FMs. In cell 2, the excitatory conductance led the inhibitory conductance for both the preferred and the null FMs. However, the lead time of excitation was even greater for the null than the preferred. In both cells, the differences in the responses to the two FMs are due largely to the differences in the shapes and amplitudes of the excitatory compared to the inhibitory conductances rather than to their relative timing

Those experiments showed several important features of the conductances evoked by the preferred and null FMs, as well as several other features of the EPSPs that would occur when the timing of inhibition is advanced or delayed. The features of the EPSPs evoked by the FMs are considered first followed by the changes in the computed EPSP amplitudes that occur as inhibition is delayed or advanced.

The first important feature is that in every cell, the excitatory conductances evoked by both the preferred and null FMs by themselves evoked a suprathreshold response. This finding is consistent with the general finding from extracellular studies that blocking inhibition reduces or eliminates the directional preferences in almost all IC cells, allowing the cells to fire to both FMs.

The second finding is that the amplitudes of the excitatory and inhibitory conductances evoked by the preferred FM and null FMs are always different. In other words, even though the spectral composition of the preferred and null FMs are identical but reversed in time, each signal does not evoke the same but time-reversed excitatory and inhibitory conductance waveforms. Rather, the excitatory and inhibitory conductance waveforms evoked by the preferred FM differ in either waveform shape or amplitude or both shape and amplitude from the conductances evoked by the null FM.

The third finding is that there was no consistent relationship between the timing of the excitatory and inhibitory conductances evoked by the preferred FM compared to the null FM.

The two cells in Fig. 8.15 illustrate two of the three features. Although not shown, the EPSPs of the preferred and null FMs computed only from the excitatory

conductances were above threshold in both cells. In addition, the waveforms of the excitatory and inhibitory conductances of the preferred and null FMs differed in shape and in peak amplitude (the exception is the peak amplitudes of the inhibitory conductances for cell 1, which were about the same). Finally, excitation and inhibition in cell 1 were virtually coincident for both the preferred and null FMs. In cell 2, in contrast, excitation led inhibition in the response to the preferred FM, but excitation led by an even greater amount of time in the response to the null FM. Since the relative timings of the excitatory and inhibitory conductances evoked by the preferred and null FMs differed from cell to cell, the relative timing of excitation and inhibition by itself provides little or no information about the preferences of these cells for the direction of an FM sweep.

Small Changes in the Timing of Inhibition Relative to Excitation Caused Only Small Changes in Response Amplitude of the Null FMs

These experiments also showed that small (i.e., 1–2 ms) changes in the relative timing of inhibition relative to excitation had only minor influences on the amplitudes of the computed EPSPs of the null FMs, where EPSP amplitudes changed by about 1.0 mV or less for each 1.0 ms advance or delay of the inhibition. These small changes in PSP amplitudes of the null FM with changes in the timing of inhibition are illustrated by cell 1 in the top panel of Fig. 8.16. Delaying inhibition caused only small changes in EPSP amplitude, in which delays of 1–5 ms increased EPSP amplitudes at a rate of ~1 mV/ms. EPSPs were always below threshold with a delay less than that required to bring the EPSP amplitude close to spike threshold; inhibition had to be either delayed by ~4 ms or advanced by considerably more than 10 ms.

The changes in null EPSP amplitudes as inhibition was advanced or delayed were even smaller for cell 2 (Fig. 8.16, lower panel), although the results of small shifts were considerably different than they were for cell 1. A significant feature of cell 2 is that the null control EPSP evoked a discharge probability of 50 %. Inhibitory delays caused only small increases in EPSP amplitudes of less than 1.0 mV/ms. Those small changes in amplitude, however, would almost certainly have caused a change in discharge probability because the membrane potential, which was already hovering at or close to threshold, would have been brought closer to or even above threshold with a timing advance of 1–2 ms. In this way, spike threshold could act to amplify the small change in membrane potential into a larger change in discharge probability evoked by the null FM. Thus, in some cells, e.g., cell 1, the inhibition in the null FM had to be advanced by 5.0 or more ms before the cell's directional preference could be changed. In other cells, e.g., cell 2, even small increments in EPSP amplitude due to small changes in timing, could induce substantial changes in the cell's directional preference.

There is a final point to be made by these manipulations. Namely, that the impact of delays or advances of inhibition relative to excitation is further complicated by

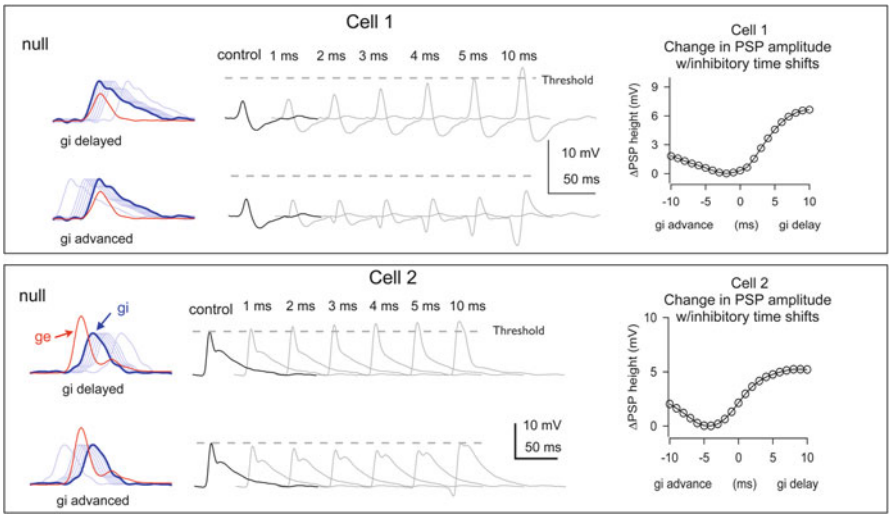


Fig. 8.16 Small temporal shifts of inhibition relative to excitation have only small effects on EPSP amplitudes of responses evoked by null FMs. The calculated excitatory conductance (ge) is shown as *red lines* and the inhibitory conductances (gi) are shown as *blue lines*. The *dark blue lines* in all records shows the temporal relationship of the inhibitory conductance evoked by the FM sweep relative to the excitatory conductance. The *lighter blue lines* show the time shifts of the inhibitory conductances. The EPSP evoked by the upward (null) FM is shown as the control response. The inhibitory conductance was then either advanced or delayed in 1.0 ms steps and predicted EPSP was then computed for each temporally shifted conductance. Cells 1 and 2 are the same cells shown in Figs. 8.14 and 8.15. In cell 1 (*top panel*) the control EPSP was so far from spike threshold that long delays of about 5.0 ms would be needed to evoke spikes. The change in EPSP amplitude with each time shift is plotted and shown in graphical form on the *far right*. In cell 2, the change in EPSP amplitude with each temporal shift was also small, but since the control EPSP was so close to or even just at threshold, even a small increase in EPSP amplitude should increase the spiking probability (adapted from Gittelman JX, Pollak GD. It's about time: how input timing is used and not used to create emergent properties in the auditory system. J Neurosci. 2011 Feb 16;31(7):2576–83 [60])

the changes in almost all features of the excitatory and inhibitory conductances that occur with almost any change in signal parameters [59, 60]. Even the upward and downward FMs presented to a given cell, which had the same durations, the same intensities, the same frequency compositions, and the same power spectra, each evoked excitatory and inhibitory conductance waveforms that differed in relative latency, waveform shape, and magnitude. Since all of those features largely determine the cell's timing sensitivity, the two FMs produce different timing sensitivities in the same cell. This differential timing sensitivity is illustrated in Fig. 8.17, which shows the changes in PSP amplitudes of the preferred and null FMs with small delays in the timing of inhibition in cell 2. Delaying the inhibition of the preferred FM by 2 ms caused PSP amplitudes to increase by ~4.0 mV, whereas the same inhibitory delay in the null FM caused PSP amplitudes to increase by only 1.4 mV. For the same delays, the increases in PSP amplitudes for the preferred were nearly

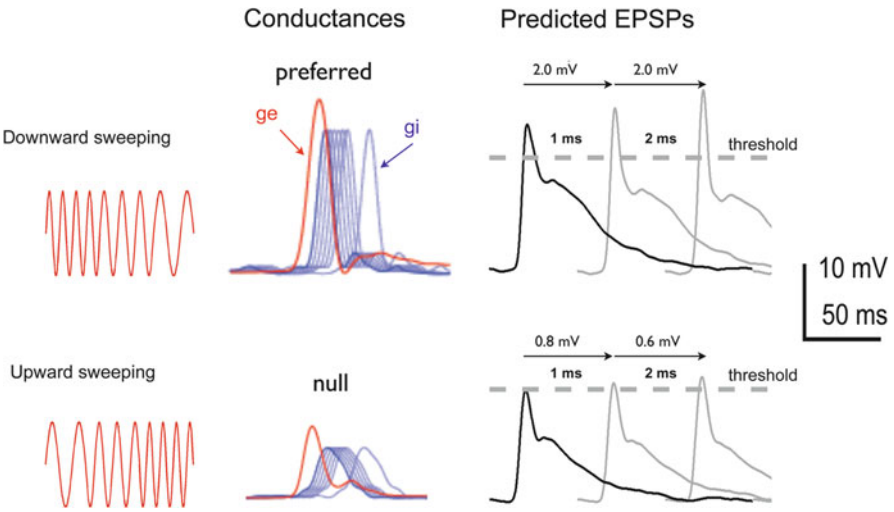


Fig. 8.17 Differences in effects of timing on EPSP amplitudes for preferred and null FMs in cell 1. Small timing delays of 1 or 2 ms caused EPSP amplitude to increase by about 2.0 mV/ms shift for the preferred FM. In marked contrast, the same timing delays in the null FM caused only very small increases of less than 1.0 mV/ms in EPSP amplitude

three times as large as they were for the null. Thus, the timing sensitivity of an IC cell is not a constant but rather varies from signal to signal, because the features of conductances change as signal parameters are varied.

Summary and Conclusions

The results of the studies we reviewed here illustrate five general features of processing in the IC. The first is the dominant role that inhibition plays in shaping the responses of IC neurons. The roles of inhibition are illustrated by the marked change in the response selectivity for communication calls when inhibition is blocked and by the prominent roles of sideband inhibition for shaping FM directionality.

The second feature is the heterogeneity of mechanisms that shape the response properties. The "filters" or STRFs that characterize IC neurons illustrate this heterogeneity. The processing in about 25 % of the IC population is linear and determined by a single filter, the STRF generated by spike-triggered averaging. Thus, simply convolving any of the calls with the cell's STRF provides a highly accurate prediction of how these neurons actually respond not only to individual syllables in a call, but to phrases or even the entire call. However, in about 75 % of IC neurons, convolving calls with the STRFs generated by spike-triggered averaging provided poor predictions to calls. These cells have multiple "filters" and thus are tuned to multiple features of the natural communication signals. The nonlinear combination of the filters defines the overall receptive field of the neuron, and thus, convolving the calls

with only one spectrotemporal feature, or filter, provides a poor response prediction which is greatly improved when the convolutions utilize with two or more filters.

The third feature follows from the above and shows that there is not a single mechanism that the IC employs to form a given response property, but rather there are multiple ways in which the same response property is formed among the IC cell population. In cells with a single filter, FM directionality is sculpted by the tilts in their excitatory and inhibitory receptive fields. This arrangement causes the cells to fire when excitation precedes inhibition, as occurs in response to the preferred FM direction, but prevents firings when inhibition and excitation are either coincident in time or when inhibition precedes excitation, features that are generated in response to the null FM. This hypothesis is supported not only by the arrangement of the excitatory and inhibitory fields but also by the decrease or elimination of directional selectivity when inhibition was blocked in these cells. In other cells, FM directionality is formed nonlinearly. The nonlinearity is most directly and clearly seen in the relationship of the excitatory and inhibitory conductances in IC cells with high input resistances. Whether these cells correspond to cells with multiple filters is unclear, but what is clear is that in cells with high input resistances, "timing" is more than just the relative latencies of excitation and inhibition. The response evoked by a signal is shaped by the interaction among the temporal features of the inputs, the relative latencies of excitation and inhibition, and the shapes of the excitatory and inhibitory conductances together with their magnitudes. Each of those features not only shapes the response evoked by a particular signal, but each also shapes the degree to which EPSP amplitudes change due to delaying or advancing inhibition relative to excitation.

The fourth feature is the nonlinear influence spike threshold. In cells with high input resistance, changes in relative timing between excitation and inhibition caused only small changes in EPSP amplitudes on the order <1.0 mV/ms. How the change in EPSP amplitude influenced discharge probability depended in large part on how close the EPSP was to spike threshold. If the EPSP amplitude is far from threshold, even timing changes of 4–5 ms might result in EPSP amplitude increases that would have little or no effect on spike probability. Conversely, if the EPSP amplitude is near threshold, then even an increase or decrease in EPSP amplitude as small as a fraction of a millivolt could affect discharge probability, and thus modulate the cell's spiking directional selectivity.

The fifth feature is the close correspondence between neural tuning and acoustic properties of conspecific communication signals. In Mexican free-tailed bats at least, this correspondence suggests that IC neurons are specifically encoding features of these signals through the neural computations that generate FM selectivity. Moreover, it is clear that the various selectivities expressed by IC neurons for communication calls are a consequence of the multiple ways in which their selectivities for features of acoustic signals, such as the direction and rate of FM sweeps, are created. The advantage conferred by the multiple formations of response properties in the IC is to amplify differential response selectivities for complex signals. The amplification is expressed by different and unique patters of activity among the neuronal population in the IC that are evoked even by signals with only subtle differences in their spectrotemporal features.

References

1. Griffin DR (1986) Listening in the dark. Cornell University Press, Ithaca, NY
2. Behr O, von Helversen O (2004) Bat serenades-complex courtship songs of the sac-winged bat (*Saccopteryx bilineata*). Behav Ecol Sociobiol 56:106–115
3. Bohn KM, Schmidt-French B, Schwartz C, Smotherman M, Pollak GD (2009) Versatility and stereotypy of free-tailed bat songs. PLoS One 4(8):e6746
4. Boughman JW (1998) Vocal learning by greater spear-nosed bats. Proc R Soc Lond B 265(1392):227–233
5. Esser K-H (1994) Audio-vocal learning in a nonhuman mammal—the lesser spear-nosed bat *Phyllostomus discolor*. Neuroreport 5(14):1718–1720
6. Knornschild M, Nagy M, Metz M, Mayer F, von Helversen O (2009) Complex vocal imitation during ontogeny in a bat. Biol Lett 6(2):156–159
7. Knornschild M, Behr O, von Helversen O (2006) Babbling behavior in the sac-winged bat (Saccopteryx bilineata). Naturwissenschaften 93(9):451–454
8. Kanwal JS (1999) Processing species-specific calls combination-sensitive neurons in an echolocating bat. In: Hauser MD, Konishi M (eds) The design of animal communication. MIT Press, Cambridge, MA, pp 133–156
9. Esser KH, Condon CJ, Suga N, Kanwal JS (1997) Syntax processing by auditory cortical neurons in the FM-FM area of the mustached bat Pteronotus parnellii. Proc Natl Acad Sci U S A 94(25):14019–14024
10. Bohn KM, Schmidt-French B, Ma ST, Pollak GD (2008) Syllable acoustics, temporal patterns, and call composition vary with behavioral context in Mexican free-tailed bats. J Acoust Soc Am 124(3):1838–1848
11. Schwartz C, Tressler J, Keller H, Vanzant M, Ezell S, Smotherman M (2007) The tiny difference between foraging and communication buzzes uttered by the Mexican free-tailed bat, Tadarida brasiliensis. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 193(8):853–863
12. Pollak GD, Winer JA, O'Neill WE (1995) Perspectives on the functional organization of the mammalian auditory system: why bats are good models. In: Popper AN, Fay RR (eds) Hearing by bats, Springer handbook of auditory research. Springer, New York, pp 481–498
13. Winer JA, Larue DT, Pollak GD (1995) GABA and glycine in the central auditory system of the mustache bat: structural substrates for inhibitory neuronal organization. J Comp Neurol 355(3):317–353
14. Pollak GD, Casseday JH (1986) The neural basis of echolocation in bats. Springer, New York
15. Pollak GD, Gittelman JX, Li N, Xie R (2011) Inhibitory projections from the ventral nucleus of the lateral lemniscus and superior paraolivary nucleus create directional selectivity of frequency modulations in the inferior colliculus: a comparison of bats with other mammals. Hear Res 273(1–2):134–144
16. Zook JM, Winer JA, Pollak GD, Bodenhamer RD (1985) Topology of the central nucleus of the mustache bat's inferior colliculus: correlation of single unit properties and neuronal architecture. J Comp Neurol 231(4):530–546
17. Casseday JH, Fremouw T, Covey E (2002) The inferior colliculus: a hub for the central auditory system. In: Oertel D, Popper AN, Fay RR (eds) Integrative functions in the mammalian auditory pathway. Springer, New York, pp 238–318
18. Oliver DL, Huerta MF (1992) Inferior and superior colliculi. In: Webster DB, Popper AN, Fay RR (eds) The mammalian auditory system: neuroanatomy. Springer, New York, pp 168–221
19. Pollak GD, Xie R, Gittelman JX, Andoni S, Li N (2011) The dominance of inhibition in the inferior colliculus. Hear Res 274(1–2):27–39
20. Li N, Gittelman JX, Pollak GD (2010) Intracellular recordings reveal novel features of neurons that code interaural intensity disparities in the inferior colliculus. J Neurosci 30(43):14573–14584
21. Fremouw T, Faure PA, Casseday JH, Covey E (2005) Duration selectivity of neurons in the inferior colliculus of the big brown bat: tolerance to changes in sound level. J Neurophysiol 94(3):1869–1878

22. Sanchez JT, Gans D, Wenstrup JJ (2007) Contribution of NMDA and AMPA receptors to temporal patterning of auditory responses in the inferior colliculus. J Neurosci 27(8):1954–1963
23. French B, Lollar A (2000) Communication among Mexican free-tailed bats. Bats Bat Conserv Int 18(2):1–4
24. French B, Lollar A (1998) Observations on the reproductive behavior of captive *Tadarida brasiliensis mexicana* (Chiroptera:Molossidae). Southwest Nat 43:484–490
25. Gelfand DL, McCracken GF (1986) Individual variation in the isolated calls of Mexican free-tailed bat pups. Anim Behav 34:1078–1086
26. McCracken GF (1984) Communal nursing in Mexican free-tailed bat maternity colonies. Science 223:1090–1091
27. Catchpole CK, Slater PJB (1995) Bird song: biological themes and variations. Cambridge University Press, Cambridge
28. Payne RS, McVay S (1971) Songs of humpback whales. Science 173(3997):585–597
29. Marler P (2004) Science and birdsong: the good old days. In: Marler P, Slabbekoorn H (eds) Nature's music: the science of birdsong. Elsevier/Academic, Amsterdam
30. Balaban E (1988) Bird song syntax: learned intraspecific variation is meaningful. Proc Natl Acad Sci U S A 85(10):3657–3660
31. Pollak GD (2011) Discriminating among complex signals: the roles of inhibition for creating response selectivities. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 197(5): 625–640
32. Portfors CV, Roberts PD, Jonson K (2009) Over-representation of species-specific vocalizations in the awake mouse inferior colliculus. Neuroscience 162(2):486–500
33. Holmstrom L, Roberts PD, Portfors CV (2007) Responses to social vocalizations in the inferior colliculus of the mustached bat are influenced by secondary tuning curves. J Neurophysiol 98(6):3461–3472
34. Suta D, Kvasnak E, Popelar J, Syka J (2003) Representation of species-specific vocalizations in the inferior colliculus of the guinea pig. J Neurophysiol 90(6):3794–3808
35. Klug A, Bauer EE, Hanson JT, Hurley L, Meitzen J, Pollak GD (2002) Response selectivity for species-specific calls in the inferior colliculus of Mexican free-tailed bats is generated by inhibition. J Neurophysiol 88(4):1941–1954
36. Xie R, Meitzen J, Pollak GD (2005) Differing roles of inhibition in hierarchical processing of species-specific calls in auditory brainstem nuclei. J Neurophysiol 94(6):4019–4037
37. Andoni S, Li N, Pollak GD (2007) Spectrotemporal receptive fields in the inferior colliculus revealing selectivity for spectral motion in conspecific vocalizations. J Neurosci 27(18):4882–4893
38. Portfors CV (2004) Combination sensitivity and processing of communication calls in the inferior colliculus of the Moustached bat Pteronotus parnellii. An Acad Bras Cienc 76(2): 253–257
39. Portfors CV, Mayko ZM, Jonson K, Cha GF, Roberts PD (2011) Spatial organization of receptive fields in the auditory midbrain of awake mouse. Neuroscience 193:429–439
40. Kowalski N, Depireux DA, Shamma SA (1996) Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra. J Neurophysiol 76(5):3503–3523
41. Kowalski N, Depireux DA, Shamma SA (1996) Analysis of dynamic spectra in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary dynamic spectra. J Neurophysiol 76(5):3524–3534
42. Klein DJ, Depireux DA, Simon JZ, Shamma SA (2000) Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design. J Comput Neurosci 9(1):85–111
43. Depireux DA, Simon JZ, Klein DJ, Shamma SA (2001) Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. J Neurophysiol 85(3): 1220–1234
44. Escabi MA, Schreiner CE (2002) Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain. J Neurosci 22(10):4114–4131

45. Fuzessery ZM, Hall JC (1996) Role of GABA in shaping frequency tuning and creating FM sweep selectivity in the inferior colliculus. J Neurophysiol 76(2):1059–1073
46. Koch U, Grothe B (1998) GABAergic and glycinergic inhibition sharpens tuning for frequency modulations in the inferior colliculus of the big brown bat. J Neurophysiol 80(1):71–82
47. Casseday JH, Covey E, Grothe B (1997) Neural selectivity and tuning for sinusoidal frequency modulations in the inferior colliculus of the big brown bat, Eptesicus fuscus. J Neurophysiol 77(3):1595–1605
48. Andoni S, Pollak GD (2011) Selectivity for spectral motion as a neural computation for encoding natural communication signals in bat inferior colliculus. J Neurosci 31(46):16529–16540
49. Razak KA, Fuzessery ZM (2006) Neural mechanisms underlying selectivity for the rate and direction of frequency-modulated sweeps in the auditory cortex of the pallid bat. J Neurophysiol 96(3):1303–1319
50. Suga N (1965) Analysis of frequency modulated sounds by auditory neurons of echolocating bats. J Physiol 179:26–53
51. Voytenko SV, Galazyuk AV (2007) Intracellular recording reveals temporal integration in inferior colliculus neurons of awake bats. J Neurophysiol 97(2):1368–1378
52. Xie R, Gittelman JX, Li N, Pollak GD (2008) Whole cell recordings of intrinsic properties and sound-evoked responses from the inferior colliculus. Neuroscience 154(1):245–256
53. Rust NC, Schwartz O, Movshon JA, Simoncelli EP (2005) Spatiotemporal elements of macaque v1 receptive fields. Neuron 46(6):945–956
54. Simoncelli EP, Pillow J, Paninski L, Schwartz O, Gazzaniga M (2004) Characterization of neural responses with stochastic stimuli. In: Gazzaniga M (ed) The cognitive neurosciences. MIT Press, Cambridge, MA
55. Woolley SM, Gill PR, Theunissen FE (2006) Stimulus-dependent auditory tuning results in synchronous population coding of vocalizations in the songbird midbrain. J Neurosci 26(9):2499–2512
56. 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(11):3374–3386
57. Woolley SM, Casseday JH (2005) Processing of modulated sounds in the zebra finch auditory midbrain: responses to noise, frequency sweeps, and sinusoidal amplitude modulations. J Neurophysiol 94(2):1143–1157
58. Gittelman JX, Li N, Pollak GD (2009) Mechanisms underlying directional selectivity for frequency modulated sweeps in the inferior colliculus revealed by in vivo whole-cell recordings. J Neurosci 29:13030–13041
59. Gittelman JX, Li N (2010) FM velocity selectivity in the inferior colliculus is inherited from velocity-selective inputs and enhanced by spike threshold. J Neurophysiol 106(5):2399–2414
60. Gittelman JX, Pollak GD (2011) It's about time: how input timing is used and not used to create emergent properties in the auditory system. J Neurosci 31(7):2576–2583
61. Xie R, Gittelman JX, Pollak GD (2007) Rethinking tuning: in vivo whole-cell recordings of the inferior colliculus in awake bats. J Neurosci 27(35):9469–9481
62. Fuzessery ZM, Richardson MD, Coburn MS (2006) Neural mechanisms underlying selectivity for the rate and direction of frequency-modulated sweeps in the inferior colliculus of the pallid bat. J Neurophysiol 96(3):1320–1336
63. Suga N, Schlegel P (1973) Coding and processing in the auditory systems of FM-signal-producing bats. J Acoust Soc Am 54(1):174–190
64. Suga N (1973) Feature extraction in the auditory system of bats. In: Moller AR (ed) Basic mechanisms in hearing. Academic, New York, pp 675–744
65. Brimijoin WO, O'Neill WE (2005) On the prediction of sweep rate and directional selectivity for FM sounds from two-tone interactions in the inferior colliculus. Hear Res 210(1–2):63–79
66. Covey E, Casseday JH (1999) Timing in the auditory system of the bat. Annu Rev Physiol 61:457–476

67. Yue Q, Casseday JH, Covey E (2007) Response properties and location of neurons selective for sinusoidal frequency modulations in the inferior colliculus of the big brown bat. J Neurophysiol 98(3):1364–1373
68. Ye CQ, Poo MM, Dan Y, Zhang XH (2010) Synaptic mechanisms of direction selectivity in primary auditory cortex. J Neurosci 30(5):1861–1868
69. Razak KA, Fuzessery ZM (2009) GABA shapes selectivity for the rate and direction of frequency modulated sweeps in the auditory cortex. J Neurophysiol 102:1366–1378
70. Zhang LI, Tan AY, Schreiner CE, Merzenich MM (2003) Topography and synaptic shaping of direction selectivity in primary auditory cortex. Nature 424(6945):201–205
71. Loftus WC, Bishop DC, Oliver DL (2010) Differential patterns of inputs create functional zones in central nucleus of inferior colliculus. J Neurosci 30(40):13396–13408

Chapter 9
Language Parallels in New World Primates

Charles T. Snowdon

Abstract New World monkeys are less well known and less frequently studied than Old World monkeys and apes, yet they have value as models for speech and language. All New World primates are arboreal living in dense forests. This has led to a reliance on vocal communication rather than the visual signaling most common in Old World primates and apes. As a result New World primates have evolved complex vocal repertoires that may share more parallels with human speech than other primates. Many of these species also live in small family groups akin to most human societies. I describe the rationale for studying speech and language parallels in New World primates and present some of the methods used in research. I then discuss several areas of potential parallels including vocal complexity, categorical responses to calls, signals that refer to external objects or events, syntax and rudimentary grammar, developmental processes including babbling, dialects, and vocal control. I conclude with a brief discussion of cognitive abilities in these species some of which have important parallels to human cognition. Although New World primates have not often been subjects of research relating to speech and language disorders, they provide much potential for understanding mechanisms and developmental and functional aspects of speech and language.

Keywords New World primates • Vocal complexity • Referential signals • Syntax • Babbling • Vocal control • Cognition

C.T. Snowdon, Ph.D. (✉)
Department of Psychology, University of Wisconsin, 1202 West Johnson Street, Madison, WI 53706, USA
e-mail: snowdon@wisc.edu

S.A. Helekar (ed.), *Animal Models of Speech and Language Disorders*,
DOI 10.1007/978-1-4614-8400-4_9,

What Are New World Primates?

The New World primates include all the nonhuman primate species found in the Americas. They are classified into three separate families: Callitrichidae (cooperatively breeding marmosets and tamarins), Atelidae (howler monkeys, spider monkeys, and muriquis), and Cebidae (capuchin monkeys, squirrel monkeys, and titi monkeys). All New World primates are arboreal and their diets range from primarily leaf and fruit eaters to omnivores eating a variety of plant and animal foods. They have a variety of social systems ranging from cooperative breeding, to monogamy, to large groups of mixed ages and sexes. The majority of research on communication has been carried out with marmosets, tamarins, capuchin monkeys, and squirrel monkeys, and these will be the main focus in this chapter although I will discuss a few interesting studies on other species.

Why Are New World Primates Interesting for Understanding Language?

As arboreal primates living in heavily forested areas, the New World primates as a group are much more vocal and have generally demonstrated much greater vocal complexity than terrestrial primates from the Old World. They are considerably more vocal than our nearest evolutionary ancestors, the great apes (chimpanzees, bonobos, gorillas, orangutans). Since human language is based on vocal production, highly vocal primates may prove to be better models than species closer to us in evolution.

Traditionally we think about evolution in terms of species diverging from one another over time. The shorter the time since two species diverged and the fewer branches between species, the more shared traits are to be expected. By this logic, our two closest ancestors are chimpanzees and bonobos, and one might expect these species to be the best models for human language. However, although chimpanzees and bonobos have many impressive cognitive abilities and expressive faces and other visual gestures, what we currently know of their vocal communication suggests they may not be the best models for speech or language although recent work suggests greater communicative complexity in chimpanzees [1].

An alternative type of evolution is converging evolution—when two very different species have similar adaptive problems and converge on a common solution. Peter Marler proposed in an influential paper [2] that birdsong would be a good model for understanding speech development because songbirds were much more dependent on vocal communication than many nonhuman primates. We can make a similar argument that the pressures of living in an arboreal environment (where visual signals will not be very effective and gesturing can lead to falling out of a tree) have made vocal communication more adaptive for New World primates and that these species may provide as at least as good models as great apes and other monkeys, for the vocal aspects of speech and language.

However, there has been relatively little work on using New World primates as explicit models for speech and language disorders, so in this chapter I will focus first on research methods then on several phenomena that are suggestive of parallels with speech and language and will conclude with some suggestions for how these monkeys might be used as potential models for research on speech and language.

Methods of Study

What are the best methods for studying vocal communication in New World primates? There are benefits to both field and captive research. Field studies can tell us much about the ecological basis of communication and help us understand the functions of communication, but captive studies can allow us to do experiments to test hypotheses and to study neural mechanisms of communication more readily. In the field many primate populations are not well habituated to the presence of human observers, and thus, scientists are likely to be making biased observations of the most obvious signals (those used in aggression or alerting others to predators) at the expense of more subtle, less obvious signals that might be used to communicate about subtle social relationships. Only a few researchers have been able to habituate wild populations enough so that subtle aspects of social communication may be studied. In contrast captive populations are much more likely to be habituated to human observers and allow recordings at closer distances. With the ability to identify known individuals, studies in captive environments may be able to identify some functions of communication that would be difficult or impossible to observe in the wild. Furthermore, some field researchers have concerns that experimental studies could have unintended consequences in wild populations. Thus, for example, playbacks of calls of strange individuals may lead to increased aggression by the study population toward its neighbors.

At the same time captive studies may lack ecological validity if animals are not housed in species-appropriate social groups or if the captive environment is not complex enough to allow normal expression of typical signals or normal expression of behavioral responses to signals. An ideal research strategy would meld field and captive observations to provide a full understanding of how signals are used and in what contexts as well as allowing for experimentation that does not affect the behavior of wild populations. In my own work on pygmy marmosets (the world's smallest monkey, found in the Western Amazon), I found that in complex captive environments, I could observe all of the vocalizations that I observed in the field, but recording in the field gave me a greater appreciation for the complexity of the contexts in which signals were used.

Although much research is focused on testing hypotheses, when we are studying the communication signals of other species, it is necessary to engage first in a stage of natural history research before hypotheses can be formed, let alone tested. The most basic stage involves making high-quality recordings of vocalizations and making careful behavioral observations at the same time about the context of the calling.

Context can include the general behavioral state, feeding, resting, socializing, and responding to predators or strange conspecifics, but context also may include the age, sex, and reproductive status of the communicator. Modern digital computers allow for highly detailed analysis of vocal signals using a variety of different analytical programs. Among the most common is the use of fast Fourier transfer functions that decompose a vocal signal into all of the component frequencies and intensities that can then be displayed visually, allowing for precise measurement of intensity, duration, and frequency patterns of call.

Since the function of communication is based on how the listeners respond as well, a careful researcher will also note the response given to a signal by others. Because it is much easier to observe the caller and its behavior (or the speaker and the content of speech), research on animal communication as well as on human language has focused much more often on the signal and the behavior of the communicator than the effects of the signals on the behavior of the listeners.

Once the stage of natural history is well underway, one can begin to develop hypotheses about function that can then lead to hypothesis testing. There are two main forms of hypothesis testing—observations and experiments. Many hypotheses can be tested by making predictions about under what contexts the signal will be produced and what listeners will do in response. These can be tested by careful observations of what behavior would elicit a call or how an animal should respond to a call if the hypotheses are correct. As an example, we tested whether our hypotheses about the usage of a series of very similar calls in cotton-top tamarins were correct by creating the contexts in which each call was predicted to occur and finding that the animals indeed produced the appropriate call variant in the predicted context [3].

Another important method is recording calls and playing them back when none of the appropriate context is present to determine if the sound of the call alone is sufficient to produce the predicted behavioral response. This playback method has the advantage of eliminating all other cues except the call itself and is thus a very powerful method of experimentation. As one example, we have played back similar, but subtlety different, calls to cotton-top tamarins to see if they could discriminate them, and we found a clear difference in how the monkeys responded to each form indicating that they perceived differences in the calls [4]. As another example, we played back the calls of individual pygmy marmosets from their own cages (an expected location) versus from other locations in the colony room and found that other monkeys responded most to playbacks of individuals coming from expected locations, indicating an ability to discriminate individuals on the basis of vocal cues alone [5].

Language-Like Phenomena

Let me now examine several phenomena of vocal communication in New World monkeys that suggest some parallels with speech and language. These include vocal complexity, categorical labeling of calls, ability to refer to objects or events external

to the caller, rudimentary syntactical rules, vocal development, dialects, and vocal control. After reviewing each of these, I will briefly consider some of the cognitive abilities of New World primates that suggest more complex cognitive abilities among these monkeys than has been generally accepted.

Vocal Complexity

One characteristic of human language is the large repertoire of sounds we use in speech. Many have thought that nonhuman animal had a much more limited repertoire, but we have observed subtle vocal variants among monkeys that suggest a more complex vocal repertoire. The cotton-top tamarin produces a large number of chirp-like calls (high-pitched, short calls with extensive frequency modulation) throughout the day. When we made careful spectrographic measurements of these calls, we could identify 8 variants of chirps. One appeared in mobbing contexts, another in response to hearing strangers, a third served as an alarm call, two others were used in feeding contexts, and another variant was used for affiliation between group members. Thus, what we initially heard as a single-call type actually consisted of several discrete calls that were used in very different contexts [6].

We also found similar complexity in trill vocalizations of pygmy marmosets. Trills are high-pitched, frequency-modulated calls. We identified four variants of trills. One was soft, short, and had a small frequency range. Another form was the same duration but had a larger frequency range, a third was similar in frequency range but had a longer duration, and the fourth consisted of a series of discrete notes that appeared to be trill-like, but with the trill being interrupted with discrete pulses. In captivity we readily noted that the long version of the complete trill was usually given in aggressive contexts with the mouth open, but we could not distinguish the contexts in which the other trill variants were given [7].

However, based on principles of sound localization, we could make some predictions about how these calls would be used in the wild. The short trill with a low degree of frequency modulation would be the most cryptic, whereas the long interrupted trill provided the most cues for sound localization and would be easily located over a longer distance. Since vocal communication represents a balance between making calls obvious to recipients but minimizing the likelihood that potential predators could locate the caller, we predicted that the most cryptic form of the trill would be given when animals were close to one another, that the trill with a greater frequency range would be used when animals were further apart and the interrupted trill would be used when animals were furthest apart. We went to the Peruvian Amazon and recorded many calls of each type and then calculated the distance between the caller and the nearest animal we could detect. We found a clear relationship between call type and the distance between animals. When an animal gave the most cryptic trill, we could almost always find another animal within 5–10 m, whereas when a monkey gave the interrupted trill, we generally found the nearest animal more than 20 m away. Thus, the monkeys appear to monitor the

location of one another and adjust the call structure to be cryptic when others are close by, but to take more risks and give more obvious calls when other group members are far apart [8].

We have replicated these findings with several other groups in Ecuador and have also broadcast samples of each call type through the forest and rerecorded these calls at different distances. We found that the more subtle calls are rapidly degraded as they pass through the vegetation with decreasing high frequencies being recorded and an increase in reverberation with increasing distance. It was difficult for us to detect the most cryptic call beyond 20 m and hard to detect the most obvious call at 40 m [9]. Thus, marmosets have evolved signal structures to minimize detection by predators and are able to select call structures that adjust for the distance between them and their group members.

Marmosets, tamarins, and other species are able to increase the complexity of their vocal repertoires by combining different signals into sequences, and I will say more about this later.

Categorical Responding to Signals

One major finding of human speech is our ability to categorize variations in phonemes into discrete classes. Thus, if we present a series of synthesized syllables ranging along the voice onset continuum from /ba/ to /pa/, human subjects will not hear each syllable as separate but instead will classify several as /ba/, and then there is a sharp boundary after which all remaining syllables on the continuum will be heard as /pa/. For a long time this categorical perception was thought to be uniquely human, but studies on macaques and chinchillas in the 1970s showed that other species could categorize human speech as well. But do animals show a similar classification of their own calls?

We were able to synthesize trill vocalizations of pygmy marmosets, and we varied each parameter separately—frequency range, rate of frequency modulation, center frequency, and duration. Remember that the length of a trill appeared to determine whether it would be used in an affiliative or aggressive context. When we played back synthesized calls to our marmosets, we found little influence of any variable other than duration. However, with duration we found a pattern similar to that seen in humans; all calls up to a certain duration were responded to with antiphonal calls, and with a mere 8 ms longer duration, all subsequent calls were ignored. The break in the response distribution occurred just at the upper limit of durations of the affiliative form of the call [10].

Although it is interesting to see parallels in the type of perception between marmosets and humans, categorical perception is also a bit strange from the perspective of social interactions. When organisms interact with one another, they need to understand the signal, but they often need to be able to identify the caller as well. Thus, a categorical perceptual system that treats all variants of a phoneme or a call as equivalent does not make much sense in the social context of communication.

We subsequently synthesized trills representing not the average parameters of the population, but instead the features of each individual, and we then played these back to our monkeys. We now found a very different response pattern. Instead of showing a broad perceptual categorization, monkeys now showed very narrow categories that corresponded to the range of variation found within each familiar individual. Thus, animal A may have only a 50 ms range of trill duration, and listeners familiar with A's calls identified A only within the narrower range of parameters that defined individual A's calls. This leads to a reformulation of perception of phonemes or calls. With unfamiliar individuals or in situations where the costs of failing to recognize a call may be high (predation or other danger), the categorical system will take priority, but in situations where it is important to recognize the individual due to the social nature of most communication, then a second perceptual system takes over to determine which individual is calling [11].

I am unaware of any parallel studies on humans using speech sounds. However, related findings from other species suggest that having multiple perceptual criteria for signals based on social context may be common. Thus, baboons respond to calls based on the relationship between rank and kinship of caller and recipient [12], and suricates (a social mongoose) encode both predator type and degree of urgency of response in their alarm calls [13].

Referential Signals

Referential signals are those that appear to communicate more than the emotional state of the caller and appear to signify objects or events in the environment. They might be construed as rudimentary forms of words, especially if there is no obvious connection between the structure of the call and what is being referenced. The clearest and best known examples of these types of calls have been described in Old World primates and, oddly, in ground squirrels and chickens. Both predator-specific alarm calls and food calls have been described, with Diana monkeys and vervet monkeys in Africa having a suite of predator alarm calls that appear to be specific to predator type (snakes, eagles, leopards) and chickens and ground squirrels having distinct calls for aerial versus terrestrial predators. Chimpanzees, macaques, and chickens also have specific calls that appear to be related to the discovery of food. However, in some case there have been alternative interpretations. For example, although ground squirrels have distinct calls for aerial versus terrestrial predators, these calls appear to relate to urgency of response rather than to specific predators. Thus, a hawk that is flying rapidly overhead and not searching for prey elicits the "terrestrial" alarm calls, whereas if a dog gets very close to a squirrel without being spotted, the squirrel gives the "aerial" alarm call. Rather than being specific to predator type, the calls appear to communicate urgency of danger instead [14].

Among New World primates food-associated calls have been observed in a broad range of species. Both cotton-top tamarins and golden lion tamarins produce calls that are highly specific to the presence of food [15–17]. In adults these calls are

rarely given to nonfood objects. In both species the rate of calling is proportional to the quality of food as perceived by the individual calling. Thus, each animal might have its own preference for a variety of different foods, and rate of calling by an individual was directly correlated with that individual's preference. Benz [17] reported that there were also significant differences in the structure of food-associated calls across individuals in golden lion tamarins with distinct calls for protein, dried fruit, and fresh grapes. This is a degree of specificity of food labeling not seen in any other species.

Field studies of two species of capuchin monkeys have also reported food-associated vocalizations. In one study DiBitetti [18] reported that tufted capuchins gave two different types of food calls at a much higher rate to clumped sources of fruit compared to distributed fruit and to less preferred types of foods. With playbacks of food calls, but not of other call types, monkeys approached the speaker broadcasting the calls suggesting that food calls attract other monkeys. Furthermore, capuchin monkeys gave more calls to large amounts of food and called more quickly when more individuals were nearby [19]. In a study of white-faced capuchin monkeys, Gros-Louis [20] reported that monkeys called more often to fruit than to eggs or insects. (Interestingly, capuchin monkeys rarely called when finding live prey even through live prey is highly preferred, suggesting that they can inhibit their calls with prey that might be able to detect the predator.) Capuchin monkeys also encode information about sex and individual identity in their call structure [21]. In white-fronted capuchins, monkeys were more likely to call when approached by a higher-ranking individual, and monkeys who failed to call were more likely to receive aggression from other group members [20]. In a study of captive capuchin monkeys, Pollick et al. [22] reported that animals called more often for large amounts of food and also called more when other animals were present than when either one animal or no animal was present. Thus, capuchin monkeys appear to be sensitive in their food calling to the rank, proximity, and number of other group members present. This contrasts with cotton-top tamarins that did not call more often when the mate was present or absent [23].

A study of spider monkeys [24] reported that food calls were given more often when large fruiting trees with abundant fruit were discovered and more often when dominant animals were present than when only subordinate animals were present. Thus, in all species studied, there are one or more calls that appear to be specifically associated with the presence of food. In most species the call rate is proportional to the perceived quality of the food with some indication that monkeys can inhibit calling in the presence of live prey. In only one study did call structure appear to be specific to food type, and these different structures may also indicate preference rather than denoting specific types of foods. Finally, callers in many species appear to be sensitive to the presence or absence, and/or the sex and dominance status of other group members in whether they call or not suggesting that social variables play an important role in calling.

The most impressive examples of referential signals have come from the studies of highly specific predator alarm calls, discussed at the beginning of this section. Much less work has been done on alarm calls in New World monkeys, but two studies on white-faced capuchin monkeys and another in tufted capuchin monkeys have

found generally similar results [25–27]. Capuchin monkeys have two types of alarm calls, one of which appears to be specific to aerial predators and elicits immediate descent from trees. The other type was given to a wide range of terrestrial predators and did not elicit a consistent response from listeners. Responses ranged from freezing to approaching the caller and engaging in mobbing behavior. A recent study of black-fronted titi monkeys also found different call types given to different predators, and when calls were played back, the monkeys looked upwards toward aerial predator calls and downward when hearing terrestrial predator alarm calls [28]. A study of two sympatric species of tamarins in the Peruvian Amazon [29] found that tamarins also have aerial and terrestrial predator alarm calls. In playback studies members of both species responded in an appropriate way to the alarm calls of either species—looking upwards when an aerial alarm was broadcast and looking at the ground when a terrestrial alarm was broadcast. However, New World primates do not appear to have the well-differentiated set of predator-specific calls that have been seen in several Old World species.

Grammar or Syntax

Syntax or the sequencing of calls in a standard order is one of the hallmarks of language, but some of the more impressive examples of primate syntax appear in New World monkeys. One of the first field studies to show a form of grammar was on titi monkeys, small primates that form monogamous pair bonds where both sexes engage in duetting behavior. In one species of titi monkey, there were sequences that involved four different call types that appeared in a fixed order [30]. When these calls were played back to wild groups in their normal order and in an altered sequence, the responders showed "disordered" behavior is response to the scrambled sequence compared to the natural sequence. There are many possible reasons for this response including the novelty of hearing a scrambled sequence for the first time, but the results are suggestive of some type of syntactical rules for sequencing calls.

The cotton-top tamarin has been shown in several captive studies to have syntactic rules governing its own vocal signals as well as being able to identify sequential relationships in sequences of human phonemes. The vocal repertoire of cotton-top tamarins has a large number of sequential signals [6]. There are several forms of long calls, multi-unit whistle-like vocalizations with a rising intonation often preceded by short chirp-like calls. No one has ever observed chirps appearing between successive whistles or at the end of whistle calls. There are other calls that are produced both separately and in combinations. One example is the combination of an alarm call with a contact call. These sequences are always produced after an alarming even when all animals have been quiet and appear to serve as an "all-clear" signal. Another call sequence is given during territorial interactions between adjacent groups where a chirp-like call specific to threat is given by one member of a pair and a long call is given by the other member of the pair during the early stages of an interaction and with both pair mates combining the calls together at the peak of confrontation [31].

Statistical learning is a recent discovery showing that human infants can readily perceive the statistical properties of syllables that co-occur with different probabilities. Several years ago it was discovered that young infants can learn words through a simple process of statistical learning. Infants as young as 6 months were played a 2 min sequence of nonsense syllables, with different statistical patterns. Some triads of syllables occurred with a probability of 70 %, others with a probability of 50 %, and others with no co-occurrence. When infants were tested after hearing this 2-min sequence of sounds, they were able to recognize the patterns of sounds that had been presented with probabilities of 50 and 70 % suggesting that infants can learn to recognize words through identifying statistical regularities in patterns of syllables [32]. This result has important implications for how infants learn to segment the speech stream that they hear and thus to learn how to differentiate words. It turns out that cotton-top tamarins can also detect similar sequential probabilities in sequences of human phonemes. When presented with the same stimuli as presented to infants, cotton-top tamarins also identified sequences in the same way [33]. In successive studies, three types of nonadjacent dependencies were created: (1) syllables one and three occurred together but the intervening syllable varied, (2) first and last consonants were predictable but the intervening vowel differed, or (3) the first and last vowels were predictable but the intervening consonant differed. Adult human subjects readily learned the second and third types of regularities but not the first type. Tamarins on the other hand learned the first and third types but not the second type [34]. However, when tamarins and infants were tested on a series of tasks where they had to extract increasingly more complex grammatical relationships based on statistical patterns, the tamarins were only successful with the simplest grammatical sequences [35]. Although humans and tamarins differed in some important ways in how they responded, it is remarkable that another species is able to extract any regularities from human speech sounds. Thus, statistical association of sound patterns appears to be an important mechanism for language learning, and similar processes occur in other species.

One other type of sequencing can be seen in turn-taking behavior. In pygmy marmosets animals exchange trill vocalization antiphonally, and when the identity of each caller is possible, there is a significantly higher probability of each animal calling in sequence than of an animal calling before all others have called. In addition the particular sequence of turn taking within a group appears to be consistent [36].

In both their natural vocalizations and in their ability to respond to sequences of human sounds, New World primates provide good models for understanding syntax and segmentation of sound patterns.

Development

One of the most puzzling aspects of research on primate communication has been the apparent absence of evidence of vocal learning, a key component in both human speech and birdsong. Aspects of communication can be divided into three separate

categories, production of appropriate vocalizations, perception of calls, and usage of calls in appropriate context. Early research on squirrel monkeys found little evidence of flexibility in any aspect of communication. Deafened monkeys appeared to acquire the normal sounds of their species [37], and squirrel monkeys reared in social isolation were said to produce the appropriate calls and show appropriate responses after their first exposure to threat contexts [38, 39], but experience was required to associate the alarm call with the appropriate context. Another study [40] found that infants did not show adult-like responses to alarm peeps given by others until they were nearly a year old. Furthermore, adults responded more promptly to playbacks of adult alarms than to infant alarms suggesting that there is also a maturation of the production of peeps with increasing age. Recordings of squirrel monkey isolation peeps showed maturational changes in call duration, but the basic structure of the call was present at birth [41]. More recent research across a wide range of primate species has shown that although there is little direct evidence of vocal learning in the sense of acquiring a totally arbitrary new signal, there is considerable evidence that monkeys can modify the structure of their calls within a constrained range. There is also strong evidence that monkeys must learn how to respond to calls of their own and other species and must learn which calls to give in specific contexts. Thus, a modern view of primate communication suggests a major role for learning through development.

A striking feature of human infants is their babbling behavior [42]. Marler [2] claimed that the development of birdsong showed parallels to babbling behavior since male birds undergo practice phases (plastic song and subsong) shortly before the end of their first year of life as they begin to practice adult song. However, there are serious problems with using birdsong as a model for human babbling. In temperate-zone birds, it is primarily males that sing and song is just one of many types of vocalizations in the repertoire of a bird with song functioning as a sexually attractive signal. Furthermore, the phases of plastic song and subsong that Marler identified do not occur early in life but appear during puberty. This would be similar to saying that the clumsy effort of adolescent boys to converse with girls is the primary source of developing language skills!

A more realistic model of babbling to study language development would need to look at a much earlier developmental period, to include both sexes and include a greater part of the adult vocal repertoire than sexual signals. The pygmy marmoset provides just such a model. We noted that newborn marmosets began to make long sequences of vocalizations within the first 2 weeks of life and these long vocal sequences continued until after weaning had occurred. Just as human infants produce many of the phonemes of adult language in their babbling, so do pygmy marmoset infants produce many of the call types of adults in their "babbling." Just as human infants repeat and juxtapose a variety of phonemes together, so do pygmy marmosets show repetition of a call type and juxtapose functionally different calls (from an adult perspective). Thus, two or three food calls might be followed immediately by a few alarm calls, followed by some affiliative calls followed by some threats. Just as human caregivers respond positively to a babbling infant, so do pygmy marmoset parents show increased affiliative behavior with babbling infants compared

with non-babbling infants [43]. We have observed this babbling behavior in both captive and wild marmosets.

Babbling is puzzling in terms of its putative function for marmosets. What benefits can infants gain from calling attention to themselves? One possible explanation is that a vigorously babbling infant communicates to its parents that it is healthy enough to deserve good parental care. As noted above, babbling infants are more likely to be recipients of increased affiliation from caregivers. A second possibility is that babbling is a form of vocal practice that leads to adult forms of calling more rapidly. We have found that infants that babble more frequently and with greater diversity of calls in the first month of life have more adult-like vocal structures at 5 months of age [44]. Thus, babbling may lead to more rapid vocal development.

Research with human infants [45] has demonstrated parallels in babbling. Some mothers were told to react to the sounds of their infants contingently (i.e., by interacting socially with their infants after babbling), whereas other mothers could not hear their infants and could not respond contingently. The infants whose mothers responded contingently to their babbling showed more rapid development of phonemes and at a later age more rapid development of words [46] than infants of mothers that did not respond contingently even within the short time frame of an experimental manipulation. These findings suggest that the babbling of marmosets may be a more relevant model for language development than birds.

Tamarins do not show the babbling behavior of marmosets; however, they also appear to learn both call structures and appropriate usage. When we tested tamarins by experimentally creating contexts in which we expected chirps to occur, we found that adults always gave the structure of trill appropriate for the context. However, infant tamarins when tested did not show any obvious differentiation of chirp structure. Instead infants typically gave a sequence of several trills with descending pitch in contrast to adults. Over the course of testing in the first 5 months of life, some, but never all, individuals occasionally gave adult-like chirps in the appropriate context, but having once given a chirp in a contextually appropriate context, there was a very low probability that the same animal would give that chirp again when tested in the same context [3]. Thus, tamarins differ considerably from squirrel monkeys. They rarely give appropriate types of calls in appropriate contexts and fail to produce adult-like structures over a 5-month period which would be the human equivalent of the first 7 years of life. We found a similar pattern with pygmy marmoset trills. We followed several animals longitudinally throughout the first year of life and then observed them when they were fully mature adults. The marmosets showed a progressive development of adult trill forms but prior to puberty they still did not have full control over vocalizations and produced trills with asymmetric structures, which were rare among adults [44]. Thus, both tamarins and marmosets show a rather slow developmental process in the production of species typical calls.

There are also clear developmental changes in the usage of calls. We tested cotton-top tamarins on their usage of food chirps from infancy through adulthood. In adults these calls are given almost exclusively to the presence of edible foods. We found that young tamarins produced imperfect forms of food calls. They also gave the same calls to inedible, small objects that could be manipulated like food

suggesting that they were overgeneralizing in their responses similar to young children. Furthermore, when presented with food, young tamarins also produced many other vocalizations that were not given by adults when food was present [47]. We expected that infant tamarins would improve call structure and use only food calls when feeding as they became older, but, to our surprise, young tamarins continued to use infantile forms of food calls, along with other vocalizations, and continued to overgeneralize even when they were postpuberty. This was puzzling at first, but in cooperative breeding species, older offspring play an important role in helping parents care for younger offspring and females, but not males, are reproductively suppressed. We hypothesized that postpubertal tamarins were showing their subordinate status by continuing to use infantile vocalizations. To test this hypothesis we recorded how tamarins responded to food before and immediately after they were removed from the group and paired with a mate of their own. Within a few days, they gave only food calls (and not the other call types) when tested with food, and within a few weeks, all animals were producing food calls with adult-like structure [48]. This suggests that young, postpubertal tamarins have the ability to produce adult-like calls in feeding contexts but are inhibited from producing these calls while living as subordinate helpers in a family group.

Dialects

The understanding of how dialects are acquired by birds has played an important role in understanding vocal learning. If a species demonstrates population or group-specific variation in vocal signals, it is unlikely that these differences are due to some innate genetic mechanism. Rather the presence of dialects suggests the likelihood that learning processes are involved. Dialects (or population differences) have been relatively rare among nonhuman primates, but some studies have suggested that chimpanzees have different dialects. Early studies found differences in the structure of pant-hoot vocalizations in widely separated chimpanzee populations in East Africa [49–51]. Recently studies of chimpanzees in the Tai Forest in West Africa have shown group differences in the structure of pant-hoot vocalizations of three adjacent populations and, on the assumption that there are few differences in habitat, suggested that these differences must be socially learned conventions to make group differences clear [52].

Studies of both pygmy marmosets and Wied's black-tufted-eared marmosets in captivity have identified a mechanism of vocal convergence that could account for group differences. In one study of pygmy marmosets, two colonies were combined into a single colony room where they could hear each other. Within a few weeks, animals of both groups showed convergence in trill structure with both groups increasing peak frequency and bandwidth of their trills [53]. In a second study of pygmy marmosets, trills were measured in several animals living in family groups prior to pairing, and then after new pairs were formed, vocal recording continued. Within 3 weeks of pairing, all pairs had converged on a pair-specific form of trill

[54]. When some pairs were monitored 3 years later, they still had convergent trills. In Wied's marmosets individual variation in phee call structure was studied over several weeks with the finding that individual structure was highly stable. However, when some animals were moved and housed adjacent to novel conspecifics, they significantly altered the structure of their phee calls [55]. Thus, marmosets can easily adjust their call structure in the presence of new social companions suggesting a high degree of vocal flexibility.

Returning now to dialects, wild pygmy marmosets recorded in the Ecuadorian Amazon demonstrated dialects or population-specific trill structure [56]. Five populations were spaced across an east–west transect of 300 km and a north–south transect of 100 km, and although there were clear individual and pair differences in call structure in each group, there were also clear vocal signatures for each population in two of the trill types. A discriminant analysis was able to correctly assign trill type to population with an accuracy of over 70 % compared to the chance rate of 20 %. The spectrum of ambient noise and of reverberation was measured in each of the five habitats, and while there were some differences in habitats with respect to sound transmission, these differences did not correspond to predicted effects on call structure if the variation in calls was purely determined by habitat acoustics. Interestingly, the same five populations also showed population-specific differences in tree species of preferred exudate foods that was independent of the relative availability of the species within each habitat [57]. The most parsimonious explanation for the population differences in trills is the social convergence hypothesis supported by the studies of captive marmosets.

Vocal Control

Tamarins are able to rapidly use calls in novel contexts, again deviating from the seeming innate vocal communication of squirrel monkeys. We were interested in how tamarins would react to a familiar food (tuna) that had been made aversive through the adding of invisible white pepper. Only about a third of the animals ever sampled the pepper-adulterated tuna, and once having sampled the peppered tuna, tamarins rarely sampled it again. There were clear signals given by the tamarins that sampled the tuna. There was a significant decrease in food calls after tamarins sampled the tuna, and notably, they showed a significant increase in alarm calls. We had never observed alarm calling in a feeding context before, and so the production if alarm calls on the first exposure to peppered tuna is an example of tamarins applying a call to a novel situation where the call served to keep others from sampling the food [58]. In a similar study with capuchin monkeys, there was no evidence of social learning to avoid a noxious food, but also no evidence of any communication signals to other group members [59].

We have been interested in how captive-born tamarins respond to predators and have completed two studies where we presented live boa constrictors (a natural predator) to tamarins [60, 61]. However, we found no evidence of fear and no alarm

or mobbing vocalizations to the snake. However, the captive monkeys readily gave alarm and mobbing calls to a caretaker dressed in distinctive clothing used to catch monkeys for veterinary care and to a feather duster used to clean light fixtures [62]. Thus, captive tamarins have adapted their threat and fear vocalizations to the ecology of captivity.

Another approach to determining vocal control is to present animals with a noisy background to see if they can adjust call structure in response to noise. Humans increase the amplitude and duration of speech sounds in a noisy environment and both common marmosets and cotton-top tamarins show increased amplitude and duration of calls in response to presentation of white noise [63, 64] showing that these species are also able to adjust call structures to cope with noise. In a related set of studies, short bursts of white noise were presented to cotton-top tamarins while they were producing long calls (sequences of several long notes). Rather than increasing amplitude and duration to these short bursts of noise, the noise actually interrupted the calls with the call ending at the syllable during which the noise burst had been presented [65, 66]. This was interpreted as evidence both of vocal control and that the tamarins organize their long calls in terms of individual syllables rather than having a single motor pattern for the entire call. This protocol might also serve as a possible model for stuttering. Taken together, the results from vocal control and the developmental changes show that for at least some New World primates adult call structure and usage is a developmental process involving both learning and social environment and that these monkeys have a remarkable degree of vocal control. This suggests that they have greater value in terms of studies of speech and language than many have previously thought.

Cognitive Skills

Language is not simply a complex vocal skill, but it is inseparable from cognitive skills. What cognitive skills do New World primates exhibit that may have relevance to language? There is emerging evidence that the cooperatively breeding marmosets and tamarins may have skills in social cognition that are more similar to those of humans than great apes. Although great apes have demonstrated impressive cognitive skills in many realms, learning to use symbols as equivalents of words, making and using tools, understanding complex relational tasks (equal to analogy problems), and more, in the social realm great apes appear remarkably dense, perhaps because chimpanzees, the species most commonly studied, are highly competitive and self-centered. Recently some have argued that human cognition is a product of two converging evolutionary strains—understanding physical relationships between objects and events derived from our great ape ancestors and social intelligence derived from cooperatively breeding monkeys [67].

Tamarins can readily learn to avoid a noxious food by watching others and responding to signals from those who have sampled the food (as noted above), yet other species of monkeys do not communicate to others about noxious foods.

Tamarins presented with a novel foraging task rapidly learn a new motor skill and learn which of five containers has food available within 2–4 trials through social learning, and they remembered these skills when tested 17 months later [68]. Marmosets can readily imitate the actions of another through observation [69].

The strongest evidence of teaching in nonhuman primates comes from studies of tamarins both in captivity and in the field. Tamarin adults begin sharing food with infants at the time of weaning by giving very rapid series of food calls. These calls attract infants who approach adults and are either offered food or can take food from the adult. Infants are rarely able to obtain food if the adult does not first produce the rapid food calls, and the rapid sequence of food calls is found only when adults are with infants [70]. Infant tamarins whose caregivers begin calling and food sharing at an earlier age more rapidly learn to forage on their own and also give adult-like food calls at an earlier age than those for whom food sharing starts at a later age. As the infants develop independent feeding skills, adults share food with decreasing frequency. When juvenile tamarins (who had been feeding independently for several months) were presented with a novel foraging apparatus in the presence of one of their parents who had been trained to forage successfully, the parent began giving rapid food calls again and proceeded to share food with the juvenile. However, as soon as the juvenile solved the new foraging task successfully, the parent ceased food calling and refused to share food with the infant [71]. When the same experiment was done with adults, there was no evidence of food calling and no food sharing was observed [72]. Taken together these studies suggest that adults have special vocalizations used in early stages of feeding in infants and that these calls reappear when a juvenile is confronted with a novel foraging task, but both the calls and the food sharing decrease as the animal demonstrates competence. Similar results have been seen in field studies of lion tamarins where adults give food calls to and share food with juveniles only with live prey, which is difficult for a juvenile to get on its own, and, similar to captive tamarins, adults withdraw support as young animals develop skills [73]. Sensitivity to the skills and knowledge of the learner and behavior scaffolding that change with the learner's competency is a hallmark of good teaching.

In contrast, parallel research on ant-dipping in wild chimpanzees failed to find any evidence of teaching behavior. Chimpanzees often forage on driver ants that are highly aggressive and produce severe bites. Adult chimpanzees are competent in using two types of tools and methods of capture, one for highly aggressive ant species close to their nest and the other with less aggressive ants away from the nest. Several mother infant pairs were observed over several years, and despite the consequences to the infants of being bitten by the ants, there were no signs that mothers helped infants make the right-sized tool or helped them to use the tool to minimize risk [74].

Chimpanzees rarely cooperate to help others, especially if food is visibly involved, but several studies have demonstrated that tamarins and marmosets are highly cooperative to the point of sharing food with others even without receiving food themselves. Tamarins were presented with an apparatus with two trays, one of which could be pulled from each end of the apparatus. The trays had springs so they

would return to the original position unless held out. Food was placed on the top tray such that when both trays were pulled simultaneously, the food dropped to the bottom of the apparatus where it could be retrieved. Two tamarins facing each other on the apparatus rapidly learned to pull the trays simultaneously to obtain food and continued to pull even though only one animal was rewarded each day [75]. Tamarins pulled at a higher rate both when both received food and when only one received food than did capuchin monkeys in a comparable study [76, 77].

When chimpanzees were tested in an apparatus where they could pull a tray to give themselves and a next-door neighbor food versus a tray giving food only to themselves, they appeared indifferent to the neighbor [78]. However, when marmosets were tested under more stringent conditions where they could pull a tray that would give food to a neighbor but get nothing themselves versus an empty tray, they provided food to the neighbor significantly more often than chance [79]. A modified version of study on tamarins found that tamarins also would donate food to a companion even though it received no food [80].

Chimpanzees are noted for their extensive use of tools and tool use varies among different populations in Africa. This variety in tool use across populations has lead to the description of chimpanzee culture [81]. Population variation in vocal signals and in feeding in pygmy marmosets was discussed earlier, but capuchin monkeys show traditions and tool use behavior similar to that of chimpanzees. Different groups of capuchin monkeys in Costa Rica have different foraging behaviors and social conventions [82]. In some populations in Brazil where there are seasonal decreases in food supply, capuchin monkeys have been observed to use stone tools to break open hard-shelled nuts much as chimpanzees do in West Africa [83, 84]. Thus, some New World monkeys demonstrate traditions, population variation in social signals, and other behaviors and have the capacity to use tools to obtain food. The cognitive skills involved in all of these phenomena have often been uniquely associated with humans so their presence in New World primates suggests cognitive parallels that may also be important for language use.

Conclusions

I have presented several research findings suggesting that many aspects of New World primates make them potentially interesting models for research on speech and language. These animals have not yet been used for research on speech and language disorders, but they can be of potential value. The common marmoset is increasingly being used on brain studies of communication and effects of cochlear implants [85, 86], and we have also used noninvasive fMRI to understand how olfactory signals are processed by marmosets [87]. All species place a great reliance on vocal communication, and in species that have been closely studied, there are complex vocal repertoires with the ability to combine discrete sounds into sequences that have a rudimentary syntax. Some species have evolved signals that refer to broad predator type and many have calls to indicate the discovery of food. In a very

basic way, these referential signals represent a rudimentary form of words. It has long been thought that learning plays a minor role in vocal development in nonhuman primates making birdsong a better parallel to human language, but birdsong is used mainly by males in a specific context of mate attraction and thus has limited functional parallel to language. However, the babbling behavior of young marmosets and the relatively slow process of acquiring adult repertoire coupled with the role of adult social reinforcement of infant vocalizations and the ability to utilize statistical learning to segment vocal streams all suggest the importance of New World primates as models for understanding language development. Although there are many cognitive skills of great apes not seen in New World monkeys, New World monkeys do show rapid social learning, imitation, and even teaching at a level not seen in apes, suggesting that cognitive skills involving social interactions are highly developed, especially in marmosets and tamarins. Finally in recent demonstrations of population variation in communication and other behavior, suggestive of culture, and in observations of stone tool use, similar to that of chimpanzees, New World primates may have more of the cognitive abilities underpinning language than has been previously thought. A fully comparative approach to the evolution of language should consider both homology and converging processes. As part of this enterprise, there is great value in studying New World primates as well as Old World primates and great apes for understanding the evolutionary origins of normal and disordered speech and language processes.

References

1. Slocombe K (2012) Have we underestimated great ape vocal capacities? In: Gibson KR, Tallerman M (eds) Oxford handbook of language evolution. Oxford University Press, Oxford
2. Marler P (1970) Bird song and speech development: could there be parallels? Am Sci 58:669–673
3. Castro NA, Snowdon CT (2000) Development of vocal responses in infant cotton-top tamarins. Behaviour 137:629–646
4. Bauers KA, Snowdon CT (1990) Discrimination of chirp vocalizations in the cotton-top tamarin. Am J Primatol 21:53–60
5. Snowdon CT, Cleveland J (1980) Individual recognition of contact calls in pygmy marmosets. Anim Behav 28:717–727
6. Cleveland J, Snowdon CT (1982) The complex vocal repertoire of the adult cotton-top tamarin (*Saguinus oedipus*). Z Tierpsychol 58:231–270
7. Pola YV, Snowdon CT (1975) The vocalizations of pygmy marmosets (*Cebuella pygmaea*). Anim Behav 23:826–842
8. Snowdon CT, Hodun A (1981) Acoustic adaptations in pygmy marmoset contact calls: locational cues vary with distance between conspecifics. Behav Ecol Sociobiol 9:295–300
9. De la Torre S, Snowdon CT (2002) Environmental correlates of vocal communication of wild pygmy marmosets, *Cebuella pygmaea*. Anim Behav 63:847–856
10. Snowdon CT, Pola YV (1978) Interspecific and intraspecific responses to synthesized pygmy marmoset vocalizations. Anim Behav 26:192–206
11. Snowdon CT (1987) A naturalistic view of categorical perception. In: Harnad S (ed) Categorical perception. Cambridge University Press, New York, pp 332–354
12. Bergman TJ, Beehner JC, Cheney DL, Seyfarth RM (2003) Hierarchical classification by rank and kinship in baboons. Science 302:1234–1236

13. Manser MB (2001) The acoustic structure of suricates' alarm calls varies with predator type and urgency. Proc R Soc Lond B Biol Sci 268:2315–2324
14. Owings DH, Hennessy DF (1984) The importance of variation in sciurid visual and vocal communication. In: Murie JA, Michener GR (eds) The biology of ground dwelling squirrels. University of Nebraska Press, Lincoln, pp 169–200
15. Elowson AM, Tannenbaum PL, Snowdon CT (1991) Food associated calls correlate with food preferences in cotton-top tamarins. Anim Behav 42:931–937
16. Benz JJ, Leger DW, French A (1992) Relations between food preferences and food-elicited vocalizations in golden lion tamarins (*Leontopithecus rosalia*). J Comp Psychol 106: 142–149
17. Benz JJ (1993) Food-elicited vocalizations in golden lion tamarins: design features for representational communication. Anim Behav 45:443–455
18. DiBitetti MS (2003) Food-associated calls of tufted capuchin monkeys (*Cebus apella nigritus*) are functionally referential signals. Behaviour 140:565–592
19. DiBitetti MS (2005) Food-associated calls and audience effects in tufted capuchin monkeys, *Cebus* apella *nigritus*. Anim Behav 69:911–919
20. Gros-Louis J (2004) The function of food-associated calls in white-faced capuchin monkeys, *Cebus capucinus*, from the perspective of the signaller. Anim Behav 67:431–440
21. Gros-Louis J (2006) Acoustic analysis and contextual description of food-associated calls in white-faced capuchin monkeys (*Cebus capucinus*). Int J Primatol 27:273–294
22. Pollick AS, Gouzoules H, DeWaal FBM (2005) Audience effects on food calls in captive brown capuchin monkeys. Anim Behav 70:1273–1280
23. Roush RS, Snowdon CT (2000) Quality, quantity, distribution and audience effects on food-calling in cotton-top tamarins. Ethology 106:673–690
24. Chapman CA, Lefebvre L (1990) Manipulating foraging group size: spider monkey food calls at fruiting trees. Anim Behav 39:891–896
25. Digweed SM, Fedigan LM, Rendell D (2005) Variable specificity in the anti-predator vocalizations and behaviour of the white-faced capuchin, *Cebus capucinus*. Behaviour 142: 997–1022
26. Fichtel C, Perry S, Gros-Louis J (2005) Alarm calls of white-faced capuchin monkeys: an acoustic analysis. Anim Behav 70:165–176
27. Wheeler BC (2010) Production and perception of situationally variable alarm calls in wild tufted capuchin monkeys (*Cebus apella nigritus*). Behav Ecol Sociobiol 64:989–1000
28. Casar C, Byrne RW, Hoppitt W, Young RJ, Zuberbuhler K (2012) Evidence for semantic communication in titi monkey alarm calls. Anim Behav 84:405–411
29. Kirchhof J, Hammerschmidt K (2006) Functionally referential alarm calls in tamarins (*Saguinus fuscicollis* and *Saguinus mystax*)—evidence from playback experiments. Ethology 112:346–354
30. Robinson JG (1979) An analysis of the organization of vocal communication in the titi monkey (*Callicebus moloch*). Z Tierpsychol 1979(49):381–405
31. McConnell PB, Snowdon CT (1986) Vocal interactions between unfamiliar groups of captive cotton top tamarins. Behaviour 97:273–296
32. Saffran JR, Aslin RN, Newport EL (1996) Statistical learning in by 8-month old infants. Science 274:1926–1928
33. Hauser MD, Newport EL, Aslin RN (2001) Segmentation of the speech steam in a non-human primate: statistical learning in cotton-top tamarins. Cognition 78:B53–B64
34. Newport EL, Hauser MD, Spaepen G, Aslin RN (2004) Learning at a distance: statistical learning of non-adjacent dependencies in a non-human primate. Cogn Psychol 49:85–117
35. Saffran J, Hauser M, Seibel R, Kapfhamer J, Tsao F, Cushman F (2008) Grammatical pattern learning by human infants and cotton-top tamarin monkeys. Cognition 107:479–500
36. Snowdon CT, Cleveland J (1984) "Conversations" among pygmy marmosets. Am J Primatol 7:15–20
37. Winter P, Hanley P, Ploog D, Schott D (1973) Ontogeny of squirrel monkey vocalizations under normal conditions and under acoustic isolation. Behaviour 47:230–239

38. Herzog M, Hopf S (1983) Effects of species-specific vocalizations on the behaviour of surrogate reared squirrel monkeys. Behaviour 86:197–214
39. Herzog M, Hopf S (1984) Behavioral responses to species-specific warning calls in infant squirrel monkeys reared in social isolation. Am J Primatol 7:99–106
40. McCowan B, Franceschini NV, Vicino GA (2001) Age differences and developmental trends in alarm peep responses by squirrel monkeys (*Saimiri sciureus*). Am J Primatol 53:19–31
41. Lieblich AK, Symmes D, Newman JP, Shapiro M (1980) Development of the isolation peep in laboratory bred squirrel monkeys. Anim Behav 28:1–9
42. Oller DK (2000) The emergence of the speech capacity. Lawrence Erlbaum, Mahwah, NJ
43. Elowson AM, Snowdon CT, Lazaro-Perea C (1998) Infant "babbling" in a nonhuman primate: complex sequences of vocal behavior. Behaviour 135:643–664
44. Snowdon CT, Elowson AM (2001) "Babbling" in pygmy marmosets: development after infancy. Behaviour 138:1235–1248
45. Goldstein MH, King AP, West ME (2003) Social interaction shapes babbling: testing parallels between birdsong and speech. Proc Natl Acad Sci U S A 100:8030–8035
46. Goldstein MH, Schwade JA, Briesch J, Syal S (2010) Learning while babbling: prelinguistic object-directed vocalizations signal a readiness to learn. Infancy 15:362–391
47. Roush RS, Snowdon CT (1994) Ontogeny of food associated calls in cotton-top tamarins. Anim Behav 47:263–273
48. Roush RS, Snowdon CT (1999) The effects of social status on food-associated calling behavior in captive cotton-top tamarins. Anim Behav 58:1299–1305
49. Mitani JC, Brandt K (1994) Social factors influence the acoustic variability in the long-range distance calls of male chimpanzees. Ethology 96:322–352
50. Mitani JC, Hasegawa T, Gros-Louis J, Marler P, Byrne R (1992) Dialects in wild chimpanzees? Am J Primatol 27:233–243
51. Mitani JC, Hunley KL, Murdoch ME (1999) Geographic variation in the calls of wild chimpanzees: a reassessment. Am J Primatol 47:133–151
52. Crockford C, Herbinger L, Vigilant L, Boesch C (2004) Wild chimpanzees produce group specific calls: a case for vocal learning? Ethology 110:221–243
53. Elowson AM, Snowdon CT (1994) Pygmy marmosets (*Cebuella pygmaea*) modify vocal structure in response to changed social environments. Anim Behav 47:1267–1277
54. Snowdon CT, Elowson AM (1999) Pygmy marmosets alter call structure when paired. Ethology 105:893–908
55. Rukstalis M, Fite JE, French JA (2003) Social change affects vocal structure in a Callitrichid primate (*Callithrix kuhlii*). Ethology 109:327–340
56. De la Torre S, Snowdon CT (2009) Dialects in pygmy marmosets? Population variation in call structure. Am J Primatol 71:333–342
57. Yepez P, de la Torre S, Snowdon CT (2005) Inter-population differences in exudate feeding of pygmy marmosets in Ecuadorian Amazonia. Am J Primatol 66:145–158
58. Snowdon CT, Boe CY (2003) Social communication about unpalatable foods in tamarins. J Comp Psychol 117:142–148
59. Visalberghi E, Addessi E (2000) Response to changes in food palatability in tufted capuchin monkeys *Cebus apella*. Anim Behav 59:231–238
60. Hayes SL, Snowdon CT (1990) Predator recognition in cotton-top tamarins (*Saguinus oedipus*). Am J Primatol 20:283–291
61. Campbell MW, Snowdon CT (2009) Can auditory playback condition predator mobbing in captive-reared cotton-top tamarins? Int J Primatol 30:93–102
62. Campbell MW, Snowdon CT (2007) Vocal response of captive-reared *Saguinus oedipus* during mobbing. Int J Primatol 28:257–270
63. Brumm H, Viss K, Köllmer I, Todt D (2004) Acoustic communication in noise: regulation of call characteristics in a New World monkey. J Exp Biol 207:443–448
64. Egnor SER, Hauser MD (2006) Noise-induced vocal modulation in cotton-top tamarins (*Saguinus oedipus*). Am J Primatol 68:1183–1190

65. Miller CT, Flusberg D, Hauser MD (2003) Interruptibility of long call production in tamarins: implications for vocal development. J Exp Biol 206:2629–2639
66. Egnor SER, Iguina CG (2006) Hauser MD perturbation of auditory feedback causes systematic perturbation in vocal structure in cotton-top tamarins. J Exp Biol 209:3652–3663
67. Burkart JM, van Schaik C (2010) Cognitive consequences of cooperative breeding in primates. Anim Cogn 13:1–19
68. Moscovice LR, Snowdon CT (2006) The role of social context and individual experience in novel task acquisition in cotton-top tamarins (*Saguinus oedipus*). Anim Behav 71:933–943
69. Voelkl B, Huber L (2000) True imitation in marmosets. Anim Behav 60:195–202
70. Joyce SM, Snowdon CT (2007) Developmental changes in food transfers in cotton-top tamarins (*Saguinus oedipus*). Am J Primatol 69:955–965
71. Humle T, Snowdon CT (2008) Socially biased learning in the acquisition of a complex foraging task in juvenile cottontop tamarins (*Saguinus oedipus*). Anim Behav 75:267–277
72. Dillis C, Humle T, Snowdon CT (2010) Socially biased learning among adult cottontop tamarins (*Saguinus oedipus*). Am J Primatol 72:287–295
73. Rapaport LG, Brown GR (2008) Social influences on foraging behavior: in young nonhuman primates: learning what, where, and how to eat. Evol Anthropol 17:189–201
74. Humle T, Snowdon CT, Matsuzawa T (2009) Social influences on the acquisition of ant dipping among the wild chimpanzees (*Pan troglodytes verus*) of Bossou, Guinea, West Africa. Anim Cogn 12:S37–S48
75. Cronin KA, Snowdon CT (2008) The effects of unequal reward distributions on cooperative problem solving by cottontop tamarins (*Saguinus oedipus*). Anim Behav 75:245–257
76. Mendres KA, de Waal FBM (2000) Capuchins do cooperate: the advantage of an intuitive task. Anim Behav 60:523–529
77. de Waal FBM, Berger ML (2000) Payment for labour in monkeys. Nature 404:563
78. Silk JB, Brosnan SF, Vonk J, Henrich J, Povinelli DJ, Richardson AS, Lambeth SP, Mascaro J, Schapiro SJ (2005) Chimpanzees are indifferent to the welfare of unrelated group members. Nature 437:1357–1359
79. Burkart JM, Fehr E, Efferson C, van Schaik CP (2007) Other-regarding preferences in a nonhuman primate: common marmosets provision food altruistically. Proc Natl Acad Sci U S A 104:19762–19766
80. Cronin KA, Schroeder KK, Snowdon CT (2010) Prosocial behaviour emerges independent of reciprocity in cottontop tamarins. Proc R Soc Lond B Biol Sci 277:3845–3851
81. Whitten AJ, Goodall J, McGrew WC, Nishida T, Reynolds V, Sugiyama Y, Tutin CEG, Wrangham RW, Boesch C (1999) Cultures in chimpanzees. Nature 399:682–685
82. Perry S, Baker M, Fedigan L, Gros-Louis J, Jack K, MacKinnon KC, Manson JH, Panger M, Pyle K, Rose L (2003) Social conventions in wild white-faced capuchin monkeys Evidence for traditions in a neotropical primate. Curr Anthropol 44:241–268
83. Otoni EB, Izar P (2008) Capuchin monkey tool use: overview and implications. Evol Anthropol 17:171–178
84. Visalberghi E, Fragaszy D, Ottoni EB, Izar P, Gomes de Oliveira M, Andrade FRD. Characteristics of hammer stones and anvils used by wild bearded capuchin monkeys (*Cebus libidinosus*) to crack open palm nuts. Am J Phys Anthropol. 2007;132:426–44.
85. Miller C, Dimauro A, Pistorio A, Hendry S, Wang X (2010) Vocalization induced cFos expression in marmoset cortex. Front Integr Neurosci 4:128
86. Johnson LA, Della Santina CC, Wang X (2012) Temporal bone characterization and cochlear implant feasibility in the common marmoset (*Callithrix jacchus*). Hear Res 290:37–44
87. Ferris CF, Snowdon CT, King JA, Sullivan JM Jr, Ziegler TE, Ludwig R, Schultz-Darken N, Wu Z, Olson DP, Tannenbaum PL, Einspanier A, Vaughn JT, Duong TQ (2004) Imaging neural pathways associated with stimuli for sexual arousal in nonhuman primates. J Magn Reson Imaging 19:168–175

Chapter 10
Apes, Language, and the Brain

William D. Hopkins

Abstract Language is a unique form of communication in humans and is unmatched in the animal kingdom. There are well-defined cortical regions involved in both the comprehension and production of speech including Wernicke's and Broca's area. To what extent these regions play a role in the communicative abilities of primates, notably great apes, remains a central topic of research in neuroscience, anthropology, and psychology. In this chapter, I present an overview of the cognitive foundations of gestural and vocal communication in chimpanzees, including some results from language-trained apes. I also present data on the evolution in size and lateralization of Wernicke's and Broca's area in chimpanzees. These anatomical data are combined with behavioral data to show how individual differences in gestural and vocal communication are associated with volumetric and lateralized differences in Broca's and Wernicke's areas. The collective findings are discussed within the context of language evolution and the emergence of complex motor and cognitive processes in humans after the split from the common ancestor with chimpanzees.

Keywords Language • Speech • Brain asymmetry • Handedness • Vocal communication • Gesture

W.D. Hopkins, B.A., M.D., Ph.D. (✉)
Neuroscience Institute and Language Research Center, Georgia State University, P.O. Box 5030, Atlanta, GA 30305, USA

Division of Developmental and Cognitive Neuroscience, Yerkes National Primate Research Center, 954 Gatewood Road, Atlanta, GA 30322, USA
e-mail: whopkins4@gsu.edu

S.A. Helekar (ed.), *Animal Models of Speech and Language Disorders*,
DOI 10.1007/978-1-4614-8400-4_10,

Introduction

As a form of communication, human language and speech is unmatched in the animal kingdom. The cognitive, motor, and neural foundations that distinguish human speech from other animal communication systems have been a central question in the sciences for more than 200 years and can be dated back to some of the earliest writings by psychologists, linguist, philosophers, and biologists [1–4]. Perhaps no other species have sparked more interest in the question of language origins and evolution than the great apes which include chimpanzees and bonobos (*Pan*), gorillas (*Gorilla*), and orangutans (*Pongo*) likely due to their physical resemblance and genetic similarity to humans [5]. Great apes, beginning with orangutans, diverged from more distantly related Old World monkeys and lesser apes around 15 mya. Within the ape lineage, gorillas subsequently split off from orangutans around 9 mya following by a further split around 5–6 mya with the last common ancestor of humans and chimpanzees diverging about 5–6 mya.

From a behavioral and cognitive standpoint, many have speculated as to whether apes might be capable of language and speech, and early in the twentieth century, there were several empirical attempts to teach apes to speak which included orangutans [6] and chimpanzees [7–9]. These studies attempted to teach apes to articulate speech, but it was not recognized until later that limits in the peripheral speech organs and potentially other factors prevented apes from producing intelligible sounds [1, 10]. This subsequently led to a series of studies beginning in the 1960s that used alternative communication systems with chimpanzees as a means of assessing their linguistic potential. These alternative systems included American Sign Language (ASL), plastic token, and visual-graphic symbols [11–18].

In contrast to the ape-language studies, an alternative approach to studying the evolution of language and speech has been to look for parallels in the natural communication systems of great apes with those found in humans [19]. These studies have primarily focused on the vocal and gestural communication repertoires of apes and have largely occurred within the context of competing theories on the origins of language in humans. The gestural origins theory postulates that early hominids initially possessed a gesture-based system of communication that included basic signs that were semantic in function [20, 21]. In contrast, others have hypothesized that there was direct selection on the neurobiological system that govern vocal communication in the common ancestor of apes and humans and this led to emergence of speech in modern humans [2]. Lastly, the work of McNeil and others have emphasized the interface between gestures and vocal communication in humans [22]. Some of the clearest examples of this multimodal link coming from observations that people often gesture while speaking, even when there is no apparent audience, such as when talking to someone on the telephone. There is a significant paucity of research on multimodal communication in nonhuman primates, though the topic is receiving increasing attention [23, 24].

In this chapter, the goal is not necessarily to provide evidence in support or against any of these proposed evolutionary models but to present a more general

review of the cognitive and communicative abilities of great apes with specific emphasis on chimpanzees. I have chosen to focus on chimpanzees because they are genetically the most closely related species to humans and, by far, they have been the most extensively studied of the great apes. I start by discussing the basic results that have emerged from the ape-language research, then summarize some recent findings on the gestural and vocal communication in apes. In the latter portion of the chapter, I present data on cortical organization and asymmetries in chimpanzees from two important brain regions known to be involved in language and speech in humans, notably Broca's and Wernicke's area. I conclude by highlighting the importance of research with great apes for understanding the evolution of language and speech in humans and offer some suggestions for future research.

Ape-Language Studies

After nearly 50 years of so-called ape-language research, several important characteristics of human language have emerged which have been recently summarized by Lyn [25]. First, it seems reasonable to conclude that great apes can learn that symbols can represent nouns and verbs and that, in the case of nouns, the symbols can represent something absent in time and space. In short, apes are capable of symbolic thought and can use these symbols referentially. Second, great apes can both name and comprehend the meaning of symbols. That is to say, if a human signs "apple," the apes are capable of selecting an apple from an array of objects or other foods (comprehension). Conversely, if a human experimenter holds up an apple, the apes can reliably sign "apple" (naming). Thus, apes use symbols in both the receptive and productive domains of communication and their representation of symbols is multimodal. Third, great apes can use their signs or symbols to communicate novel phrases, which suggest at least some level of generativity in their thoughts and use of symbols. For example, the chimpanzee Lana learned a symbol for the food "apple" and another symbol the color "orange." Upon first encountering her human experimenter eating an orange (which was unfamiliar to her), Lana negotiated requesting the food by typing out "please Tim give Lana piece of apple which is orange" [13]. Fourth, there are several reports that language-trained apes comprehend spoken English (and presumably other languages depending on the native speakers tongue). For example, Patterson reported that the gorilla, Koko, understands more than 500 spoken English words. In a more systematic report, Savage-Rumbaugh and colleagues have reported that two bonobos (Kanzi and Panbanisha) and one chimpanzee (Panzee) reliably comprehend at least 100 words [26–28]. Finally, the impressive abilities to comprehend spoken English are not limited to single words. Kanzi, one of the bonobos in the ape-language studies, has been shown to comprehend more than 700 novel sentences on the basis of speech cues alone [26]. These studies are particularly impressive when considered within the context of the experimental rigor that has been extended to rule out other potential explanations for his speech comprehension abilities. Indeed, building on these

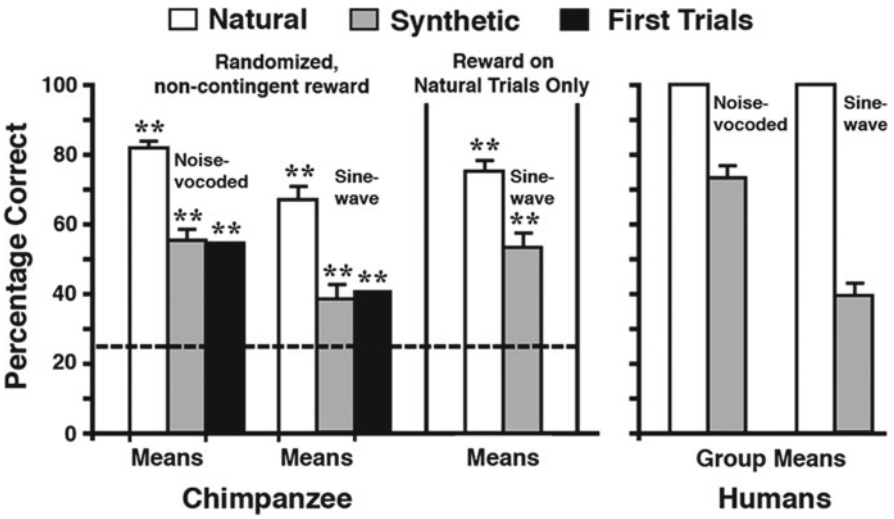


Fig. 10.1 Performance by the chimpanzee and human listeners. Means and standard errors of percentage-correct performance for 48 words heard in natural, NV, and SW forms. Experiments with the chimpanzee, Panzee, included testing each of 48 words 16 times in natural and four times in synthetic form. First trials represent the 48 first instances of the chimpanzee hearing a word in a given synthetic form. The first set of SW results shows performance with noncontingent, intermittent reward delivery and no response feedback. The second set shows performance with contingent reward received on natural trials but with no reward or response feedback on SW trials. The *dashed line* indicates the chance-performance rate of 25 % correct. Humans heard and identified all 48 words once each in natural form, followed by either NV (16 listeners) or SW (16 listeners) versions. All comparisons to chance performance were statistically significant at $p \leq 008$ and are marked by a pair of *asterisk*. *Reprinted with permission from Heimbauer L, Beran M, Owren M. A chimpanzee recognizes synthetic speech with significantly reduced acoustic cues to phonetic content. Curr. Biol. 2011;21(1210–1214)* [29]

studies, a recent study showed that Panzee, a chimpanzee who comprehends human speech, could reliably identify individually spoken English words that had been either synthesized or manipulated in such a matter to become impossibly unspeech-like [29]. The performance of Panzee was similar to results from human subjects tested using the same stimuli (see Fig. 10.1), and these findings clearly challenge some long-held views on theories of speech perception and language.

Gestural Communication in Great Apes

Studies in captive and wild great apes have documented a number of important cognitive processes and behavioral characteristics of gestural and vocal communication in great apes, and these have been reviewed previously [30, 31]. Briefly, from the standpoint of gestural communication, there is now clear evidence that great apes

gesture more frequently in the presence compared to the absence of an audience. Furthermore, apes gesture to request otherwise unattainable food items or objects; they will alternate their gaze between the referent and the social agent. This suggests that the ape gestures are intentional and that the individuals understand the function of their signals [32]. This type of triadic form of communication has been described in developing children (9–15 months of age) and appears to predict the subsequent development of language and speech [33–35]. For example, children that develop pointing in conjunction with alternation of gaze at an earlier age also develop speech at an earlier age.

Additional studies have shown that chimpanzees and orangutans can alter the type of communication signal they use in response to the attentional state of a human experimenter [36–39]. For example, Leavens et al. [40] performed an experiment in which they recorded the type of communication signal chimpanzees used toward a human experimenter who was either (a) looking at them and offering food, (b) looking at a cagemate and offering them food, or (c) offering food to a chimpanzee living in an adjacent cage. When the human experimenter was looking at the focal chimpanzee and offering the food, the majority of the chimpanzees chose to use a visual communication gesture in the form of either a begging manual gesture or lip pout to communicate with the human. In contrast, when the human experimenter was looking at another chimpanzee and offering food, a significant majority of the focal chimpanzee subjects used an auditory gesture such as attention-getting (AG) sounds (see below), clapping of the hands, or banging on the cage. Thus, the chimpanzees were able to assess the attentional state of the human experimenter and subsequently alter their communicative behavior so as to use a more effective signal within each context.

Despite these impressive abilities, there are some limits and differences in the inherent gestural communication systems of apes and humans [41, 42]. First, though species-specific gestures appear to denote some specific meaning to the other conspecifics, there is little evidence of more sophisticated gestures such as iconic gestures [42]. Further to this point, there have been several attempts to examine whether captive and wild apes put gestures together in consistent sequences, and again though some evidence can be found [43–45], the complexity of multi-gestures sequences are not terribly sophisticated. The exception appears to be the integration of gestures with symbol use in some of the language-trained apes [46]. Thus, differences in early rearing or sociolinguistic experiences may have some influence on how apes use gestures. Second, some have suggested that all apes gestures are request based or imperative in function, whereas human pointing is both imperative (request based) and declarative (produced for the sake of sharing information) [47]. This criticism has also been leveled at the ape-language research with the claim being that apes only produce signs or keyboard utterances to request and seldom to simply share information [15]. This alleged distinction between apes and humans has not been adequately investigated in apes, but there is at least one report of the use of declarative symbols by language-trained apes [48], so some further research is warranted before making any definitive conclusions. Some have even suggested that there is no real psychological distinction between imperative and declarative

pointing [49]. Lastly, there is very little evidence that ape gestures are socially learned, like many human languages. For example, Gentry and colleagues [50] recently found that the repertoire and types of gestures produced by captive lowland gorillas were similar to those observed in wild populations. Similar findings have been reported in wild chimpanzees [51]. In a different study, Tomasello and colleagues [52] taught a novel gesture to a single chimpanzee living in the group and then assessed the proliferation in the use of the gesture in the other group members and found virtually no transfer in the use of the signal. Again, this area of research has not been extensively studied, so it is difficult to make any strong conclusions. Moreover, the lack of social learning of gestures is quite inconsistent with the remarkable social learning seen in apes in other domains such as in foraging or tool use [53–56].

Vocal Communication

Far less research has been conducted on vocal communication in great apes beyond the basic descriptive ethograms of the different types and contexts in which species-specific sounds are produced [57, 58]. More recently, several studies in both captive and wild apes have revealed some new and important findings regarding the cognitive and motivational foundations of vocal communication as it pertains to the evolution of human speech. First, several recent studies have shown that chimpanzees demonstrate the audience effect for vocal signals and may use them to inform conspecifics with specific information for which they may be ignorant. For example, in 2005, Slocombe and Zuberbuhler reported that a captive chimpanzee produced different vocalizations to indicate the type of food found at different localization in the subject's enclosure [59]. Additional studies in wild chimpanzees have shown that the audience effect can be found in the use of food calls [60] as well as pant-grunts [61]. Even more impressive are the findings presented by Crockford et al. [62] who found that wild chimpanzees selectively produce alarm "hoos" in the presence of snakes more frequently when surrounded by individuals who are ignorant to the presence of the snake than when surrounded by individuals who know the snake is there. These findings are significant for two reasons. First, they suggest that the chimpanzees are attempting to *inform* the other chimpanzees about the presence of the snake rather than simply producing an involuntary emotional response to the stimulus. This observation is not trivial because it appears to differ from reports in the use of so-called "semantic" vocalizations by some primate species, such as vervet monkeys. Seyfarth et al. [63] showed through playback studies that vervet monkeys differentially respond to three different alarm calls to predators including leopards, snakes, and raptors. Cheney and Seyfarth [64–66] have suggested that when the monkeys produce these alarms calls, the intent of the signaler is not to inform the other group members of the predators. Instead, the sight of the predators elicits the production of the vocalization and the group members have learned to differentially respond to the calls. Second, the findings suggest that the chimpanzees have

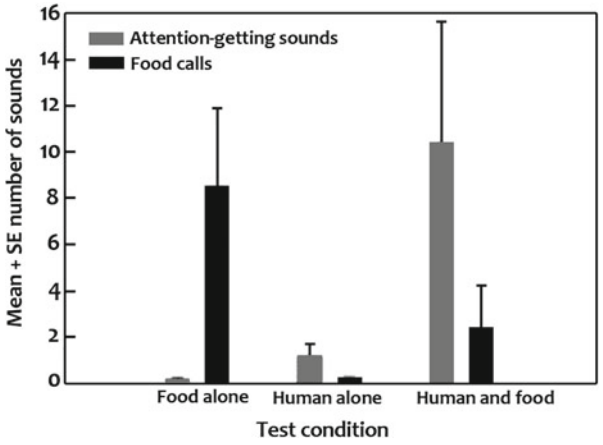


Fig. 10.2 Mean + SE number of food calls and attention-getting sounds as a function of whether food was presented alone, a human was present alone or a human and food were presented at the same time. *From Hopkins WD, Taglialatela JP, Leavens DA. Do chimpanzees have voluntary control of their facial expressions and vocalizations? In: Vilain A, Schwartz J-L, Abry C, Vauclair J, eds. Primate communication and human language: Vocalisation, gestures, imitation and deixis in humans and non-humans. Amsterdam: John Benjamins Publishing Company; 2011:71–90* [68]

voluntary control over their vocalizations, a result that challenges many historical and contemporary views of primate vocalizations [67, 68]. In other words, the chimpanzees can choose to produce the sounds or inhibit their production.

Results from captive apes have revealed similar evidence of voluntary control of vocalizations and potentially evidence that the use of some sounds may be socially learned. As noted previously, experimental studies in captive chimpanzees and orangutans have shown that they produce attention-getting (AG) sounds to capture the attention of an otherwise inattentive audience. For example, Hopkins et al. [69] recorded the types of vocalizations chimpanzees made when presented with bananas alone, a human experimenter alone, or when a banana was present with a human experimenter. When the food was presented alone, the chimpanzees predominantly made food calls, whereas when a human was present in conjunction with the food, the chimpanzees predominantly made AG sounds (see Fig. 10.2). These findings also suggest that the apes have voluntary control over their facial expressions and the production of these attention-getting sounds. In addition, some of the attention-getting sounds produced by captive apes have been described as idiosyncratic and appeared to be individually learned. For instance, in chimpanzees, three attention-getting sounds have been described including the raspberry, kiss squeak, and extended food grunt. The raspberry and kiss squeak are unvoiced, while the extended food grunt is a voiced sound and involves the use of the larynx. The raspberry, in particular, has not been described extensively in wild chimpanzees and appears to be individually learned. Most impressive is the report that several captive orangutans learned to whistle to capture the attention of their human experimenter [70]. The level of orofacial motor control required to whistle and learn novel sounds such

as the raspberry should leave little doubt the apes have voluntary control over their orofacial musculature, vocal apparatus, and air flow. Lastly, there is some evidence that the use of attention-getting sounds is socially learned. Taglialatela et al. [71] examined the occurrence of attention-getting sounds in a sample of 158 chimpanzees and found that significantly higher proportion of chimpanzees that produced attention-getting sounds were born to and raised by females who also made AG sounds. Chimpanzees born to females who produce AG sounds who were taken at birth and raised by humans did not consistently produce AG sounds.

The flexibility in vocal communication observed in chimpanzees has also been reported in bonobos and orangutans. Bonobo vocalizations have a much higher frequency than most chimpanzee vocalizations, and contextually they are used very differently [72–74]. Perhaps most impressive are the data on vocal communication in the language-trained bonobo Kanzi who produces a number of vocalizations that are not found in the species-typical repertoire of his species [75]. An additional study of the vocal communication of Kanzi showed that he produced at least four different classes of vocalization that were reliably associated with different semantic categories and he used these vocalizations in dyadic interactions with human experimenters. Thus, the production of these vocalizations by Kanzi was used conversationally, in this sense that they were elicited in response to human-specific queries [76, 77].

The Neural Substrates of Language and Speech

One important feature of human language and speech is that it is strongly lateralized in the human brain. A significant majority of humans are left-hemisphere dominant for language and speech. Interestingly, in humans, lateralization for language is somewhat modified by their handedness. For example, it has been shown that approximately 96 % of right-handed individuals are left-hemisphere dominant for language compared to 70 % of left-handed people [78, 79]. The link between handedness and lateralization for speech has been the foundation for a number of genetic models of hemispheric specialization, which hypothesize that different genes code for language lateralization and that right-handedness is a consequence of the expression of lateralization. For example, Annett [80] has proposed the right-shift (RS) theory. In the RS model, individuals inherit a single allele that codes for left language lateralization (rs+) or they do not (rs−). Based on simple Mendelian genetics, individuals are either homozygous RS (rs+, rs+), heterozygous (rs+, rs−), or lack a copy of the allele (rs−, rs−). The RS theory proposes that anyone who inherits the rs+ gene will be left-hemisphere dominant for language (75 % of the population), while lateralization for those homozygous rs− is randomly determined (12.5 % left and 12.5 % right). Thus, approximately 88 % of humans are left-hemisphere dominant for language due to the rs+ gene, and this value approximates the observed neuropsychological data. According to the RS theory, right-handedness is a consequence of left-hemisphere dominance for fine motor control implicitly needed for

speech. In other words, because speech requires fine motor control and it lateralized to the left hemisphere, the right hand by default becomes the preferred hand for skill actions in most people. Furthermore, because the assumption is that lateralization for speech is the driving mechanisms for right-handedness, the RS theory presupposes that species-level right-handedness should only be evident in humans.

Because handedness is linked to language lateralization (albeit weakly), many have argued that hemispheric specialization is unique to hominoid evolution. The evolutionary argument goes something like this: (1) Language is unique to humans, (2) only humans show population-level right-handedness, (3) animals do not have language, and (4) animals do not show population-level hand (or limb) preferences. Ergo, only humans have hemispheric specialization and this is explicitly linked to the evolution of language and speech [81]. In the past 20 years, this long-held perspective has been challenged from data many species. Specifically, data from a variety of vertebrates and invertebrates have begun to demonstrate evidence of population-level behavioral asymmetries [82–87]. For instance, right-limp preference has been reported in toads when removing a foreign object on their back or when righting themselves when floating upside down in water [88]. A variety of bird species and some fish show eye dominance when viewing different types of stimuli [89]. There is also some evidence of population-level asymmetries footedness in birds [90]. These studies clearly challenge the long-held belief that population-level behavioral asymmetries are unique to humans and, indeed, raise questions regarding the assumption that language is a necessary condition for the emergence of asymmetrical functions within the brain; however, our understanding of the evolutionary factors that are specifically selected for asymmetries in the primate lineage and which culminated in the degree of lateralization seen for language and handedness in modern humans remains unclear.

Behavioral Precursors: Lateralization for Manual Gestures and Orofacial Asymmetries

In primates, as a means of addressing the notion that language lateralization and handedness are linked, the focus has overwhelmingly been on studies of hand preference for a variety of different kind of tasks [91–95], and the findings have not been clear. Early on, many studies focused on handedness for simple reaching, and the results suggested that primates showed no evidence of population-level handedness [96]. In other words, though a significant majority of the subjects showed individual hand preferences, there were roughly equal numbers of left- and right-handed monkeys or apes. It was not until the early 1990s that a number of studies began to emerge that focused on handedness for complex actions such as coordinated bimanual actions, tool use, and prehensile grasping, and these latter studies appear to have revealed at least some evidence of population-level handedness for some tasks [93, 96].

Rather than focus on handedness for manual actions that were directed toward objects, several investigators have focused on the question of asymmetries in

Fig. 10.3 Mean handedness scores (HI) for manual gestures and simple reaching in human children, chimpanzees, and baboons [97]. *Modified from*: **Baboon**: *Papademetriou E*, *Sheu CF*, *Michel GF*. *A meta-analysis of primate hand preferences for reaching and other hand-use measures*. *J. Comp. Psychol. 2005*;*119*:*33–48*. [96]; **Chimpanzee**: *Hopkins WD*, *Russell JL*, *Freeman H*, *Buehler N*, *Reynolds E*, *Schapiro SJ*. *The distribution and development of handedness for manual gestures in captive chimpanzees* (*Pan troglodytes*). *Psychological Science 2005*;*16*(*6*):*487–493*. [146]; **Human**: *Cochet H*, *Vauclair J*. *Features of spontaneous pointing gestures in toddlers*. *Gesture*. *2010*;*10*(*1*):*86–107* [100]

communicative motor actions including manual gestures and vocal communication. Because human language and speech is strongly left lateralized, the basic theoretical foundation of this research has been to focus on lateralization in primate behaviors that are communicative in function. As noted above, primates, and particularly apes, have rich manual gestural repertoires, and therefore focusing on quantifying handedness for these kinds of behaviors has been the focus of some research. Small but significant population-level right-handedness has been found for manual gestures in chimpanzees, gorillas, bonobos, orangutans, and baboons, a species of Old World monkey [97–99]. More recently, similar patterns of right-handedness have been reported in developing preverbal human children [100, 101]. One interesting observation from these studies is that handedness for manual gestures appears to elicit significantly greater right-handedness than manual actions that are not communicative in function, such as simple reaching (see Fig. 10.3). These findings suggest that the potential neural foundation for gestures may differ from other simple manual actions that are not communicative in function and, at face value, lend some support to the gestural origins theory of language [102].

In contrast to manual gestures, there have also been attempts to quantify orofacial asymmetries during the expression of species-specific and learned vocalizations in monkeys and apes. In one of the first studies, Hook-Costigan and Rogers [103] measured the lateralization in magnitude of orofacial expressions during the production of vocalizations in a sample of marmosets, a New World monkey species. For fear-based calls, the monkeys showed a left hemiface asymmetry but showed a

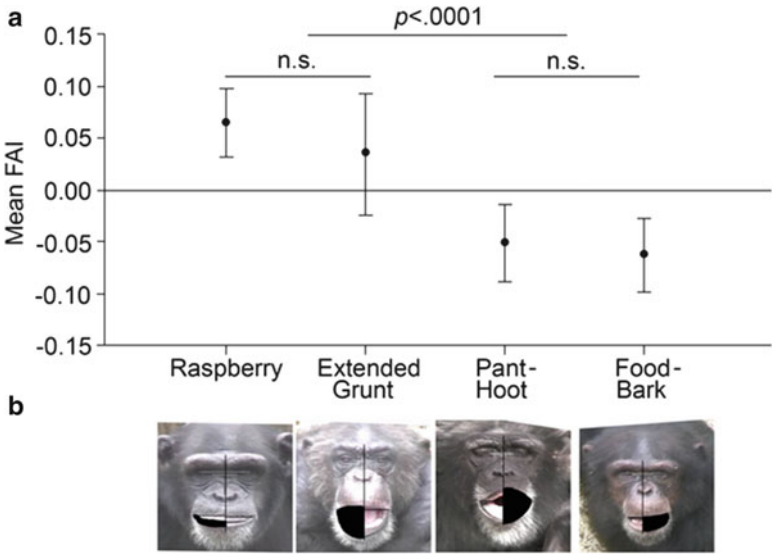


Fig. 10.4 (**a**) Least squares means of FAI scores for the raspberry, extended grunt, pant-hoot, and food-bark expressions along with 95 % confidence intervals for these values. Positive FAI scores represent right hemi-mouth biases and negative values reflect left hemi-mouth biases. (**b**) Illustration of hemi-mouth area calculation procedure on representative images of the raspberry, extended grunt, pant-hoot, and food bark under their corresponding mean FAI values. *Reprinted with permission from Losin ER*, *Freeman H*, *Russell JL*, *Meguerditchian A*, *Hopkins WD. Left hemisphere specialization for oro-facial movements of learned vocal signals by captive chimpanzees PlosONE. 2008*;*3*:*1–7* [107]

right hemiface asymmetry for social contact calls. In another study, Wallez and colleagues [104, 105] measured orofacial asymmetries in adult and infant macaques and baboons and reported a right hemisphere bias for several expressions. Adopting similar procedures to those used by Hook-Costigan and Rogers [103], Fernandez-Carriba et al. [106], and Losin et al. [107] measured asymmetries in the facial expressions of chimpanzees associated with species-specific calls as well as for the production of the AG sounds previously described (see above). For species-specific vocalizations, a left orofacial asymmetry was found for hooting, play, silent-barred teeth, screams, and food calls. In contrast, for AG sounds, a right orofacial asymmetry was found, suggesting that learning novel sounds that require novel manipulation of the orofacial musculature involves the left hemisphere (see Fig. 10.4). Thus, vocalizations typically produced in the context of emotional expressions appear to be under the control of the right hemisphere, whereas the production novel learned sounds appear to be under the control of left half of the brain.

Finally, the simultaneous production of speech and gesture has been reported by a number of individuals observing humans while in a conversation [108]. That is to say, people often move their hands while talking. It has also been reported that right-hand actions are much more commonly produced than left-hand actions

when people are speaking. One interpretation for this observation is that, because the left hemisphere is dominant for speech, it results in greater manual actions produced by the right hand during speech production. In chimpanzees, we have also found a temporal relationship between the production of AG sounds and manual gestures. In some chimpanzees, manual gestures are often produced in conjunction with the production of AG sounds, and the onset of both the manual gesture and AG sound are linked close in time. Thus, like in humans, a homologous pattern of results has been found with AG sound production more often accompanied by right compared to left-hand gestures [109]. This observation in both humans and chimpanzees highlights the interrelationship between manual and orofacial actions and reinforces the view that language is multimodal.

Neuroanatomical Precursors: Broca's and Wernicke's Area Homologs

There are two well-known cortical regions that have been implicated in a variety of linguistic functions in the human brain including Broca's and Wernicke's areas. Broca's area is located in the inferior frontal gyrus (IFG). Anatomically, based on specific sulci landmarks, the IFG in humans is divided into three regions including the pars opercularis, pars triangularis, and pars orbitalis [110]. The anterior portion of the pars opercularis is defined by the ascending ramus, while the posterior landmark is defined by the precentral inferior sulcus. The anterior border of the pars triangularis is the horizontal ramus, and the posterior border is the ascending ramus. The homolog to Broca's area in the chimpanzee brain is also located in the IFG, but the anatomy of this region differs from humans (see Fig. 10.5) [111]. Specifically, like the human brain, the pars opercularis can be anatomically defined in chimpanzee brain with the fronto-orbital sulcus serving as the anterior border and the precentral inferior sulcus defining the posterior border; however, with rare exception, the pars triangularis cannot be anatomically defined in the chimpanzee brain [112]. Thus, there is increased cortical folding and gyrification within the IFG of the human compared to chimpanzee brain, and this is likely a consequence of increased selection for cortical representation of language and speech in humans, after the split from the common ancestor with chimpanzees.

Keller et al. [111] compared the grey matter volume of the pars opercularis in a sample of 30 humans and 30 chimpanzees using the same landmarks and scanning procedures. These authors found that the pars opercularis is approximately three times larger in humans than in chimpanzees. Though it is often reported that humans show a left anatomical asymmetry for either the pars opercularis or pars triangularis [113–115], several recent studies that have specifically quantified the grey matter have failed to find evidence of population-level asymmetries in either the pars opercularis or pars triangularis [111, 116]. Similarly, for the pars opercularis in humans and chimpanzees, Keller et al. [110] failed to find population-level asymmetries in grey matter volume in either species.

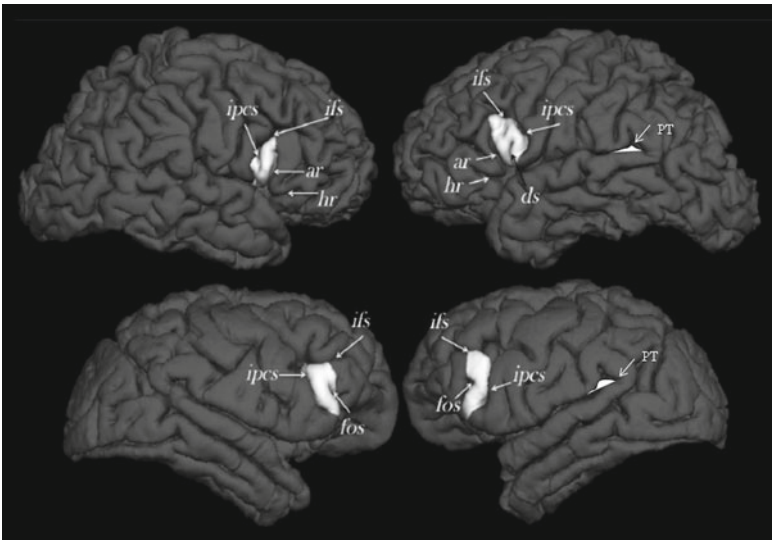


Fig. 10.5 External morphology of the right and left cerebral hemispheres of a randomly selected human (*top*) and chimpanzee (*bottom*) from the study sample indicating the location of the frontal operculum (*white*) and the defining sulcal contours (not to scale). Cortical reconstructions and labeling of the frontal operculum was performed using Freesurfer software (http://surfer.nmr.mgh.harvard.edu/). *ar* Anterior ascending ramus of the Sylvian fissure, *ds* diagonal sulcus, *fos* fronto-orbital sulcus, *ifs*, inferior frontal sulcus, *ipcs* inferior precentral sulcus. *From Keller SS, Roberts N, Hopkins WD. A comparative magnetic resonance imaging study of the anatomy, variability and asymmetry of Broca's area in the human and chimpanzee brain J. Neurosci. 2009;29:14607–14616* [111]

Beside the morphology of Broca's area, there have been several studies on the cytoarchitectonic regions of Broca's area in chimpanzees and humans. Broca's area is made up of Brodmann's area 44 and 45 (Ar44 and Ar45, respectively) [117, 118]. Recently, Schenker et al. [119] quantified Ar44 and Ar45 in a sample of 12 postmortem chimpanzee brains. Schenker et al. [119] found that Ar44 neurons were primarily found in the cortex immediately anterior to the precentral inferior sulcus, whereas Ar45 was more prevalent in the tissue lying anterior to the fronto-orbital sulcus. As evidence to the potential importance of Broca's area to language and speech, Schenker et al. [119] further found that Ar44 and Ar45 were both roughly 6–7 times larger in the human compared to chimpanzee brain, particularly for the left hemisphere. In comparison to other cytoarchitectonic regions, such as Area 4 (primary motor cortex) and Area 13 (limbic region of the prefrontal cortex) which are relatively much smaller in humans compared to chimpanzees, these findings strongly suggest that there has been selective expansion of the Ar44 and Ar45 in the human brain after the spit for the spit from the common ancestor with chimpanzees [120].

In addition to the size of Ar44 and Ar45, there are also potential species differences in lateralization in these regions between humans and chimpanzees, though given the rather limited number of postmortem brains that have been examined, some caution is warranted. Specifically, Uylings et al. [121] measured the volume

of Ar44 and Ar45 in ten human brains and found significant leftward asymmetries for both regions. Schenker et al. [119] failed to find asymmetries in either the volume, neuron density of neuron number for Ar44 and Ar45 in the chimpanzees. However, it should be noted that there were significant individual differences within the chimpanzee sample and sex differences approached conventional levels of statistical significance. Specifically, there were six male and six female chimpanzees in the Schenker et al. [119] study, and one individual (specimen C0491) showed a pronounced rightward asymmetry in region Ar44 to the point where box plots indicated that he was an outlier. If this specimen is removed, then significant sex differences were found in Ar44 with males showing a leftward asymmetry and females showing a rightward bias.

Planum Temporale

In contrast to the IFG, there appears to be much more homology in the cortical organization and lateralization of the PT between humans and chimpanzees. The PT is the flat bank of tissue that lies posterior to Heschl's gyrus or the primary auditory cortex (see Fig. 10.5) and overlaps with Wernicke's area. Numerous functional imaging and clinical studies have implicated the posterior superior temporal gyrus in the comprehension of language and speech as well as a number of other functions [122, 123]. Morphologically, the PT exhibits perhaps one of the most consistent and robust asymmetries in the human brain. Numerous studies that have quantified the PT surface area from postmortem brains or more recently MRI have reported that approximately 75 % of human brains show a leftward asymmetry [124]. Recent studies using MRI have also found significant leftward asymmetries in grey matter volume of the PT in humans, particularly among right-handed individuals [116, 125].

Quantifying the PT in great ape brains can be accomplished using the same landmarks as those used in human brains. Applying identical landmarks and procedure, significant leftward asymmetries in the PT surface area, and grey matter volume have been found in chimpanzees [126–128]. Indeed, in both humans and chimpanzees, approximately 70–75 % of individuals show a leftward asymmetry in the PT. Anatomically, the PT is made up primarily of Brodmann's area 22 (sometimes referred to as area Tpt or BA22). Like Ar44 and Ar45, humans have a relatively large Tpt compared to chimpanzees but the fold difference is not nearly as robust, suggesting that this region is more evolutionarily conserved than Broca's area. Presently, much less is known about cytoarchitectonic asymmetries in area Tpt in humans, but one report in three human brain specimens reported leftward asymmetries for all three subjects [129]. In chimpanzees, Spocter et al. [130] found a significant leftward asymmetry in the area Tpt volume in a sample of 12 chimpanzee brains. There is also one report of leftward asymmetries in area Tpt in six rhesus monkeys, [131] which is interesting from a comparative perspective, because asymmetries in the surface area and grey matter volume of the posterior temporal lobe are absent in this species [132]. Thus, unlike humans and chimpanzees, there appears to

be some discrepancies in the presence of asymmetries in area Tpt in more distantly related macaques depending on the method and level of analysis (macrostructural versus miscrostructural).

Behavioral and Brain Asymmetries with an Emphasis on Gestural and Vocal Communication

In our laboratory, we have also tested for association between brain asymmetries and lateralization in manual gestures and vocal communication in the chimpanzees. In one study, Taglialatela and colleagues [133] found that chimpanzees who preferred to gesture with their right hand had significantly larger leftward asymmetries in the IFG compared to chimpanzees who were non-right-handed in their gestures. Interestingly, when these same chimpanzees were compared for asymmetries in another brain region, the motor-hand region of the precentral gyrus, no differences were found between left- and right-handed individuals. Thus, individual differences in hand use for manual gestures were linked specifically to asymmetries in the IFG but not other cortical regions within the frontal lobe. Similarly, chimpanzees who prefer to gesture with the right hand showed greater leftward asymmetries in grey matter volume of the PT compared to left-handed individuals [127]. Cantalupo et al. [134] found that right-handed chimpanzees showed a significantly lower ratio in white-to-grey in perisylvian regions of the left hemisphere compared to non-right-handed individuals (this findings would imply greater grey matter in these regions).

Functional Asymmetries

The previous summary of findings on neuroanatomical asymmetries focused largely on morphology and cytoarchitectonics of the cortex and not on function. It is possible that different brain regions might be specialized for specific functions but necessarily exhibit robust morphological asymmetries. Therefore, in this final section, I present some recent findings on functional asymmetries in chimpanzees derived from in vivo positron emission tomography (PET). The feasibility of functional imaging of the chimpanzee brain is relatively new, and therefore these types of studies are in their infancy. In two separate studies in chimpanzees, subjects were PET imaged when they were either (1) producing AG sounds while simultaneously requesting an out of reach food [135] or (2) passively hearing two different classes of species-specific vocalizations (either broadcast or proximal calls) [136]. With specific reference to the IFG, for the task in which the chimpanzees were producing AG sounds and gesturing, among a number of regions, Taglialatela et al. [135] found significant left hemisphere clusters in the dorsal portion of the IFG and in the cortex immediately anterior to the fronto-orbital sulcus (see Fig. 10.6). In the follow-up study and analysis, Taglialatela et al. [137] found that the left hemisphere asymmetries within the IFG were largely attributable to the production of the

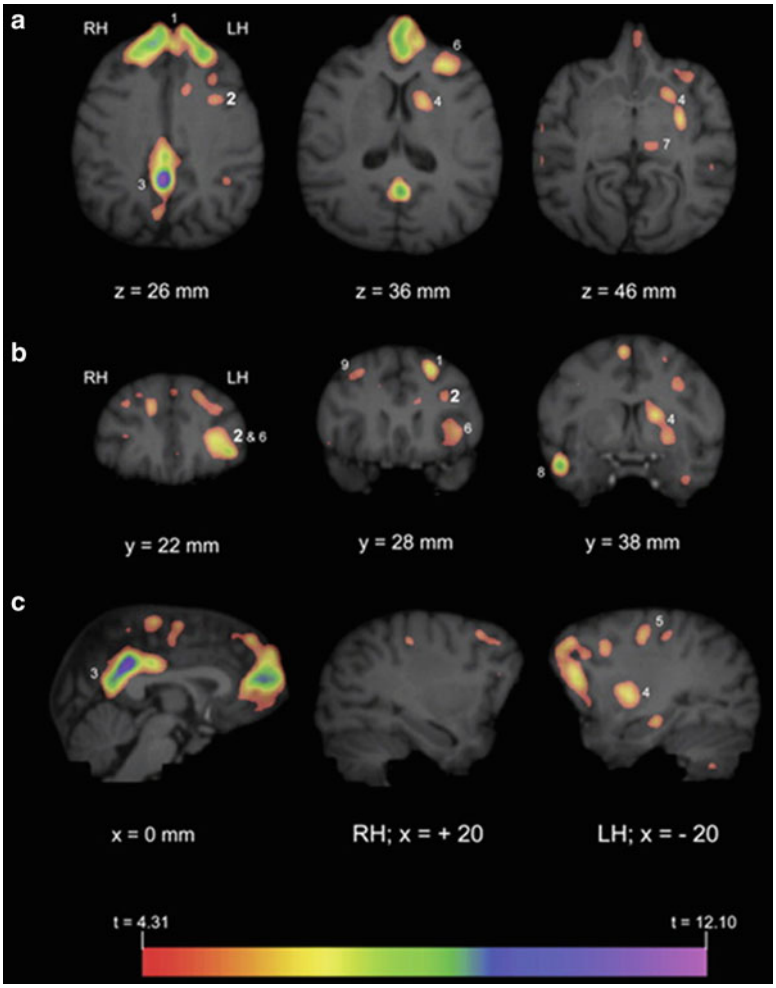


Fig. 10.6 Significant areas of activation for communicative production PET activation (GV > BL) were overlaid on MR images of representative chimpanzee brain. *x*, *y*, and *z* refer to the planes. Measurements refer to the depth from the dorsal tip of the brain (*z*, dorsal to ventral), distance from frontal pole (*y*, anterior to posterior), or distance from midsagittal (*x*, ascending positive values correspond to the right hemisphere, medial to lateral; ascending negative values correspond to the left hemisphere, medial to lateral). Panels display axial (**a**), coronal (**b**), and sagittal (**c**) views of MR images with significant GV (Gesture-Vocal) > BL (Baseline) activation. Numbers correspond to the following anatomical locations: 1, bilateral superior frontal gyrus; 2, left inferior frontal gyrus (depicted in large *bold* type); 3, bilateral posterior cingulate gyrus; 4, left caudate/putamen; 5, left medial pre- and postcentral gyrus; 6, left frontal orbital gyrus; 7, left thalamus; 8, right middle temporal gyrus; 9, right middle frontal gyrus. Note that not all areas of activation are labeled in all planes. *From Taglialatela JP*, *Russell JL*, *Schaeffer JA*, *Hopkins WD*. *Communicative signaling activates "Broca's" homologue in chimpanzees*. *Curr*. *Biol*. *2008*; *18*:*343–348* [135]

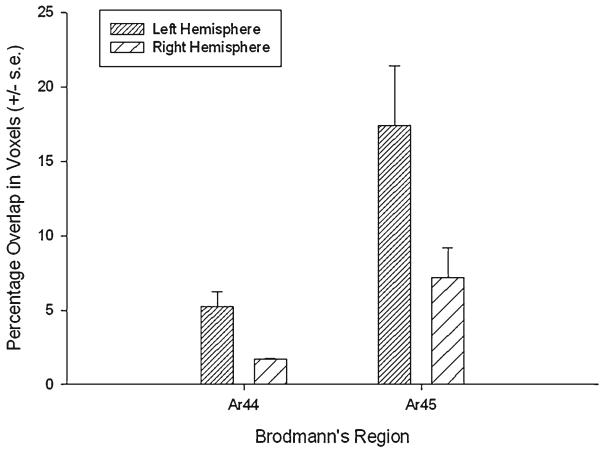


Fig. 10.7 Mean percentage (±SE) in voxels between Ar44 and Ar45 cytoarchitectonic maps with significant PET activation clusters found in chimpanzees when producing manual gestures and attention-getting vocalizations

AG sounds. In terms of the perception of the species-specific calls, Taglialatela et al. [136] found significant right hemisphere asymmetries in area Tpt in the processing of species-specific calls, particularly for proximal calls. Moreover, the magnitude of activation across all brain regions was far more pronounced when the chimpanzees were hearing proximal compared to distal calls.

Recently, my colleagues and I have attempted to examine, to what extent, the IFG regions active during the PET studies correspond to Broca's area defined by the previously described cytoarchitectonic analyses. Our interest in this question stems from similar types of voxel-of-interest-based studies in humans. For example, Horwitz et al. [138] used PET to assess metabolic activation during speech in humans. As regions of interest in their study, Horwitz et al. used previously published cytoarchitectonic maps of Ar44 and Ar45 to define the regions of interest within Broca's area in relation to the use of signs and speech. Ar45 was more active during sign and speech, whereas Ar44 was found to be active during complex articulatory movements.

To explore this association, we initially took the individual cytoarchitectonic maps of Ar44 and Ar45 for the 12 chimpanzee postmortem samples examined by Schenker et al. [119] and registered them to a template of the chimpanzee brain. Registering each individual cytoarchitectonic map to the template placed them in the same stereotaxic space. The average chimpanzee PET scans for the Taglialatela et al. [135] study were also aligned to the template and thereby also placed in the same stereotaxic space as the individual cytoarchitectonic maps. For Ar44 and Ar45, each individual cytoarchitectonic map was placed on the PET volume representing the difference in gesture-vocal activation minus the grasping condition (see Fig. 10.6). The number of significant PET voxels that fell within the Ar44 and Ar45 cytoarchitectonic maps was then calculated for each of the 12 postmortem specimens. The mean percentage of PET voxels found within Ar44 and Ar45 in the left and right hemisphere is shown in Fig. 10.7. Two observations are worth noting.

First, the percentage of active voxels in Ar45 was significantly higher than Ar44. Second, for both Ar44 and Ar45, the percentage of significantly active voxels was greater in the left compared to right hemisphere.

Conclusions

Chimpanzees are the closest living relative to humans. The fossil record does not allow for direct reconstruction of either the brain or associated behavioral or communicative repertoires of early extinct hominids; therefore, from a anthropological perspective, the chimpanzee is a model species for evaluating the neurological, cognitive, and behavioral traits of the common ancestor of humans and chimpanzees approximately 6 mya. From the standpoint of the cognitive foundations of language, chimpanzees exhibit some of the requisite skills that support linguistic functions including symbolic thought and intentional and referential communication. The manner in which chimpanzee use gestures and attention-getting sounds certainly suggests that they understand the function of communication signals and use them to bring about changes in the behavior of the recipients of those signals. Indeed, there are substantial parallels in the nonverbal communication of chimpanzees and typically developing preverbal children, at least with respect to the initiation of joint attention in the requesting domain. Sometime between 30 and 36 months in developing children, the orofacial motor skills that underlie speech emerge and this becomes the dominant mode of communication in our species, which sets us motorically apart from chimpanzees.

Among nonhuman primates, there are several features of captive and wild chimpanzee vocal communication that are unique. First, chimpanzee alarm "hoo" vocalizations are produced as a means of informing conspecifics to dangers who are otherwise ignorant. This appears to differ quite dramatically from the motivation of alarm-type vocalizations in more distantly related primates. Furthermore, captive chimpanzees can learn to make AG sounds (or learn to whistle), and they appear to use them in functionally meaningful ways. Moreover, producing AG sounds appears to be socially learned, and the evidence overwhelmingly suggests that the apes have voluntary control over their orofacial musculature and vocal apparatus. I believe that these recently documented abilities in great apes (or at least chimpanzees and orangutans) challenge many historical and some more contemporary views of primate vocalizations and facial expressions and represent some very novel findings with respect to the evolution of human speech. These collective results suggest that the common ancestor of chimpanzees and humans may have possessed and used some type of primitive sounds or grunts as a form of intraspecific communication either alone or in conjunction with a small gestural repertoire.

From a neurological standpoint, the cortical regions corresponding to Broca's area are disproportionately large in the human compared to chimpanzee brain, particularly within the left hemisphere. In comparison, area Tpt though disproportionately larger in humans compared to chimpanzees, the magnitude in relative size

differences is smaller compared to Ar44 and Ar45. Moreover, asymmetries are more consistent between human and chimpanzees for area Tpt compared to Ar44 and Ar45. These findings suggest that there has been intense selection for increased cortical expansion of Broca's area after humans split from chimpanzees and these changes likely reflect specific adaptation to selection for increasing motor and cognitive control of the orofacial musculature and peripheral speech organs in early hominids. Consistent with these observations are additional comparative findings showing that the volume and grey level index of the facial and hypoglossi nuclei in humans and apes are higher than would be predicted for species of their brain size and this likely explains the greater flexibility in facial expressions in apes compared to more distantly related monkeys [139, 140]. The fact that differences in cortical organization are found in chimpanzees that have learned to produce and use AG sounds suggests that an initial step in the evolution of language and speech in humans was the developmental of voluntary control of the orofacial musculature to produce novel sounds.

Avenues of Further Investigation

There have been substantial advances in the our understanding of the evolution of language and speech over the past 20 years, and with increasing technology and collaborative efforts, the scientific community is poised to make further advancement. There are at least four areas of research that merit continuous investigation. First, the parallel joint attention skills seen in early human preverbal communication and that of chimpanzees are quite remarkable. Raising apes in human sociolinguistic environments can enhance these abilities, but nonetheless, there is very little understanding of the neurological and genetic foundations of these abilities. Second, much has been made of alleged differences in the motivation to communicate or engage in joint attention between humans and apes. Notably, the claim that only humans engage in declarative signaling (see above) warrants further investigation. One of the main challenges with these types of comparative analyses is that, often times, humans are raised and have received very different rearing experiences and reinforcement histories compared to a chimpanzee living in a standard laboratory setting. In the absence of better experimentally controlled comparisons between the species, the claim of species differences becomes difficult to make [141]. Third, neurologically, most studies have focused on the anatomy and morphology of the homologs to Broca's and Wernicke's areas, but few have examined cortical and functional connectivity using diffusion tensor imaging or resting state fMRI [142–145]. These imaging methods provide ways to assess connectivity which may lead to novel and important insights on phylogenetic changes in the brains of humans and apes. Finally, the evidence that chimpanzees and other great apes can learn to produce AG sounds clearly challenges many historical and contemporary views on primate vocalizations. Preliminary studies further suggest that areas within Broca's area are involved in the production of AG sounds and that individual who

reliably produce these sounds show different patterns in cortical organization compared to those who do not. What is unclear from these findings is whether the changes in cortical organization and functional activation are a *consequence* of learning AG sounds or whether these apes have acquired AG sounds because their brains were already organized differently. The potential for identifying causal links between the acquisition of AG sounds and associated global and localized changes in cortical organization may potentially lead to a greater understanding of how experiential factors influence brain development and connectivity. This, in turn, may provide important clues on the evolutionary factors that shaped the primate brain, including that of humans.

Acknowledgements This research was supported in part by NIH grants NS-42867, NS-73134, HD-60563, and HD-56232. American Psychological Association and Institute of Medicine guidelines for the ethical treatment of chimpanzees in research were adhered to during all aspects of this study. I am grateful to the helpful assistance of the entire veterinary staff at the Yerkes Center for their assistance in collection of the MRI scans. The invaluable contributions of Jennifer Schaeffer, Jamie Russell, Dr. Jared Taglialatela, and Dr. Stephanie Bogart are most appreciated. Correspondence regarding this chapter can be sent to William D. Hopkins, Division of Developmental and Cognitive Neuroscience, Yerkes National Primate Research Center, 954 Gatewood Road, Atlanta, GA 30322. Email: whopkin@emory.edu.

References

1. MacNeilage PF (2008) The origin of speech. Oxford University Press, Oxford
2. Fitch WT (2000) The evolution of speech: a comparative review. Trends Cogn Sci 4:258–267
3. Hauser MD (1996) The evolution of communication. MIT Press, Cambridge
4. Tallerman M, Gibson K (eds) (2012) The Oxford handbook of language evolution. Oxford University Press, Oxford
5. Consortium TCSaA (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 437:69–87
6. Furness W (1916) Observations on the mentality of chimpanzees and orangutans. Proc Am Philos Soc 45:281–290
7. Hayes C (1951) The ape in our house. Harper, New York
8. Kellogg WN, Kellogg LA (1933) The ape and the child: a study of early environmental influence upon early behavior. McGraw-Hill, New York
9. Kellogg WN (1968) Communication and language in the home-raised chimpanzee. Science 162:423–427
10. Lieberman P (2007) The evolution of human speech: its anatomical and neural bases. Curr Anthropol 48(1):39–66
11. Gardner RA, Gardner BT (1969) Teaching sign language to a chimpanzee. Science 165(3894):664–672
12. Premack D (1971) Language in Chimpanzee? Science 172:808–822
13. Rumbaugh DM (1977) Language learning by a chimpanzee: the Lana project. Academic, New York
14. Savage-Rumbaugh ES (1986) Ape language: from conditioned response to symbol. Columbia University Press, New York
15. Terrace HS (1979) Nim. Knopf, New York
16. Miles LH (1983) Apes and language: the search for communicative competence. In: de Luce J, Wilder HT (eds) Language in primates: perspectives and implications. Springer, New York

17. Patterson FG (1978) The gestures of a gorilla: language acquisition in another pongid. Brain Lang 5:72–97
18. Savage-Rumbaugh ES, Lewin R (1994) Kanzi: the ape at the brink of the human mind. Wiley, New York
19. Zuberbuhler K (2005) The phylogenetic roots of language. Evidence from primate communication and cognition. Curr Dir Psychol Sci 14(3):126–130
20. Hewes GW (1976) The current status of the gestural theory of language origin. Ann N Y Acad Sci 280:482–504
21. Corballis MC (2002) From hand to mouth: the origins of language. Princeton University Press, Princeton
22. McNeil D (1992) Hand and mind: what gestures reveal about thought. University of Chicago Press, Chicago
23. Leavens DA, Russell JL, Hopkins WD (2010) Multi-modal communication and its social contextual use in captive chimpanzees (Pan troglodytes). Anim Cogn 13:33–40
24. Slocombe KE, Waller BM, Liebal K (2011) The language void: the need for multimodality in primate communication research. Anim Behav 81(5):919–924
25. Lyn H (2012) Apes and the evolution of language: taking stock of 40 years of research. In: Vonk J, Shackelford T (eds) Oxford handbook of comparative evolutionary psychology. Oxford University Press, Oxford
26. Savage-Rumbaugh ES, Murphy J, Sevcik RA, Brakke KE, Williams SL, Rumbaugh DM (1993) Language comprehension in ape and child. Monogr Soc Res Child Dev 58:1–256
27. Brakke KE, Savage-Rumbaugh ES (1995) The development of language skills in bonobo and chimpanzee—I. Comprehension. Lang Commun 15:121–148
28. Sevcik RA, Savage-Rumabugh ES (1994) Language comprehension and use by great apes. Lang Commun 14(1):37–58
29. Heimbauer L, Beran M, Owren M (2011) A chimpanzee recognizes synthetic speech with significantly reduced acoustic cues to phonetic content. Curr Biol 21:1210–1214
30. Call J, Tomasello M (2007) The gestural communication of monkeys and apes. Psychology Press, Oxford
31. Arbib MA, Liebal K, Pika S (2008) Primate vocalization, gesture and the evolution of human language. Curr Anthropol 49(6):1053–1076
32. Leavens DA (2012) Joint attention: twelve myths. In: Seemann A (ed) Joint attention: new developments in psychology, philosophy of mind and social neuroscience. MIT Press, Cambridge, pp 43–72
33. Iverson JM, Goldin-Meadow S (2005) Gesture paves the way for language development. Psychol Sci 16:367–371
34. Bates E, O'Connell B, Shore C (1987) Language and communication in infancy. In: Osofsky J (ed) Handbook of infant development. Wiley, New York, pp 149–203
35. Bates E, Camaioni L, Volterra V (1975) Performatives prior to speech. Merrill Palmer Q 21:205–226
36. Cartmill E, Byrne RW (2007) Orangutans modify their gestural signaling according to their audience's comprehension. Curr Biol 17:1–14
37. Liebal K, Pika S, Call J, Tomasello M (2004) To move or not to move: how apes adjust to the attentional state of others. Interact Stud 5:199–219
38. Hostetter AB, Cantero M, Hopkins WD (2001) Differential use of vocal and gestural communication by chimpanzees (Pan troglodytes) in response to the attentional status of a human (Homo sapiens). J Comp Psychol 115(4):337–343
39. Hostetter AB, Russell JL, Freeman H, Hopkins WD (2007) Now you see me, now you don't: evidence that chimpanzees understand the role of the eyes in attention. Anim Cogn 10:55–62
40. Leavens DA, Hostetter AB, Wesley MJ, Hopkins WD (2004) Tactical use of unimodal and bimodal communication by chimpanzees, Pan troglodytes. Anim Behav 67: 467–476

41. Pika S (2008) Gestures of apes and prelinguistic human children: similar or different? First Lang 28(2):116–140
42. Cartmill E, Beilock S, Goldin-Meadow S (2012) A word in the hand: action, gesture and mental representation in humans and non-human primates. Philos Trans R Soc Lond B Biol Sci 367(1585):129–143
43. Gentry E, Byrne RW (2010) Why do gorillas make sequences of gestures? Anim Cogn 13(2):287–301
44. Liebal K, Call J, Tomasello C (2004) Use of gesture sequences in chimpanzees. Am J Primatol 64:377–396
45. Hobaiter C, Byrne RW (2011) Serial gesturing by wild chimpanzees: its nature and function for communication. Anim Cogn 14(6):827–838
46. Greenfield PM, Lyn H, Savage-Rumbaugh ES (2008) Protolanguage in ontogeny and phylogeny: combining deixis and representation. Interact Stud 9(1):34–50
47. Tomasello M (2008) Origins of human communication. MIT Press, Cambridge
48. Lyn H, Greenfield PM, Savage-Rumbaugh ES, Gilliespie-Lynch K, Hopkins WD (2011) Nonhuman primates do declare! A comparison of declarative symbol and gesture use in two children, two bonobos, and a chimpanzee. Lang Commun 31:63–74
49. Leavens DA, Racine TP (2009) Joint attention in apes and humans: are humans unique? J Conscious Stud 16:240–267
50. Gentry E, Breuer T, Hobaiter K, Byrnw RW (2009) Gestural communication of the gorilla (Gorilla gorilla): repertoire, intentionality, and possible origins. Anim Cogn 12:527–546
51. Hobaiter C, Byrne RW (2011) The gestural repertoire of wild chimpanzees. Anim Cogn 14:745–767
52. Tomasello M, Call J, Nagell K, Olguin R, Carpenter M (1994) The learning and use of gestural signals by young chimpanzees: a trans-generational study. Primates 35:137–154
53. Bonnie KE, Horner V, Whiten A, de Waal FB (2007) Spread of arbitrary conventions among chimpanzees: a controlled experiment. Proc Biol Sci 274(1608):367–372
54. Whiten A, Spiteri A, Horner V et al (2007) Transmission of multiple traditions within and between chimpanzee groups. Curr Biol 17(12):1038–1043
55. Horner V, Whiten A, Flynn E, de Waal FB (2006) Faithful replication of foraging techniques along cultural transmission chains by chimpanzees and children. Proc Natl Acad Sci U S A 103(37):13878–13883
56. Van Schaik CP, Ancrenaz M, Borgen G et al (2003) Orangutan culture and the evolution of material culture. Science 299:102–205
57. Goodall J (1986) The chimpanzees of Gombe: patterns of behavior. Harvard University Press, Cambridge
58. Fossey D (1983) Gorillas in the mist. Hodder & Stoughton, London
59. Slocombe KE, Zuberbühler K (2005) Agonistic screams in wild chimpanzees (Pan troglodytes schweinfurthii) vary as a function of social role. J Comp Psychol 119(1):67–77
60. Slocombe KE, Kaller T, Turman L, Townsend SW, Papworth S, Zuberbuhler K (2010) Production of food-associated calls in wild male chimpanzees is dependent on the composition of the audience. Behav Ecol Sociobiol 64(12):1959–1966
61. Laporte MNC, Zuberbühler K (2010) Vocal greeting behaviour in wild chimpanzee females. Anim Behav 80(3):467–473
62. Crockford C, Wittig RM, Mundry R, Zuberbuhler K (2012) Wild chimpanzees inform ignorant group members of danger. Curr Biol 22:1–5
63. Seyfarth RM, Cheney DL, Marler P (1980) Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. Science 210:801–803
64. Cheney DL, Seyfarth RM (1990) Attending to behaviour versus attending to knowledge: examining monkeys' attribution of mental states. Anim Behav 40:742–753
65. Cheney DL, Seyfarth RM (1999) Mechanisms underlying the vocalizations of nonhuman primates. In: Hauser MD, Konishi M (eds) Design of animal communication. MIT Press, Cambridge

66. Seyfarth RM, Cheney DL (2010) Production, usage, and comprehension in animal vocalizations. Brain Lang 115:92–100
67. Premack D (2004) Is language the key to human intelligence? Science 303:318–320
68. Hopkins WD, Taglialatela JP, Leavens DA (2011) Do chimpanzees have voluntary control of their facial expressions and vocalizations? In: Vilain A, Schwartz J-L, Abry C, Vauclair J (eds) Primate communication and human language: vocalisation, gestures, imitation and deixis in humans and non-humans. John Benjamins, Amsterdam, pp 71–90
69. Hopkins WD, Taglialatela JP, Leavens DA (2007) Chimpanzees differentially produce novel vocalizations to capture the attention of a human. Anim Behav 73:281–286
70. Wich SA, Swartz KB, Hardus ME, Lameira AR, Stromberg E, Shumaker RW (2009) A case of spontaneous acquisition of a human sound by an orangutan. Primates 50(1):56–64
71. Taglialatela JP, Reamer L, Schapiro SJ, Hopkins WD (2012) Social learning of a communicative signal in captive chimpanzees. Biol Lett 8:498–501
72. de Waal FBM (1989) Behavioral contrasts between bonobo and chimpanzee. In: Heltne PG, Marquardt LA (eds) Understanding chimpanzees. Harvard University Press, Cambridge, pp 154–175
73. Parish AR, De Waal FBM, Haig D (2000) The other "closest living relative": how bonobos (Pan paniscus) challenge traditional assumptions about females, dominance, intra- and intersexual interactions, and hominid evolution. Ann N Y Acad Sci 907(1):97–113
74. Pollick AS, de Waal FMB (2007) Ape gestures and language evolution. Proc Natl Acad Sci U S A 104:8184–8189
75. Hopkins WD, Savage-Rumbaugh ES (1991) Vocal communication as a function of differential rearing experiences in Pan paniscus: a preliminary report. Int J Primatol 12:559–583
76. Taglialatela JP, Savage-Rumbaugh ES, Baker LA (2003) Vocal production by a language-competent Pan paniscus. Int J Primatol 24(1):1–17
77. Taglialatela JP (2004) Functional asymmetries for bonobo vocal communication [dissertation]. Department of Biology, Georgia State University, Atlanta
78. Knecht S, Drager B, Deppe M et al (2000) Handedness and hemispheric language dominance in healthy humans. Brain 123(12):2512–2518
79. Rasmussen T, Milner B (1977) The role of early left-brain injury in determining lateralization of cerebral speech function. Ann N Y Acad Sci 299:355–369
80. Annett M (2002) Handedness and brain asymmetry: the right shift theory. Psychology Press, Hove
81. Warren JM (1980) Handedness and laterality in humans and other animals. Physiol Psychol 8:351–359
82. Bradshaw JL, Rogers LJ (1993) The evolution of lateral asymmetries, language, tool use, and intellect. Academic, San Diego
83. Rogers LJ, Andrew JR (2002) Comparative vertebrate lateralization. Cambridge University Press, Cambridge
84. Strockens F, Gunturkun O, Sebastian O. Limb preferences in non-human vertebrates. Laterality (in press)
85. Fransnelli E, Vallortigara G, Rogers LJ (2012) Left-right asymmetries of behaviour and nervous system in invertebrates. Neurosci Biobehav Rev 36(4):1273–1291
86. MacNeilage PF, Rogers LJ, Vallortigara G (2009) Evolutionary origins of your right and left brain. Sci Am 301:60–67
87. Rogers LJ, Vallortigara G, Andrew RJ (2013) Divided brains: the biology and behaviour of brain asymmetries. Cambridge University Press, New York
88. Bisazza A, Cantalupo C, Robins A, Rogers LJ, Vallortigara G (1996) Right pawededness in toads. Nature 379:404
89. Güntürkün O (1997) Avian visual lateralization: a review. Neuroreport 8:3–11
90. Randler C (2007) Foot preferences during resting in wildfowl and waders. Laterality 12(2):191–197
91. Cashmore L, Uomini N, Chapelain A (2008) The evolution of handedness in humans and great apes: a review and current issues. J Anthropol Sci 86:7–35

92. Fagot J, Vauclair J (1991) Manual laterality in nonhuman primates: a distinction between handedness and manual specialization. Psychol Bull 109(1):76–89
93. Hopkins WD (2006) Comparative and familial analysis of handedness in great apes. Psychol Bull 132(4):538–559
94. MacNeilage PF, Studdert-Kennedy MG, Lindblom B (1987) Primate handedness reconsidered. Behav Brain Sci 10:247–303
95. McGrew WC, Marchant LF (1997) On the other hand: current issues in and meta-analysis of the behavioral laterality of hand function in non-human primates. Yearb Phys Anthropol 40:201–232
96. Papademetriou E, Sheu CF, Michel GF (2005) A meta-analysis of primate hand preferences for reaching and other hand-use measures. J Comp Psychol 119:33–48
97. Meguerditchian A, Vauclair J (2006) Baboons communicate with their right hand. Behav Brain Res 171:170–174
98. Meguerditchian A, Vauclair J, Hopkins WD (2010) Captive chimpanzees use their right hand to communicate with each other: implications for the origins of hemispheric specialization for language. Cortex 46:40–48
99. Hopkins WD, Pika S, Liebal K et al (2012) Handedness for manual gestures in great apes: a meta-analysis. In: Pika S, Liebal K (eds) Current developments in non-human primate gesture research. John Benjamins, Amsterdam
100. Cochet H, Vauclair J (2010) Features of spontaneous pointing gestures in toddlers. Gesture 10(1):86–107
101. Cochet H, Vauclair J (2010) Pointing gestures produced by toddlers from 15 to 30 months: different functions, hand shapes and laterality patterns. Inf Behav Develop 33:431–441
102. Corballis MC, Badzakova-Trajkov G, Haberling IS (2012) Right hand, left brain: genetic and evolutionary basis of cerebral asymmetries for language and manual action. Wiley Interdiscip Rev Cogn Sci 3:1–17
103. Hook-Costigan MA, Rogers LJ (1998) Lateralized use of the mouth in production of vocalizations by marmosets. Neuropsychologia 36(12):1265–1273
104. Wallez C, Vauclair J (2011) Right hemisphere dominance for emotion processing in baboons. Brain Cogn 75(2):164–169
105. Wallez C, Vauclair J (2012) First evidence of population-level oro-facial asymmetries during the production of distress calls by macaque (Macaca mulatta) and baboon (Papio anubis) infants. Behav Brain Res 234(1):69–75
106. Fernandez-Carriba S, Loeches A, Morcillo A, Hopkins WD (2002) Asymmetry in facial expression of emotions by chimpanzees. Neuropsychologia 40(9):1523–1533
107. Losin ER, Freeman H, Russell JL, Meguerditchian A, Hopkins WD (2008) Left hemisphere specialization for oro-facial movements of learned vocal signals by captive chimpanzees. PLoS One 3:1–7
108. Kimura D (1993) Neuromotor mechanisms in human communication. Oxford University Press, Oxford
109. Hopkins WD, Cantero M (2003) From hand to mouth in the evolution of language: the influence of vocal behavior on lateralized hand use in manual gestures by chimpanzees (Pan troglodytes). Dev Sci 6(1):55–61
110. Keller SS, Crow TJ, Foundas AL, Amunts K, Roberts N (2009) Broca's area: nomenclature, anatomy, typology and asymmetry. Brain Lang 109:29–48
111. Keller SS, Roberts N, Hopkins WD (2009) A comparative magnetic resonance imaging study of the anatomy, variability and asymmetry of Broca's area in the human and chimpanzee brain. J Neurosci 29:14607–14616
112. Keller SS, Deppe M, Herbin M, Gilissen E (2012) Variability and asymmetry of the sulcal contours defining Broca's area homologue in the chimpanzee brain. J Comp Neurol 520: 1165–1180
113. Foundas A, Leonard C, Heilman K (1995) Morphological cerebral asymmetries and handedness: the pars triangularis and planum temporale. Arch Neurol 52:501–508

114. Foundas AL, Eure KF, Luevano LF, Weinberger DR (1998) MRI asymmetries of Broca's area: the pars triangularis and pars opercularis. Brain Lang 64(3):282
115. Tomaiuolo F, MacDonald JD, Caramanos Z et al (1999) Morphology, morphometry and probability mapping of the pars opercularis of the inferior frontal gyrus: an in vivo MRI analysis. Eur J Neurosci 11:3033–3046
116. Knaus TA, Corey DM, Bollich AM, Lemen LC, Foundas AL (2007) Anatomical asymmetries of anterior perisylvian speech-language regions. Cortex 43:499–510
117. Amunts K, Lenzen M, Friederici AD et al (2010) Broca's region: novel organizational principles and multiple receptor mapping. PLoS Biol 8(9):1000489
118. Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings HB, Zilles K (1999) Broca's region revisited: cytoarchitecture and intersubject variability. J Comp Neurol 412(2):319–341
119. Schenker NM, Hopkins WD, Spocter MA et al (2010) Broca's area homologue in chimpanzees (Pan troglodytes): probabilistic mapping, asymmetry and comparison to humans. Cereb Cortex 20:730–742
120. Sherwood CC, Baurernfeind AL, Bianchi S, Raghanti MA, Hof PR (2012) Human brain evolution writ large and small. In: Hofman MA, Falk D (eds) Progress in brain research. Elsevier, New York
121. Uylings H, Jacobsen A, Zilles K, Amunts K (2006) Left-right asymmetry in volume and number of neurons in adult Broca's area. Cortex 42(4):652–658
122. Beaton AA (1997) The relation of planum temporale asymmetry and morphology of the corpus callosum to handedness, gender and dyslexia: a review of the evidence. Brain Lang 60:255–322
123. Josse G, Mazoyer B, Crivello F, Tzourio-Mazoyer N (2003) Left planum temporale: an anatomical marker of left hemispheric specialization for language comprehension. Cogn Brain Res 18:1–14
124. Shapleske J, Rossell SL, Woodruff PW, David AS (1999) The planum temporale: a systematic, quantitative review of its structural, functional and clinical significance. Brain Res Rev 29:26–49
125. Knaus TA, Bollich AM, Corey DM, Lemen LC, Foundas AL (2006) Variability in perisylvian brain anatomy in healthy adults. Brain Lang 97:219–232
126. Gannon PJ, Holloway RL, Broadfield DC, Braun AR (1998) Asymmetry of chimpanzee planum temporale: humanlike pattern of Wernicke's language area homolog. Science 279:220–222
127. Hopkins WD, Nir T (2010) Planum temporale surface area and grey matter asymmetries in chimpanzees (Pan troglodytes): the effect of handedness and comparison within findings in humans. Behav Brain Res 208(2):436–443
128. Gilissen E (2001) Structural symmetries and asymmetries in human and chimpanzee brains. In: Falk D, Gibson KR (eds) Evolutionary anatomy of the primate cerebral cortex. Cambridge University, Cambridge, pp 187–215
129. Galaburda AM, Sanides F (1980) Cytoarchitectonic organization of the human auditory cortex. J Comp Neurol 190:597–610
130. Spocter MA, Hopkins WD, Garrison AR et al (2010) Wernicke's area homolog in chimpanzees (Pan troglodytes): probabilistic mapping, asymmetry and comparison with humans. Proc R Soc Lond B Biol Sci 277:2165–2174
131. Gannon PJ, Kheck N, Hof PR (2008) Leftward interhemispheric asymmetry of macaque monkey temporal lobe language area homolog is evident at the cytoarchitectural, but not gross anatomic level. Brain Res 1199:62–73
132. Lyn HL, Pierre P, Bennett AJ, Fears SC, Woods RP, Hopkins WD (2011) Planum temporale grey matter asymmetries in chimpanzees (Pan troglodytes), vervet (Chlorocebus aethiops sabaeus), rhesus (Macaca mulatta) and bonnet (Macaca radiata) monkeys. Neuropsychologia 49:2004–2012
133. Taglialatela JP, Cantalupo C, Hopkins WD (2006) Gesture handedness predicts asymmetry in the chimpanzee inferior frontal gyrus. Neuroreport 17(9):923–927

134. Cantalupo C, Oliver J, Smith J, Nir T, Taglialatela JP, Hopkins WD (2009) The chimpanzee show human-like perisylvian asymmetries in white matter. Eur J Neurosci 30:431–438
135. Taglialatela JP, Russell JL, Schaeffer JA, Hopkins WD (2008) Communicative signaling activates "Broca's" homologue in chimpanzees. Curr Biol 18:343–348
136. Taglialatela JP, Russell JL, Schaeffer JA, Hopkins WD (2009) Visualizing vocal perception in the chimpanzee brain. Cereb Cortex 19(5):1151–1157
137. Taglialatela JP, Russell JL, Schaeffer JA, Hopkins WD (2011) Chimpanzee vocal signaling points to a multimodal origin of human language. PLoS One 6:e18852
138. Horwitz B, Amunts K, Bhattacharyya R et al (1999) Activation of Broca's area during the production of spoken and signed language: a combined cytoarchitectonic mapping and PET analysis. Neuropsychologia 41(14):1868–1876
139. Sherwood CC, Hof PR, Holloway RL et al (2005) Evolution of the brainstem orofacial motor system in primates: a comparative study of trigeminal, facial and hypoglossal nuclei. J Hum Evol 48:45–84
140. Sherwood CC, Holloway RL, Erwin JM, Hof PR (2004) Cortical orofacial motor representation in Old Word monkeys, great apes and humans: II. Sterological analysis of chemoarchitecture. Brain Behav Evol 63:82–106
141. Leavens DA, Bard KA (2011) Environmental influences on joint attention in great apes: implications for human cognition. J Cogn Educ Psychol 10:9–31
142. Glasser MF, Rilling JK (2008) DTI tractography of the human brain's language pathways. Cereb Cortex 18(11):2471–2482
143. Li L, Preuss TM, Rilling JK et al (2009) Chimpanzee pre-central corticospinal system asymmetry and handedness: a diffusion magnetic resonance imaging study. PLoS One 5(9):e12886
144. Rilling JK, Glasser MF, Jbabdi S, Andersson J, Preuss TM (2012) Continuity, divergence and the evolution of brain language pathways. Front Evol Neurosci 3(11):1–6
145. Rilling JK, Glasser MF, Preuss TM et al (2008) The evolution of the arcuate fasciculus revealed with comparative DTI. Nat Neurosci 11:426–428
146. Hopkins WD, Russell JL, Freeman H, Buehler N, Reynolds E, Schapiro SJ (2005) The distribution and development of handedness for manual gestures in captive chimpanzees (Pan troglodytes). Psychol Sci 16(6):487–493

Index

S.A. Helekar (ed.), *Animal Models of Speech and Language Disorders,*
DOI 10.1007/978-1-4614-8400-4, © Springer Science+Business Media New York 2013