



Published in final edited form as:

Curr Opin Neurobiol. 2020 October ; 64: 24–31. doi:10.1016/j.conb.2020.01.004.

The neurobiology of innate and learned vocalizations in rodents and songbirds

Richard Mooney

Department of Neurobiology, Duke University School of Medicine, Durham, NC 27705, United States

Abstract

Vocalizations are an important medium for sexual and social signaling in mammals and birds. In most mammals other than humans, vocalizations are specified by innate mechanisms and develop normally in the absence of auditory experience. By contrast, juvenile songbirds memorize and copy the songs of adult tutors, a process with many parallels to human speech learning. Despite the centrality of vocal learning to human speech, vocal production in humans as well as in songbirds exploits ancestral circuitry for innate vocalizations, and effective vocal communication depends on the fluent blending of innate and learned elements. This review covers recent advances in our understanding of central mechanisms for learned and innate vocalizations in birds and mice, including brainstem mechanisms that help to ‘gate’ vocalizations on or off, cortical involvement in learned and innate vocalizations, and the delineation of circuits that evaluate and reinforce song performance to facilitate vocal learning.

Introduction

Learned versus innate courtship vocalizations

Vocalizations are produced in many different contexts, ranging from social and sexual encounters to painful stimuli and physical threat. Here the focus is confined to courtship vocalizations, or songs, which are often produced by males to attract and court females and to defend territory from other males. Songbirds produce courtship songs comprising several to tens of individual syllables, each 50–150 ms in duration, and organized into stereotyped sequences referred to as a phrase, or motif, that lasts from several hundreds of milliseconds to tens of seconds [1,2]. Juvenile songbirds learn the acoustic pattern and order of syllables in their motif by memorizing and copying songs produced by one or more tutors, constituting a paradigmatic example of a sensitive period-limited form of vocal learning [2-8]. As adults, males (and in some species, females) sing in response to visual and auditory cues provided by other birds, and also as a function of the singer’s reproductive state. Although many features of birdsong are learned, innate vocal elements can figure prominently in the song display, serving as introductory notes at the onset of song or as

Corresponding author: Mooney, Richard (mooney@neuro.duke.edu).

Conflict of interest statement

Nothing declared.

passing notes that link together several renditions of the song motif within a longer song bout [2,9].

Mice produce an extensive vocal repertoire including both audible (<20 kHz) and ultrasonic elements. In response to the presence of nearby females or female odorants, adult male mice produce ultrasonic vocalizations (USVs) comprising individual syllables of 5–200 ms duration, organized into highly variable sequences lasting one to several seconds [10,11]. Unlike birdsong, the acoustic patterns of these ‘songs’ are specified by innate mechanisms, as shown by the finding that mice genetically engineered to be deaf produce a normal USV repertoire [12,13]. Furthermore, whereas male birdsong is necessary for successful courtship and mating, the USVs produced by male mice may render him more attractive to the female but appear to be unnecessary for successful mating, at least in a laboratory setting [14,15]. Thus, the courtship songs of bird and mice differ in their learned versus innate qualities and the extent to which they serve a crucial role in courtship.

Peripheral mechanisms of song production: voiced versus whistled

In addition to their learned versus innate qualities, the songs of bird and mice also differ in their peripheral mechanisms of production. Similar to the vocal cord vibrations that give rise to human speech, birdsong arises from ‘voiced’ oscillatory vibrations of membranes in the syrinx, a bipartite structure situated caudal to the larynx and where the trachea bifurcates into the two bronchial tubes [16]. By contrast, mouse and other rodent USVs result from an aerodynamic whistle formed in the larynx [17]. Indeed, a recent structural MRI analysis revealed that murine rodents that produce USVs possess a specialized ventral laryngeal pouch that is the likely source of this whistle [18]. Despite their contrasting peripheral mechanics, both birdsong and rodent USVs require precise integration of respiratory and vocal muscle activity, and in either case respiratory pressure must be carefully regulated to maintain either voiced or whistled outputs [19–21]. Thus, central mechanisms for producing courtship vocalizations in birds and mice are likely to share many features and ultimately must solve similar problems to control and integrate respiratory and vocal muscle activity.

Neural mechanisms for singing and countersinging in birds and mice

Because vocalization requires integrated activity of vocal and respiratory muscles, the terminal neurons in central circuits for vocalization are the respiratory and vocal motor neuron pools in the spinal cord and brainstem. The production of innate vocalizations in birds and mammals is under the control of pattern generating circuitry in the rostroventral lateral medulla (RVL) that accesses these motor neuron pools and that in turn is gated by descending projections from neurons in the caudolateral part of the midbrain periaqueductal gray (PAG) [22–25]. In many mammals, these brainstem pathways are sufficient to produce a normal vocal repertoire [26,27]. In mammals, descending pathways from the cingulate cortex and from subcortical structures, including the hypothalamus, act via the PAG to regulate vocalization as a function of social context and reproductive state [24]. Bilateral lesions of the PAG abolish innate vocalizations in non-human primates and other mammals, and damage to a similar region in human subjects results in the loss of speech as well as innate emotional vocalizations, such as laughter and crying [24,28]. (Although an early study found that ‘vocal’ PAG lesions in canaries did not prevent singing [29], whether the

PAG serves an obligatory role for vocalization in songbirds warrants further investigation.) As the PAG is a complex structure that serves a wide variety of functions, including nociception, reproduction, defense, and autonomic regulation [30-33], a longstanding challenge has been to determine whether vocalization-specific neurons exist in the PAG and, if so, to disentangle them from the heterogeneous network of other PAG neurons, in which they are embedded. In fact, the recent application of genetic tagging methods has identified a population of PAG neurons in the mouse specifically involved in gating USVs, enabling genetic tracing and manipulation experiments that will provide new light into how the PAG integrates a wide variety of information and gates downstream vocal pattern generator circuitry [34*].

Given that most mammals can produce vocalizations without cortical input [24], an exceptional quality of birdsong (and human speech) is the remarkable degree to which cortical circuits contribute to vocal patterning. In songbirds, the telencephalic nucleus HVC, which may be considered as an analogue to either the primary or premotor motor cortex in mammals, is essential to the production of learned but not innate vocalizations [9,29]. A variety of correlative and causal studies support the idea that HVC neurons that project to the song motor nucleus RA (HVC_{RA} neurons) encode temporal aspects of song by firing in a temporally precise sequence during song [35-38]. In fact, bilaterally cooling HVC slows song timing without altering song's spectral features [37,39], which has been interpreted as evidence that HVC_{RA} neurons form a synfire chain that encodes both global and local aspects of song timing. While other regions outside of HVC also contribute to song timing [39-41], there is widespread consensus that the sequential activity of HVC_{RA} neurons is an essential component of the song timing mechanism.

Of course, sequential activity emanating from HVC must be read out by downstream elements of the vocal respiratory network, including the RVL. Moreover, just as the capacity to vocalize is built on a more ancient capacity to breathe, the capacity for producing learned vocalizations is presumably built upon more ancestral circuitry for producing innate vocalizations. Indeed, in addition to containing respiratory premotor neurons, the RVL contains neurons that give rise to a recurrent pathway that ultimately innervates HVC [42]. Presumably, this pathway allows respiratory and possibly vocal pattern generating circuits in the brainstem to modulate HVC activity during singing. Given this recurrent circuitry, current debate has centered on the extent to which birdsong results from a strict top-down influence of HVC on 'subservient' brainstem machinery, or instead involves more of a reciprocal interaction between HVC and the brainstem.

Part of the answer to this debate rests on obtaining a more complete picture of how HVC_{RA} neurons function at a population level during singing. Experimental constraints are that HVC_{RA} neurons are notoriously difficult to isolate and 'hold' with an electrode in singing birds and that there are ~40 000 of them in each HVC. A recent study using widefield calcium imaging from identified populations of HVC_{RA} neurons has proven to be a game-changer, not only confirming the canonical sequential activity of a subset of HVC_{RA} neurons during song production but also revealing that other HVC_{RA} cells undergo prolonged ramping activity beginning hundreds of milliseconds or even seconds before song onset [43*]. Moreover, similar pre-song increases in activity are also detected in RA and in the

respiratory system, suggesting that increased activity in HVC_{RA} ‘pre-song’ neurons prepares the respiratory system for the transition from silence to song, which may then act via recurrent projections to trigger sequential firing in HVC_{RA} ‘song’ neurons. Whether HVC_{RA} ‘pre-song’ neurons are necessary to singing remains to be determined, but their existence suggests that they sit at or near the point where the decision to sing is being made in the brain. The rich and complex convergence of sensory, sensorimotor, and neuromodulatory inputs to HVC provide numerous candidates for providing information salient to that decision.

Consistent with the idea that complex networks give rise to neurons with complex properties, some HVC neurons that display song premotor activity also respond to playback of the bird’s own song and other acoustically similar songs [44-46]. In certain cases, the playback-evoked and premotor activity patterns of individual HVC neurons are highly similar, suggesting that HVC functions as part of an inverse model to convert auditory stimuli into a song ‘action.’ In support of this idea, prior studies in anesthetized or sleeping birds established that song playback can sufficiently excite HVC so as to entrain respiration and syringeal motor neurons and muscles [47,48], hinting that auditory stimulation can effectively engage the song motor machinery. A recent study by Bush *et al.* provides direct support for this idea, showing that playback-evoked syringeal EMGs recorded during sleep are highly similar to the syringeal EMGs recorded from the same bird during singing [49**]. Furthermore, their clever use of synthetic stimuli showed that even acoustically degraded songs could drive syringeal EMG response patterns in an all or none fashion, albeit with a lower response probability, behavior that is consistent with an attractor network [50].

Such robust auditory to motor transformations are especially well-suited to the antiphonal (call and response) vocal behaviors exhibited by many songbirds, primates, and even certain rodent species, because they can help trigger vocal production in response to a partner’s vocalizations, thus avoiding temporal overlap between the two signals. This capacity may seem simple, but it is essential to our own spoken conversation and requires careful monitoring of the partner’s vocal signal as well as the auditory feedback created by one’s own voice. The duet songs of male and female white-browed sparrow weavers involve a series of introductory notes produced by one partner followed by a rapid and precise alternating exchange of duetting syllables between the partners [51]. In a heroic field study conducted in the African savannah, Hoffman *et al.* fitted male and female sparrow weavers with radio transmitter backpacks, head-fixed miniature microphones and microelectrodes in HVC [52*]. The acoustic recordings reveal that the rate at which introductory notes are sung (~4 Hz) are highly similar to the rate at which the duetting syllables are produced, requiring the initiating bird to halve its syllable rate once the other bird joins the duet, and also requiring the responding partner to join the duet at a rate half of that produced by the initiator. An attractive idea is that this is accomplished by auditory entrainment of HVC premotor activity, and in fact sparrow weavers will duet in response to song playback. Furthermore, when one bird produces an introductory note sequence, its HVC neurons burst ~4 Hz, but then drop to ~2 Hz once the other bird sings, possibly because auditory stimulation generated by the responding partner entrains the timing of HVC premotor activity in the initiating bird.

In contrast to songbirds and humans, where cortical lesions can completely abolish learned vocalizations, decorticate male laboratory mice still produce a normal USV repertoire in response to a nearby female [53]. Thus for USVs produced by mice, as with vocalizations produced by most other mammals including non-human primates, the motor cortex does not play any obvious role in vocal patterning. Nonetheless, classical studies in monkeys have highlighted a pathway from the anterior cingulate cortex to the periaqueductal gray that is important in the volitional control of vocalization as a function of social motivation [24]. This pathway appears to be a well-conserved feature of the mammalian brain as a recent study established that electrical stimulation in the posterior prelimbic cortex (PLC) of the rat (a region that includes the cingulate cortex) can reliably trigger USVs as well as other vocalizations [54]. Moreover, transynaptic tracing revealed that PLC neurons in this vocal hotspot project to regions of the PAG that in turn innervate the RVL, providing a route via which the cortex could modulate vocalization. While the behavioral relevance of this pathway remains uncertain, an intriguing recent study in another murid rodent, the central American singing mouse, reveals a prominent role for the orofacial motor cortex (OFC) in regulating vocalization as a function of social and auditory context [55**]. In this species, males sing either spontaneously when socially isolated or antiphonally when they encounter each other, and these antiphonal songs are precisely controlled to minimize temporal overlap. Notably, this antiphonal singing can also be elicited by song playback, underscoring that auditory rather than social cues play a prominent role in triggering the antiphonal behavior. When the OFC is inactivated with muscimol, males continue to sing spontaneously but not in response to playback, pointing to the cortex as a key element that helps to recruit subcortical vocal pattern generators to facilitate antiphonal singing. As the birdsong nucleus RA plays a similar role in regulating the timing of innate antiphonal calls [56], a likely idea is that cortical neurons that provide input to brainstem vocal gating and patterning networks also help to bridge auditory and motor systems to enable adaptive and flexible forms of vocalization, as required for antiphonal vocal behaviors.

Circuit mechanisms for birdsong learning

A related idea is that such an auditory to motor bridge provides the foundation for the even more flexible sensorimotor interactions necessary to vocal learning. In songbirds, HVC is a major site where auditory and motor systems for song converge. Prior studies established that exposure to a singing tutor rapidly potentiates and stabilizes synapses onto HVC neurons in a juvenile "pupil" [57,58], supporting the idea that auditory synapses on HVC neurons provide part of the substrate for encoding tutor song memories. An innovative recent study directly tested this idea by optogenetically stimulating sensorimotor inputs to HVC in tutor-naïve juvenile zebra finches in different temporal patterns [59**]. Different juveniles 'trained' with different patterns grew up to produce songs with temporal characteristics similar to the pattern of stimulation they were subjected to when young, consistent with the idea that artificial stimulation of the sensorimotor input to HVC is sufficient to drive song copying in the absence of any tutor.

In nature, however, juvenile zebra finches learn most readily by socializing with singing tutors, and copy little or not at all from songs played through a speaker [60,61]. Thus, the songbird's brain presumably integrates both social and auditory cues from the singing tutor,

providing a failsafe mechanism to ensure that copying is made from a socially suitable model. Part of this integrator involves dopamine neurons in the midbrain A11 cell group that provide input to HVC: In juveniles, these neurons are strongly activated during encounters with a singing tutor, but not when the juvenile encounters non-singing birds or simply listens to song playback [62**]. Furthermore, dopamine release into the pupil's HVC is necessary for song copying from a live tutor and pairing dopamine terminal stimulation with song playback promotes copying even in the absence of a live tutor. Lastly, during an initial tutoring session, dopamine acts in HVC to promote temporally precise auditory-evoked responses to the tutor song, similar to the temporally precise premotor activity that typifies HVC neuron activity in singing birds. Therefore, the integration of social and auditory information in HVC rapidly builds a circuit that translates auditory stimuli into a motor framework, ultimately promoting song copying (Figure 1).

Forming an auditory memory of a tutor song is only half the battle, though, as the juvenile must then translate this memory into a suitable vocal copy. A decades-old idea is that dopamine inputs from the ventral tegmental area and substantia nigra pars compacta (VTA/SNc) to a song-specialized region of the basal ganglia (Area X) adaptively reinforce song renditions that more closely match the song memory [63]. This longstanding idea has finally gained solid experimental support, primarily from studies using adult birds exposed to syllable-triggered noise. In this protocol, noise delivered as a function of a target syllable's pitch (i.e. fundamental frequency) slowly drives a shift in the syllable's pitch away from the region targeted by noise, a process referred to as pitch learning [64]. Importantly, in a regime where singing triggers intermittent noise, VTA neurons that project to Area X (VTA_X neurons) 'learn' to encode reward prediction error, a key element in reinforcement learning mechanisms [65]. Moreover, optogenetic manipulation of VTA_X terminal activity is sufficient to drive pitch learning, with optogenetic excitation increasing the probability that a syllable paired with stimulation will be retained in the repertoire and with suppression decreasing this probability [66,67]. An important caveat is that these studies were performed in adult birds that sing otherwise stable songs, not in juveniles in the process of copying a tutor song. However, both juvenile copying and pitch learning depend on VTA/SNc neurons and dopamine release into Area X [66,68], suggesting that they do involve a common underlying mechanism.

A remaining issue is how VTA_X neurons 'know' which song renditions are good copies and thus should be reinforced, and which renditions are poor copies and should be discarded. Ultimately, this evaluation depends on the juvenile's ability to compare singing-related auditory feedback with a memory of the tutor song. Although the circuit basis for this comparison remains a mystery, a reasonable guess is that the circuit spans the region between HVC and the VTA/SNc complex. Two structures immediately upstream of the VTA are the ventral pallidum (VP) and the intermediate ventral arcopallium (Aiv) [69,70], the latter of which may be analogous to higher levels of the auditory cortex [71]. In fact, an elegant prior study established that Aiv neurons that project to the VTA (Aiv_{VTA} neurons) detect distortions in singing-related auditory feedback, and thus are well suited to detect vocal errors [72]. A recent study tested this idea by combining noise-evoked pitch learning with optogenetic stimulation of Aiv_{VTA} terminals, revealing that rendition-by-rendition variations in Aiv_{VTA} terminal activity during auditory feedback but not premotor windows

associated with the target syllable were necessary to drive pitch learning [73**]. Moreover, pitch-contingent optogenetic stimulation of Aiv_{VTA} terminals was sufficient to negatively reinforce syllable pitch, similar to the effects of pitch-contingent noise. This same study established that VP_{VTA} terminals also function as part of the auditory feedback-dependent evaluation circuit, but that their activity positively reinforced syllable pitch, paralleling another recent finding that a subset of VP_{VTA} neurons can respond to ‘good’ performances and some also encode positive reward prediction error [74**]. These various findings advance a model, in which Aiv and VP provide information that collectively steers the VTA to reinforce song performance, an arrangement that may optimize song learning rates (Figure 2). Consistent with the idea that Aiv and VP function cooperatively to drive learning, the learning rates achieved with pitch contingent optogenetic stimulation of Aiv_{VTA} or VP_{VTA} terminals, or by optogenetic excitation of VTA_X terminals, are approximately half that measured in response to pitch-contingent noise [66,73**]. Thus, the fastest learning presumably occurs when VTA_X neuron activity can be bidirectionally modulated, a process that most likely requires the coordinated and possibly parallel actions of both Aiv and VP.

Acknowledgements

The author received helpful comments on an earlier version of the manuscript from Drs Mor Ben-Tov (Duke University) and Steve Shea (CSHL). The work was supported by funding to the author from N.I.H. R01DC013826, R01MH117778, R01NS099288, and NSF 1354962.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
 - of outstanding interest
1. Greenwalt C: *Birdsong: Acoustics and Physiology*. Washinton D. C: Smithsonian University Press; 1968.
 2. Immelmann K: Song development in zebra finch and other Estrildid finches In *Bird Vocalisations*. Edited by Hinde RA. London: Cambridge University Press; 1969:61–74.
 3. Funabiki Y, Konishi M: Long memory in song learning by zebra finches. *J Neurosci* 2003, 23:6928–6935. [PubMed: 12890787]
 4. Konishi M: The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Z Tierpsychol* 1965, 22:770–783. [PubMed: 5874921]
 5. Marler P, Peters S: Sparrows learn adult song and more from memory. *Science* 1981, 213:780–782. [PubMed: 17834587]
 6. Marler P, Tamura M: Culturally transmitted patterns of vocal behavior in sparrows. *Science* 1964, 146:1483–1486. [PubMed: 14208581]
 7. Marler P, Waser MS: Role of auditory feedback in canary song development. *J Comp Physiol Psychol* 1977, 91:8–16. [PubMed: 838918]
 8. Thorpe W: The learning of song patterns by birds, with especial reference to the song of the chaffinch. *Ibis* 1958, 100:535–570.
 9. Aronov D, Andalman AS, Fee MS: A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science* 2008, 320:630–634. [PubMed: 18451295]
 10. Holy TE, Guo Z: Ultrasonic songs of male mice. *PLoS Biol* 2005, 3:e386. [PubMed: 16248680]
 11. Sewell GD: Ultrasound in adult rodents. *Nature* 1967, 215:512.

12. Hammerschmidt K, Reisinger E, Westekemper K et al.: Mice do not require auditory input for the normal development of their ultrasonic vocalizations. *BMC Neurosci* 2012, 13:40. [PubMed: 22533376]
13. Mahrt EJ, Perkel DJ, Tong L et al.: Engineered deafness reveals that mouse courtship vocalizations do not require auditory experience. *J Neurosci* 2013, 33:5573–5583. [PubMed: 23536072]
14. Nunez AA, Pomerantz SM, Bean NJ, Youngstrom TG: Effects of laryngeal denervation on ultrasound production and male sexual behavior in rodents. *Physiol Behav* 1985, 34:901–905. [PubMed: 4059379]
15. Pomerantz SM, Nunez AA, Bean NJ: Female behavior is affected by male ultrasonic vocalizations in house mice. *Physiol Behav* 1983, 31:91–96. [PubMed: 6685321]
16. Suthers RA, Goller F, Pytte C: The neuromuscular control of birdsong. *Philos Trans R Soc Lond B Biol Sci* 1999, 354:927–939. [PubMed: 10382225]
17. Roberts LH: The rodent ultrasound production mechanism. *Ultrasonics* 1975, 13:83–88. [PubMed: 1167711]
18. Riede T, Borgard HL, Pasch B: Laryngeal airway reconstruction indicates that rodent ultrasonic vocalizations are produced by an edge-tone mechanism. *R Soc Open Sci* 2017, 4:170976. [PubMed: 29291091]
19. Goller F, Suthers RA: Role of syringeal muscles in gating airflow and sound production in singing brown thrashers. *J Neurophysiol* 1996, 75:867–876. [PubMed: 8714659]
20. Riede T: Peripheral vocal motor dynamics and combinatorial call complexity of ultrasonic vocal production in rats In *Handbook of Ultrasonic Vocalizations: A Window into the Emotional Brain*. Edited by Brudzynski SM. Elsevier; 2018:45–60.
21. Riede T, Goller F: Peripheral mechanisms for vocal production in birds - differences and similarities to human speech and singing. *Brain Lang* 2010, 115:69–80. [PubMed: 20153887]
22. Hage SR, Jurgens U: On the role of the pontine brainstem in vocal pattern generation: a telemetric single-unit recording study in the squirrel monkey. *J Neurosci* 2006, 26:7105–7115. [PubMed: 16807339]
23. Jurgens U, Hage SR: On the role of the reticular formation in vocal pattern generation. *Behav Brain Res* 2007, 182:308–314. [PubMed: 17173983]
24. Jurgens U: The neural control of vocalization in mammals: a review. *J Voice* 2009, 23:1–10. [PubMed: 18207362]
25. Wild JM: The avian nucleus retroambiguus: a nucleus for breathing, singing and calling. *Brain Res* 1993, 606:319–324. [PubMed: 8387862]
26. Jurgens U: The role of the periaqueductal grey in vocal behaviour. *Behav Brain Res* 1994, 62:107–117. [PubMed: 7945960]
27. Jurgens U: Neural pathways underlying vocal control. *Neurosci Biobehav Rev* 2002, 26:235–258. [PubMed: 11856561]
28. Esposito A, Demeurisse G, Alberti B, Fabbro F: Complete mutism after midbrain periaqueductal gray lesion. *Neuroreport* 1999, 10:681–685. [PubMed: 10208530]
29. Nottebohm F, Stokes TM, Leonard CM: Central control of song in the canary, *Serinus canarius*. *J Comp Neurol* 1976, 165:457–486. [PubMed: 1262540]
30. Carrive P: The periaqueductal gray and defensive behavior: functional representation and neuronal organization. *Behav Brain Res* 1993, 58:27–47. [PubMed: 8136048]
31. Evans DA, Stempel AV, Vale R et al.: A synaptic threshold mechanism for computing escape decisions. *Nature* 2018, 558:590–594. [PubMed: 29925954]
32. Holstege G: The periaqueductal gray controls brainstem emotional motor systems including respiration. *Prog Brain Res* 2014, 209:379–405. [PubMed: 24746059]
33. Li Y, Zeng J, Zhang J et al.: Hypothalamic circuits for predation and evasion. *Neuron* 2018, 97:911–924 e915. [PubMed: 29398361]
- 34. Tschida K, Michael V, Takatoh J et al.: A specialized neural circuit gates social vocalizations in the mouse. *Neuron* 2019, 103:459–472 e454. [PubMed: 31204083] The authors use an activity-dependent intersectional genetic tagging method to identify PAG neurons that are selectively activated when male mice produce courtship USVs. Optogenetic and chemogenetic

manipulations establishes that the activity of these PAG-USV neurons is necessary and sufficient to trigger USVs, and genetic tracing maps their afferent and efferent connections, providing an important step forward in understanding the neural basis of vocal control.

35. Fee MS, Kozhevnikov AA, Hahnloser RH: Neural mechanisms of vocal sequence generation in the songbird. *Ann N Y Acad Sci* 2004, 1016:153–170. [PubMed: 15313774]
36. Hahnloser RH, Kozhevnikov AA, Fee MS: An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* 2002, 419:65–70. [PubMed: 12214232]
37. Long MA, Fee MS: Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* 2008, 456:189–194. [PubMed: 19005546]
38. Long MA, Jin DZ, Fee MS: Support for a synaptic chain model of neuronal sequence generation. *Nature* 2010, 468:394–399. [PubMed: 20972420]
39. Hamaguchi K, Tanaka M, Mooney R: A distributed recurrent network contributes to temporally precise vocalizations. *Neuron* 2016, 91:680–693. [PubMed: 27397518]
40. Ashmore RC, Wild JM, Schmidt MF: Brainstem and forebrain contributions to the generation of learned motor behaviors for song. *J Neurosci* 2005, 25:8543–8554. [PubMed: 16162936]
41. Goldin MA, Alonso LM, Allende JA et al.: Temperature induced syllable breaking unveils nonlinearly interacting timescales in birdsong motor pathway. *PLoS One* 2013, 8:e67814. [PubMed: 23818988]
42. Wild J, Williams M, Suthers R: Neural pathways for bilateral vocal control in songbirds. *J Comp Neurol* 2000, 423:413–426. [PubMed: 10713571]
- 43. Daliparthi VK, Tachibana RO, Cooper BG et al.: Transitioning between preparatory and precisely sequenced neuronal activity in production of a skilled behavior. *eLife* 2019, 8. This study describes the use of intersectional methods to image populations of song premotor neurons during bouts of singing. In addition to confirming earlier observations of sequential activity in these premotor neurons during song production, these imaging experiments reveal that other subsets of premotor neurons can increase their activity hundreds of milliseconds to seconds before song onset, suggestive of preparatory activity that may help to synchronize respiration before vocalization.
44. Fujimoto H, Hasegawa T, Watanabe D: Neural coding of syntactic structure in learned vocalizations in the songbird. *J Neurosci* 2011, 31:10023–10033. [PubMed: 21734294]
45. Hamaguchi K, Tschida KA, Yoon I et al.: Auditory synapses to song premotor neurons are gated off during vocalization in zebra finches. *eLife* 2014, 3:e01833. [PubMed: 24550254]
46. Prather JF, Peters S, Nowicki S, Mooney R: Precise auditoryvocal mirroring in neurons for learned vocal communication. *Nature* 2008, 451:305–310. [PubMed: 18202651]
47. Sturdy CB, Wild JM, Mooney R: Respiratory and telencephalic modulation of vocal motor neurons in the zebra finch. *J Neurosci* 2003, 23:1072–1086. [PubMed: 12574438]
48. Williams H, Nottebohm F: Auditory responses in avian vocal motor neurons: a motor theory for song perception in birds. *Science* 1985, 229:279–282. [PubMed: 4012321]
- 49. Bush A, Doppler JF, Goller F, Mindlin GB: Syringeal EMGs and synthetic stimuli reveal a switch-like activation of the songbird's vocal motor program. *Proc Natl Acad Sci U S A* 2018, 115:8436–8441. [PubMed: 30068604] Studies in sleeping and singing birds show that auditory-evoked activity in vocal muscles during sleep closely resemble vocal muscle activity during singing. Moreover, synthetic song stimuli, in which acoustic cues are systematically degraded show that the vocal muscle responds in an all or none fashion even to degraded songs, behavior that is reminiscent of an attractor network. As auditory-evoked activity in the vocal muscle depends on the song premotor nucleus HVC, HVC emerges as a likely site where auditory stimuli are converted into vocal action.
50. Amit DJ: *Modeling Brain Function: the World of Attractor Neural Networks*. Cambridge England; New York: Cambridge University Press; 1989.
51. Voigt C, Leitner S, Gahr M: Repertoire and structure of duet and solo songs in cooperatively breeding white-browed sparrow weavers. *Behaviour* 2006, 143:159–182.
- 52. Hoffmann S, Trost L, Voigt C et al.: Duets recorded in the wild reveal that interindividually coordinated motor control enables cooperative behavior. *Nat Commun* 2019, 10:2577. [PubMed: 31189912] Neurobiological field work of the most challenging sort, in which simultaneous

acoustic and neurophysiological recordings were made from male and female duetting birds, reveals rapid and precise synchronization of song premotor activity in response to a partner's vocalizations, suggesting that auditory stimulation can entrain vocal motor activity.

53. Hammerschmidt K, Whelan G, Eichele G, Fischer J: Mice lacking the cerebral cortex develop normal song: insights into the foundations of vocal learning. *Sci Rep* 2015, 5:8808. [PubMed: 25744204]
54. Bennett P, Maier E, Brecht M: Involvement of rat posterior prelimbic and cingulate area 2 in vocalization control. *Eur J Neurosci* 2019, 00:1–17.
- 55. Okobi DE Jr, Banerjee A, Matheson AMM et al.: Motor cortical control of vocal interaction in neotropical singing mice. *Science* 2019, 363:983–988. [PubMed: 30819963] Behavioral and electromyographic recordings of antiphonal singing in a Central American mouse, along with functional mapping and manipulation of cortical activity establishes that a part of the mouse's motor cortex is necessary for antiphonal singing between two mice, even though it does not play a role in patterning or producing spontaneous singing when the mouse is isolated. This finding emphasizes that a fundamental role of the cortex may be to help trigger vocalization as a function of social and auditory context, a capacity that may have helped give rise to vocal imitation.
56. Benichov JI, Benezra SE, Vallentin D et al.: The forebrain song system mediates predictive call timing in female and male zebra finches. *Curr Biol* 2016, 26:309–318. [PubMed: 26774786]
57. Roberts TF, Gobes SM, Murugan M et al.: Motor circuits are required to encode a sensory model for imitative learning. *Nat Neurosci* 2012, 15:1454–1459. [PubMed: 22983208]
58. Roberts TF, Tschida KA, Klein ME, Mooney R: Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning. *Nature* 2010, 463:948–952. [PubMed: 20164928]
- 59. Zhao W, Garcia-Oscos F, Dinh D, Roberts T: Inception of memories that guide vocal learning in the songbird. *Science* 2019, 366:83–89. [PubMed: 31604306] Optogenetic stimulation of auditory afferents to the song premotor nucleus HVC in tutor-naïve juvenile zebra finches can drive them to form songs with temporal patterns resembling the stimulus template. These findings emphasize the role of HVC in encoding tutor song memories while also highlighting that other structures presumably play a role in helping to encode spectral features of the tutor song.
60. Chen Y, Matheson LE, Sakata JT: Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proc Natl Acad Sci US A* 2016, 113:6641–6646.
61. Deregnaucourt S, Poirier C, Kant AV et al.: Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. *J Physiol Paris* 2013, 107:210–218. [PubMed: 22982543]
- 62. Tanaka M, Sun F, Li Y, Mooney R: A mesocortical dopamine circuit enables the cultural transmission of vocal behaviour. *Nature* 2018, 563:117–120. [PubMed: 30333629] This study illuminates how auditory and social cues are integrated in the song premotor nucleus HVC to help encode auditory memories of a social tutor and promote copying of the tutor's song behavior. The involvement of the midbrain A11 cell group, which provides dopaminergic input to HVC, in signaling the presence of a singing tutor may point to a conserved role for this cell group in social imitation, as this midbrain population is highly conserved in both birds and mammals.
63. Doya K, Sejnowski T: A computational model of avian song learning In *The New Cognitive Neurosciences*. Edited by Cambridge Gazzaniga M.: MIT Press; 1999:469–482.
64. Tumer EC, Brainard MS: Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. *Nature* 2007, 450:1240–1244. [PubMed: 18097411]
65. Gadagkar V, Puzerey PA, Chen R et al.: Dopamine neurons encode performance error in singing birds. *Science* 2016, 354:1278–1282. [PubMed: 27940871]
66. Hisey E, Kearney MG, Mooney R: A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. *Nat Neurosci* 2018, 21:589–597. [PubMed: 29483664]
67. Xiao L, Chattree G, Oscos FG et al.: A basal ganglia circuit sufficient to guide birdsong learning. *Neuron* 2018, 98:208–221 e205. [PubMed: 29551492]

68. Hoffmann LA, Saravanan V, Wood AN et al.: Dopaminergic contributions to vocal learning. *J Neurosci* 2016, 36:2176–2189. [PubMed: 26888928]
69. Gale SD, Person AL, Perkel DJ: A novel basal ganglia pathway forms a loop linking a vocal learning circuit with its dopaminergic input. *J Comp Neurol* 2008, 508:824–839. [PubMed: 18398824]
70. Person AL, Gale SD, Farries MA, Perkel DJ: Organization of the songbird basal ganglia, including area X. *J Comp Neurol* 2008, 508:840–866. [PubMed: 18398825]
71. Vates GE, Broome BM, Mello CV, Nottebohm F: Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches. *J Comp Neurol* 1996, 366:613–642. [PubMed: 8833113]
72. Mandelblat-Cerf Y, Las L, Denisenko N, Fee MS: A role for descending auditory cortical projections in songbird vocal learning. *eLife* 2014, 3.
- 73. Kearney MG, Warren TL, Hisey E et al: Discrete evaluative and premotor circuits enable vocal learning in songbirds. *Neuron* 2019, 104:559–575. [PubMed: 31447169] The use of dual closed-loop optogenetic and behavioral manipulations in singing birds establish that two inputs to the VTA both function to evaluate song performance, but exert effects on song learning of opposing valence. Moreover, a similar approach dissociates these evaluative circuits from downstream premotor circuits that also play a role in song learning. This and the following study sharpen the focus for circuits that compare vocalization-related auditory feedback with tutor song memories to regions upstream of the VTA.
- 74. Chen R, Puzerey PA, Roeser AC et al.: Songbird ventral pallidum sends diverse performance error signals to dopaminergic midbrain. *Neuron* 2019, 103:266–276 e264. [PubMed: 31153647] Painstaking recordings of ventral pallidal neurons that project to the VTA establish that they some are excited and others are suppressed by distorted singing-related feedback, indicating the VP can signal both good and bad aspects of performance, and that some VP_{VTA} neurons also encode positive reward prediction error. Thus, the VP could help to provide information to the VTA that is used in computing both positive and negative reinforcement signals for vocal learning.
75. Nicholson DA, Roberts TF, Sober SJ: Thalamostriatal and cerebellothalamic pathways in a songbird, the *Bengalese finch*. *J Comp Neurol* 2018, 526:1550–1570. [PubMed: 29520771]
76. Pidoux L, Le Blanc P, Levenes C, Leblois A: A subcortical circuit linking the cerebellum to the basal ganglia engaged in vocal learning. *eLife* 2018, 7.

Summary and future directions

Over the past two years, many important advances have been made in understanding mechanisms of vocal gating, cortical contributions to innate and learned vocalizations, neural mechanisms underlying the formation of song memories, and mechanisms that evaluate and reinforce vocal performance. Going forward, major goals include:

1. In both rodents and songbirds, exploring how the PAG gates vocal and respiratory pattern generating networks and how it also integrates disparate sources of information to help determine when and what type of vocalizations should be produced.
2. In rodents, exploring how frontal cortical regions enable vocalizations to be regulated flexibly and adaptively as a function of social context and in response to auditory stimulation, a capacity that may have ultimately enabled the evolution of vocal learning.
3. In songbirds, exploring how circuits upstream of the VTA function compare auditory feedback with song memories to evaluate vocal performance.
4. In songbirds, to more broadly and agnostically map circuits for song learning; the recent evidence [75,76] of cerebellar connections to song nuclei that contribute song learning underscores that such a broader perspective will be necessary.

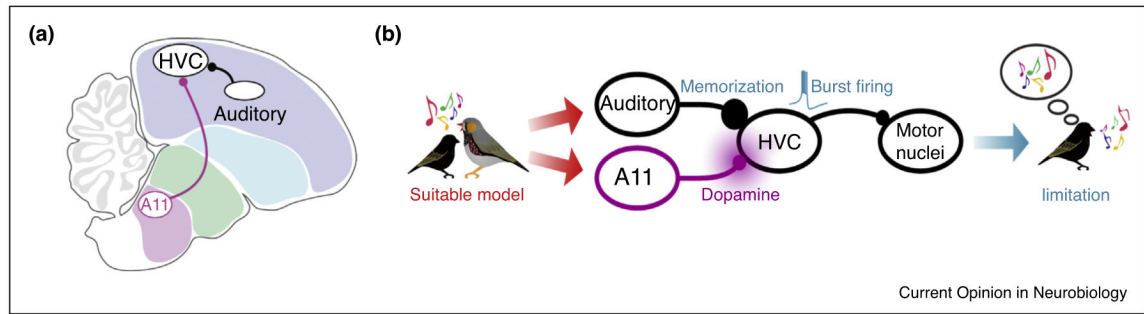


Figure 1.

Integration of social and auditory cues in the juvenile songbird's brain. **(a)** A sagittal view of the songbird brain emphasizing the convergence in HVC of dopaminergic inputs (from the midbrain cell group A11) and auditory inputs from the sensorimotor nucleus NIf. **(b)** The juvenile's encounter with a signing tutor triggers coincident activity of auditory and dopamine inputs to HVC, promoting song memorization, burst firing in the motor network, and song imitation. Reproduced from Tanaka *et al.* [62**].

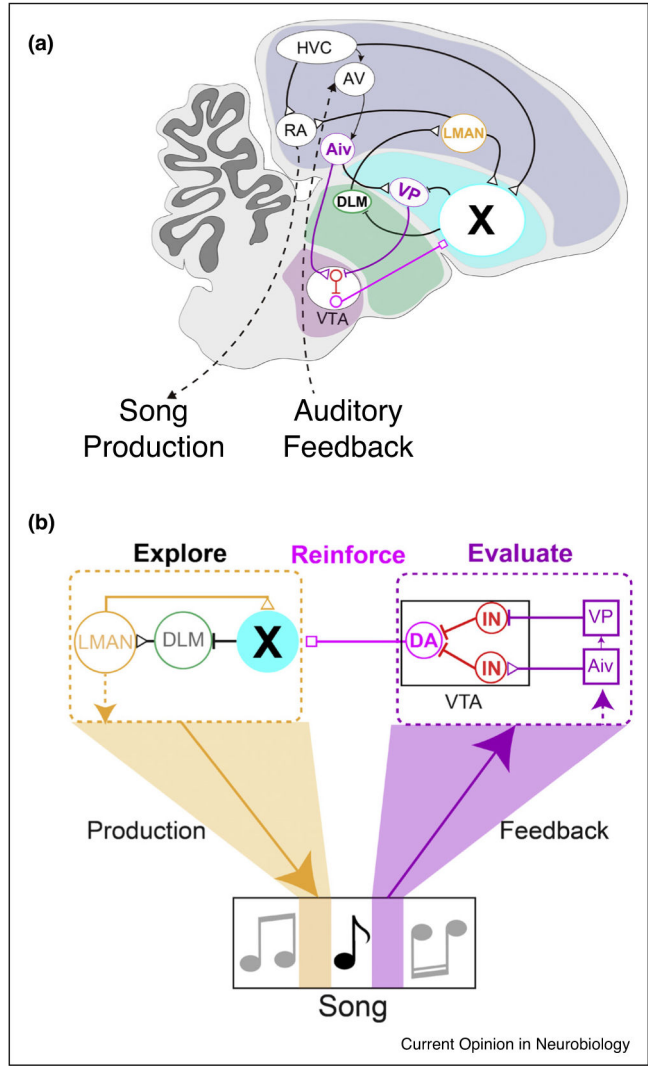


Figure 2. Neural circuits for song production, vocal exploration, reinforcement, and auditory feedback-dependent performance evaluation in the songbird. **(a)** A sagittal view of the songbird brain showing some of the structures important to song learning. These include a song production pathway (HVC (used as a proper name), RA (robust nucleus of the arcopallium), and RA’s descending inputs to the vocal-respiratory brainstem (shown as a dashed line)); a cortico-basal ganglia loop that enables vocal exploration important to song motor learning (Area X (a song specialized region of the basal ganglia), DLM (dorsolateral thalamic nucleus) and LMAN (lateral magnocellular nucleus of the anterior nidopallium)); reinforcing dopamine signals from the VTA; and circuits interposed between HVC and VTA that evaluate song performance (AV (avalanche), Aiv (ventral part of intermediate arcopallium), VP (ventral pallidum)). Major brain divisions are colored: pallium = gray; basal ganglia = cyan; thalamus = green; midbrain = purple. **(b)** A block diagram summarizing the organization of the portions of the song circuit important to song motor

exploration, reinforcement, and performance evaluation. IN = interneurons. DA = dopamine. Modified from Kearney *et al.* [73**].

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript