

Review



Cite this article: Nieder A, Mooney R. 2019

The neurobiology of innate, volitional and learned vocalizations in mammals and birds. *Phil. Trans. R. Soc. B* **375**: 20190054. <http://dx.doi.org/10.1098/rstb.2019.0054>

Accepted: 15 July 2019

One contribution of 15 to a theme issue ‘What can animal communication teach us about human language?’

Subject Areas:

behaviour, cognition, neuroscience

Keywords:

vocalization, vocal pathways, songbird, mouse, monkey

Authors for correspondence:

Andreas Nieder

e-mail: andreas.nieder@uni-tuebingen.de

Richard Mooney

e-mail: mooney@neuro.duke.edu

The neurobiology of innate, volitional and learned vocalizations in mammals and birds

Andreas Nieder¹ and Richard Mooney²

¹Animal Physiology Unit, Institute of Neurobiology, University Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

²Department of Neurobiology, Duke University School of Medicine, Durham, NC 27710, USA

AN, 0000-0001-6381-0375

Vocalization is an ancient vertebrate trait essential to many forms of communication, ranging from courtship calls to free verse. Vocalizations may be entirely innate and evoked by sexual cues or emotional state, as with many types of calls made in primates, rodents and birds; volitional, as with innate calls that, following extensive training, can be evoked by arbitrary sensory cues in non-human primates and corvid songbirds; or learned, acoustically flexible and complex, as with human speech and the courtship songs of oscine songbirds. This review compares and contrasts the neural mechanisms underlying innate, volitional and learned vocalizations, with an emphasis on functional studies in primates, rodents and songbirds. This comparison reveals both highly conserved and convergent mechanisms of vocal production in these different groups, despite their often vast phylogenetic separation. This similarity of central mechanisms for different forms of vocal production presents experimentalists with useful avenues for gaining detailed mechanistic insight into how vocalizations are employed for social and sexual signalling, and how they can be modified through experience to yield new vocal repertoires customized to the individual's social group.

This article is part of the theme issue ‘What can animal communication teach us about human language?’

1. Introduction

Vocalizations play a fundamental role in intraspecific communication in birds and mammals. As in any communication system, vocal signals need to be produced by a sender and deciphered by a receiver. Thus, both the production and the perception of vocalizations are integral parts of vocal communication. This article focuses on the neural mechanisms of vocal production from a comparative point of view in songbirds, in primates and laboratory mice, the latter species providing the most genetically tractable vertebrate organism in which to map and manipulate central circuits for vocal control.

In discussing the behavioural and neural foundations of vocal production, two issues are of special importance. The first issue concerns whether vocalizations are innately programmed, or instead acquired through learning. The second issue is whether vocalizations are elicited solely through emotional (affective) mechanisms or can also be volitionally controlled. These two aspects constitute orthogonal axes in a ‘vocal feature space’ because any combinations between degrees of learning and control are possible, from innate-affective (e.g. alarm calls), innate-volitional (instructed calls), learned-affective (birdsong), to learned-volitional (human speech).

(a) Innate versus learned vocalizations

The vast majority of avian and mammalian species produce only innate vocalizations, defined here as vocalizations that are closely controlled by hard-wired brain structures and their underlying genetic programmes with little or

no environmental influence. These innate vocalizations occur spontaneously in all healthy members of a species whenever they are exposed to a certain stimulus, and are not learned through cultural experience or through vocal practice and performance evaluation. They help to communicate about food, mediate social interactions or signal the presence of different predators. For instance, chickadees produce a high-frequency, low-amplitude ‘seet’ alarm call in response to flying raptors, but a loud, broadband ‘chick-a-dee’ alarm call in response to a perched or stationary predator [1]. Similarly, vervet monkeys are well known to elicit three types of predator-specific alarm calls: ‘leopard alarm calls’ are short tonal calls produced in a series of inhalations and exhalations, ‘eagle alarm calls’ are low-pitched grunts, while ‘python alarm calls’ are high-pitched ‘chutters’ [2] (but see [3] for evidence that both cognitive appraisal of the situation and internal state contribute to the variation in call usage and structure).

Despite the complex ways that monkeys can use different vocalizations, Kaspar-Hauser experiments in squirrel monkeys [4], hybridization studies in gibbons [5] and cross-fostering experiments in macaques [6] underscore that the underlying vocal patterns produced by non-human primates are innate. Cross-fostering experiments and early deafening experiments in mice also indicate that their vocal repertoires are innate ([7–9], but see [10]). Humans also produce a wide range of vocalizations, such as cries or laughter, even when they are congenitally deaf, indicating that these vocalizations are innate rather than learned [11].

While vocal utterances in non-human primates are innate, recent behavioural experiments in the marmoset, a new-world monkey, demonstrate the influence of social experience on vocal maturation. Marmoset infants initially produce a high ratio of ‘cries’ to ‘phee’ calls in the first month of post-natal life, then transition to producing a repertoire comprising predominantly phee calls. This transition occurs earlier in development in infants whose phee calls more frequently elicited phee calls from their parents [12–14]. Moreover, this change cannot be attributed to effects of social experience on overall growth rates; rather, an infant marmoset appears to select and refine elements from its innate vocal repertoire via social–auditory reinforcement from its parents. In support of this view, a study of twin marmoset infants found that the twin that received greater contingent ‘phee’-back played from a speaker transitioned to phee calls earlier in life [15]. Caveats include that infant marmosets with less responsive parents still transition to adult proportions of phee calls and that the effects of social–vocal experience on factors other than gross body weight that could influence vocal development, such as steroid levels in the juvenile’s brain and larynx, have yet to be explored. Nonetheless, the social reinforcement-based change in marmoset vocalizations may be regarded as vocal production learning in a very broad sense [16–18]. In that respect, it resembles aspects of early (prelinguistic) vocal development in humans [19,20], songbirds [21,22] and bats [23,24]. This simpler form of vocal plasticity suggests that marmosets may serve as valuable models for rudimentary aspects of speech development and evolution, in addition to their established utility for studying how vocalizations are used for social communication and recognition (see table in [25] for a list of studies of plastic primate vocal behaviours). However, such a broad definition of vocal production learning needs to be clearly differentiated from vocal learning via imitation.

As more narrowly (and accurately) defined, vocal production learning involves imitation of the vocal patterns produced by another individual, including that of another species or even artificial (man-made) sounds. Vocal imitation requires that the ‘pupil’ hear an appropriate model and engage in extensive vocal practice and evaluation through auditory feedback, and involves enhanced forebrain control of the vocal organ [26]. In fact, surprisingly few groups of vertebrates learn their species-typical vocalizations: other than for speech learning in humans, evidence of vocal learning exists only in a few different orders of birds (oscine songbirds [27–29], parrots [30–32] and certain hummingbirds [33]) and of mammals (pinniped carnivores [34], cetaceans [35], bats [36] and elephants [37]). Even then, systematic experimental evidence of the necessity of models, practice and auditory feedback-dependent performance evaluation is lacking in pinnipeds, whales, bats and elephants. In conclusion, despite the fact that the vocalizations of non-human primates change over development, the limited vocal modifications seen in nonhuman primates, including chimpanzees, are categorically different from vocal production learning in this more narrow and traditional sense [38].

Although the capacity for vocal learning emerged independently in songbirds and humans, birdsong and speech learning share many striking similarities. Most notably, songbirds and humans learn their species-typical vocalizations during a juvenile sensitive period in a process that depends on auditory experience of an appropriate vocal model, extensive vocal practice, and self-evaluation and adaptive modification of vocal performance through auditory feedback [39–43]. In addition, both humans and songbirds have evolved a complex hierarchy of specialized forebrain sensorimotor structures important to the acquisition and production of learned vocalizations, and that exert their influence on vocalizations through the same brainstem regions important to the production of innate vocalizations [44–47]. Indeed, innate and learned vocalizations depend on the activity of common pools of motor neurons and a single vocal end organ, as will be discussed later in this review in §2. Thus, an understanding of the neural mechanisms underlying innate and learned vocalizations in songbirds and humans can inform how innate and learned motor patterns coexist and cooperate with each other to generate a full vocal repertoire.

(b) Internal state-based versus volitional control of vocalizations

Many and perhaps most vocalizations uttered by birds and mammals are a consequence of specific internal states. The alarm calls made by chickadees and vervet monkeys just described are illustrative examples of such internal state-based vocalizations. Here, the internal state change in the presence of a predator is sufficient to explain the vocal utterances. The internal state is used as an umbrella term that contains different intrinsic factors. According to one framework, the internal state consists of an affective component (related to the individual’s evaluation of the environment), a motivational component (related to the individual’s action tendencies) and an arousal component (related to the individual’s likelihood and urgency to respond) [48]. Thus, affect, motivation and arousal all describe different, sometimes orthogonal internal states that might elicit vocalizations.

Without a direct measure of physiological parameters in controlled vocalization contexts, it is virtually impossible to figure out how these various internal states relate to the production of vocalizations [49]. In one rare attempt in which a direct link between specific call types and internal state could be established in squirrel monkeys, the monkeys were trained to increase or avoid electric stimulation of specific brain areas by switching between different compartments in a cage [50]. Aversive situations in which the monkeys avoided stimulation were predictably associated with specific call types, whereas appetitive situations reliably elicited other innate types of vocalizations [50]. The results indicate that calls are loosely tied to different affective or motivational states.

A recent study in marmosets used noninvasive surface electromyography to measure heart rates as a proxy for arousal while the monkeys produced different vocalizations that were elicited by systematically manipulating social context [51]. It was found that arousal levels were correlated with changes in different acoustic features and the production of different vocalizations. These findings are in agreement with the notion of arousal level changes resulting in call type differences. However, this study also found that the production of these different call types is also affected by extrinsic factors such as the timing of a conspecific's vocalization in contexts where marmosets are interacting.

While internal state is a natural driver of many human vocalizations, we can also vocalize under volitional control, such as when we engage in spoken conversation. In fact, a distinction between volitional and affective (spontaneous, emotional) movements has long been recognized in clinical neurology, especially for orofacial movements that are indispensable to vocalization [52,53]. Patients with facial paralysis owing to damage to the descending pathways from the motor cortex have considerable difficulty smiling or frowning in response to a neurologist's commands (a condition called 'voluntary facial paresis'), even though they smile or frown naturally in response to their emotional state. Similar dissociations have been observed for vocalizations, where some patients with neurological insults may lose volitional control of their speech (despite intact speech comprehension), but can still laugh, scream or groan when they are happy, frightened or in pain. For example, following damage to the inferior frontal gyrus, non-verbal vocal utterances remain intact despite devastating impairments in speech and language production (Broca's aphasia) [54–57]. Moreover, patients with a clinical diagnosis of primary progressive aphasia develop abnormal laughter-like vocalizations that increasingly replace speech in the context of progressive speech output impairment leading to mutism, until ultimately laughter-like vocalizations are the only extended utterance produced by these patients [58]. Neuropsychological studies in these patients further help to show that affective and volitional vocalizations are mediated by distinct neural pathways, as will be discussed in further detail in §5.

Whether non-human animals share the capacity for volitional control of vocalization is a matter of intense debate. Despite claims of volitional vocal control in some animals, particularly in nonhuman primates [59], these claims remain controversial. Without specific experimental efforts to control the affective status of an animal under investigation, for instance, affective causes will always remain the simplest explanation for vocal initiation, particularly when working with spontaneously behaving animals.

A key issue is how to assess whether non-human animals can exert volitional control of their vocalizations. Clinical neurology provides three criteria useful for making such an assessment. First, vocalizations need to be uttered in response to an arbitrary instruction stimulus that is neutral in its value or emotional valence. Second, vocalizations need to be uttered in manner that is temporally contingent to the instruction stimulus. Third, vocalizations need to be produced reliably after the presentation of the instructive stimulus, and withheld in its absence. Note that these criteria for volitional vocalization are not identical to current definitions of 'intentional communication' in animals, which primarily rely on the effects that communicative signals may exert on a potential recipient [59].

Indeed, recent studies determined that monkeys and crows (corvid songbirds) can be trained in a manner that satisfies these three criteria, and thus display volitional vocal control [60,61]. There, the animals were trained in a computer-controlled Go/Nogo-protocol to vocalize as soon as they saw a visual stimulus, which was delivered with an unpredictable onset time. The red or blue squares used as instructive stimuli were arbitrary and contained no inherent hedonic value that might trigger vocalization through an affective process, satisfying the first criterion. Second, the animals vocalized promptly after the onset of the instruction stimulus (approx. 1.5 s). Third, the animals responded reliably to the instruction stimulus with high hit rates after presentation of the instruction stimulus, and displayed low false alarm rates in the absence of the instruction stimulus [60,61]. Finally, one monkey even learned to switch between two distinct call types from trial to trial in response to different visual cues [60].

These findings indicate that monkeys and songbirds are able to volitionally initiate vocal production and thus can instrumentalize their vocal behaviour in a goal-directed manner. Interestingly, however, this volitional control was developmentally restricted in monkeys: while both monkeys reliably vocalized on command during juvenile periods, they discontinued this controlled vocal behaviour in adulthood [62]. This developmental loss of volitional vocal production contrasts with persistent affective vocal production, as both monkeys continued to vocalize spontaneously as adults. Furthermore, both monkeys could continue to make volitional manual gestures as adults. We speculate that the loss of voluntary vocalizations reflects maturational changes in the monkeys' brains that are unrelated to vocal practice or production *per se*. For instance, hormonal changes associated with sexual maturation contribute to adolescent-typical behavioural changes that necessarily have an impact on large-scale networks, so that functions beneficial during childhood may become inhibited during adulthood. Moreover, synaptic elimination during adolescence likely involves adjustment of the excitatory/inhibitory balance on individual neurons and within networks, given that excitatory synapses are selectively pruned whereas inhibitory synapses are spared [63]. Whatever the reasons, volitional vocal control is a recent evolutionary invention in primates, and only fully available throughout life in humans.

2. Peripheral vocal production mechanisms

Vocalization in birds and mammals requires precisely coordinated activity of many different respiratory and vocal

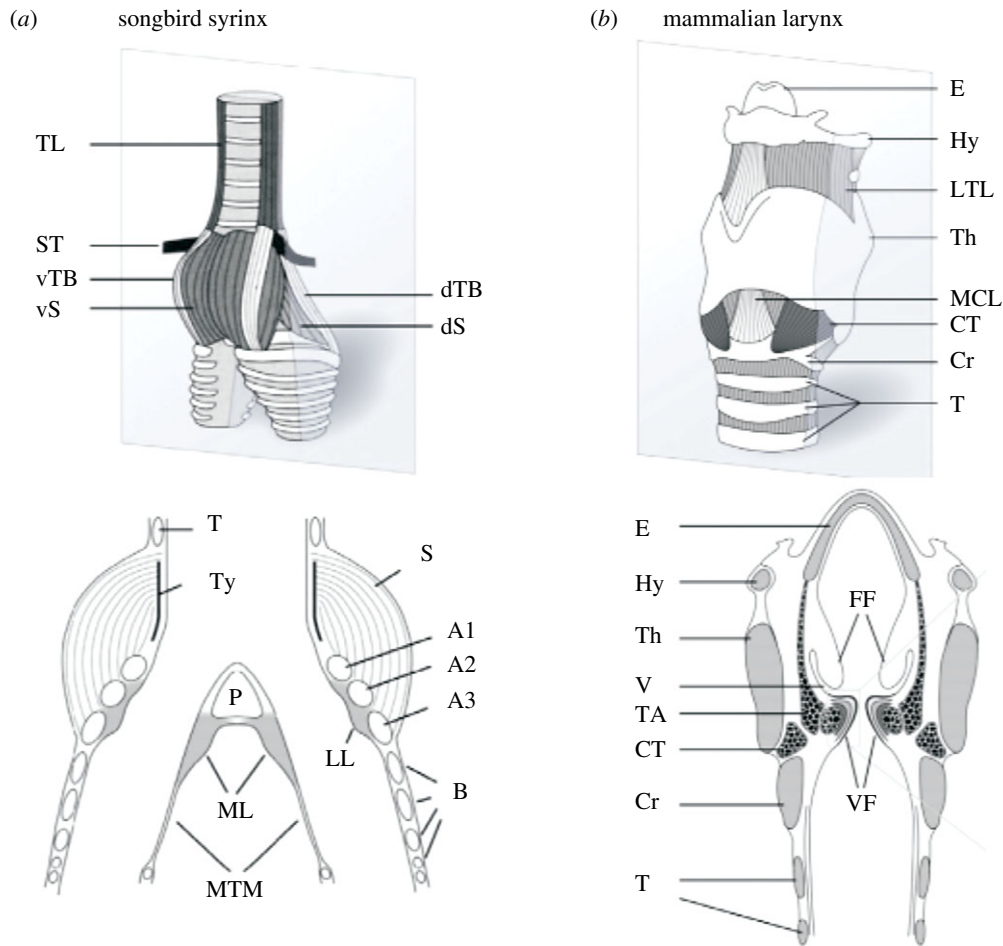


Figure 1. Songbird syrinx (*a*) and mammalian larynx (*b*) shown in en face (top) and frontal cross-section (bottom) views. The syrinx is located at the juncture of the trachea with bronchi. Abbreviations, left two panels: TL, m. tracheolateralis; ST, m. sternotrachealis; vTB, m. tracheobronchialis ventralis; dTB, m. tracheobronchialis dorsalis; dS, m. syringealis dorsalis; vS, m. syringealis ventralis; T, tracheal ring; Ty, tympanum; P, pessulus; ML, medial labium; MTM, medial tympaniform membrane; LL, lateral labium; S, syringeal muscle; A1–A3, tracheo-bronchial semi-rings; B, bronchial rings. Abbreviations, right two panels: E, epiglottis; Hy, hyoid bone; LTL, lateral thyrohyoid ligament; Th, thyroid cartilage; MCL, median cricothyroid ligament; CT, cricothyroid muscle; Cr, cricoid cartilage; T, tracheal rings; FF, ventricular or false folds; V, lateral laryngeal ventricle; TA, thyroarytenoid muscle; VF, vocal folds. Figures reproduced with permission from [67].

muscles, the latter of which includes the muscles of the sound-generating organ (the intrinsic muscles of the larynx in mammals or of the syrinx in birds) as well as those of the upper vocal tract, which can further filter and amplify components of these vocal sounds [64–69]. This integration can be enormously complex: human speech is estimated to involve the coordinated activity of approximately 100 distinct vocal, orofacial and respiratory muscles [47]. Moreover, the intrinsic vocal (laryngeal or syringeal, respectively) muscles are active not only during vocalization but also during quiet breathing [70–72], where they help regulate and gate respiratory airflow [73]. In fact, vocal motor networks are embedded in and likely derived from brainstem respiratory pattern-generating networks [74–77]. Thus, vocal and respiratory pattern generating networks are strongly intertwined and may not be wholly separable.

One important consideration when comparing vocal communication in birds and mammals is that they vocalize through functionally similar but anatomically distinct vocal end-organs [67,73] (figure 1). Birds sing and call through their syrinx, a bipartite muscular organ located at the juncture of the trachea and the two bronchi, whereas mammals vocalize through the larynx, a unipartite structure rostral to this juncture, closer to the cranium. Notably, many songbird species exploit the bipartite architecture of the syrinx to

simultaneously generate two independent and often harmonically unrelated voices [78,79], something that is not possible to achieve with laryngeal vocalizations (although *harmonically* related two-voice effects can be achieved by skilled human vocalists who learn to accent the fundamental and certain of its harmonics through upper vocal tract filtering [80]).

Despite structural differences in their vocal end organs, many functional aspects of the vocal periphery are similar in birds and mammals, as reviewed in depth by [67]. In most instances, vocalizations are emitted during the expiratory phase of the respiratory cycle. Furthermore, to phonate, mammals and birds must simultaneously increase airway pressure and contract the intrinsic laryngeal or syringeal muscles to apply tension to the vocal folds and the labia, respectively. This synergistic respiratory and vocal activity causes periodic vibrations of the vocal folds or labia, which in turn cause periodic vibrations in surrounding air molecules. These airborne vibrations manifest as relatively low frequency (less than 10 kHz) ‘voiced’ harmonic sounds, the fundamental frequency of which (100–200 Hz in adult humans) directly correlates with the activity of certain intrinsic laryngeal or syringeal muscles (i.e. the cricothyroid muscle in mammals and the ventral syringeal muscle in birds) [68,81]. Other similarities include the ability to emit a wide range of vocal frequencies by exploiting different vocal tract dynamics,

spanning from very low frequencies ('vocal fry' or 'pulse tone register') to higher frequencies ('modal register'); the capacity to utter multiple vocal notes, elements, calls or syllables during a single exhalation (a 'fusion' event that figures prominently in human speech); and the accenting and filtering of certain harmonics through the active manipulation of the resonant properties of the upper vocal tract.

Upper vocal tract filtering ('articulation') plays an especially prominent role in speech and birdsong, and thus warrants further attention. When humans or birds produce voiced sounds in a helium-oxygen (heliox) atmosphere, which increases the resonant frequency of the upper vocal tract without affecting the oscillatory frequency of the vocal folds or syringeal labia, previously suppressed harmonics can be detected, providing experimental evidence of this upper vocal tract filter [82,83]. In humans, the articulatory muscles of the face, tongue, jaw and pharynx contribute to the filter; in birds, the muscles that control the oropharyngeal-esophageal cavity (OEC) appear to figure prominently in upper vocal tract filtering [67,84,85]. Songbirds also actively modulate the gape of their beaks during singing, although most likely to optimize acoustic transmission, rather than for spectral filtering [86].

In addition to producing audible vocalizations ('squeaks') through vibrations of their vocal folds, many rodents, including rats and laboratory mice, emit ultrasonic vocalizations (USVs; greater than 30 kHz) [87–89]. Notably, USVs figure prominently in social affiliation and courtship, whereas lower frequency vocalizations are produced more commonly in aggressive encounters, or in response to stress or pain [90]. While rodent 'squeaks' are produced by mechanical vibrations of the vocal folds, USVs are produced by an aerodynamic whistle. Support for this idea comes from the knowledge that the fundamental frequency of an aerodynamic whistle depends on the density of the surrounding gases, whereas the fundamental frequency of a mechanical oscillator (such as the vibrating vocal folds) does not. Consistent with this distinction, the fundamental frequency of rodent USVs shifts upward in a heliox atmosphere, whereas the fundamental frequency of voiced sounds remains unchanged [89]. Moreover, structural analysis of the rodent airway supports the idea that USVs result from an edge-tone whistle effect generated by airflow over a ventral laryngeal pouch [91]. Despite their distinct physical mechanisms, both 'whistled' USVs and voiced vocalizations originate in the larynx and depend on a precise coordination of respiratory and vocal activity. In fact, to produce USVs, rats (and presumably mice) must maintain respiratory pressure below the level that induces vocal fold oscillations, and must precisely adjust the position of the larynx to sustain airflow necessary for generating the ultrasonic whistle [92].

In summary, many striking parallels exist in how vocal sounds are produced in different mammals and birds. Thus, a comparative approach using non-human primates, rodents and birds is likely to be relevant to understanding the neuromuscular control of vocalization more broadly, including for human speech. Even the whistles that give rise to rodent USVs have their analogues in the whistle languages used by humans living in mountainous regions [93]. And while to date the cognitive and semantic roles served by human speech are without obvious parallels in other animal vocalizations, research in songbirds and non-human mammals can provide insights into many neural

processes relevant to human speech, including respiratory–vocal integration in the brainstem, the modulation of these brainstem networks by the forebrain to facilitate vocal learning and volitional control of vocalizations, and the control of vocal sequences necessary to phonological syntax.

3. Brainstem networks for vocalizations

Humans, other primates, mice and songbirds can produce innate vocalizations of normal acoustical structure even in the absence of forebrain inputs to the brainstem. Indeed, while the motivation and ability to gate innate vocalizations as a function of social or environmental context may require an intact forebrain in some species, the pattern of these innate vocalizations is unaffected when the forebrain inputs are removed by surgical transection, focal forebrain lesions or genetic mutations that result in anencephaly [8,55,94,95]. In contrast, speech and birdsong depend intimately on the forebrain: focal lesions in speech- or song-related regions of the human or songbird cortex render individuals unable to speak or sing, respectively, while their ability to utter acoustically normal innate vocalizations remains intact [8,55,94–96]. Here we take a 'bottom up' approach to exploring these brainstem networks.

(a) Vocal and respiratory motor neurons

Fluent vocalization depends on the precise coordination of numerous vocal and respiratory motor neuron pools. Motor neurons for breathing are largely located in the spinal cord and serve similar functions in mammals in birds, except that birds lack a diaphragm and the associated phrenic motor nucleus [97]. Indeed, one important distinction is that avian respiration involves active inspiration and active expiration (rather than passive expiration, as in mammals), mediated in part by a collection of bellows-like air sacs that perfuse a relatively rigid lung [65,97]. Vocal motor neurons in both birds and mammals are located in the caudal medulla, with the intrinsic laryngeal muscles in mammals receiving input from motor neurons in the nucleus ambiguus and syringeal muscles in songbirds receiving their innervation from motor neurons located in the tracheosyringeal part of the hypoglossal motor nucleus (interestingly, the muscles of the tongue, which play a prominent role in human speech (but not in birdsong), are innervated by the lingual part of the hypoglossal nucleus) [97,98]. A notable feature of vocal motor neurons in mammals and birds is that they also are active in normal respiration and, at least in mammals, to effect swallowing [76,99].

Motor neurons important to the muscles of the upper vocal tract originate from multiple sources, including from motor neurons in the nucleus ambiguus, the hypoglossal motor nucleus, the facial motor nucleus and the trigeminal motor nucleus [98,100,101]. Because all of these motor neurons lack axon collaterals and only innervate their respective target muscles, they serve only to read out rather than directly participate in vocal pattern-generating circuits. Consequently, vocal pattern generators must synchronize activity of motor neurons located over a wide extent of the brainstem, rather than relying on reciprocal connections between disparate motor neuron pools to achieve vocal and respiratory coordination.

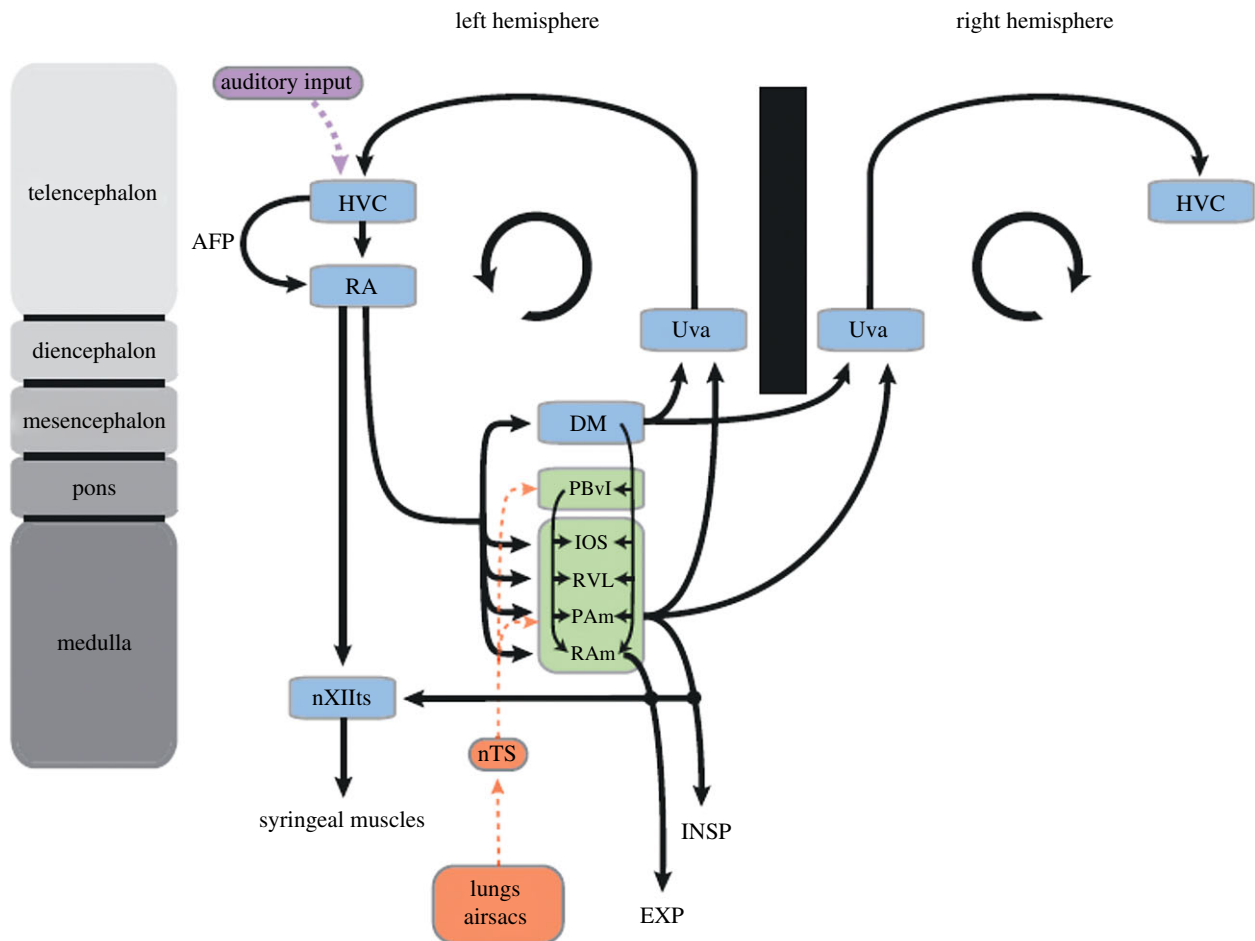


Figure 2. Schematic of the vocal control system of the songbird, emphasizing descending and ascending (recurrent) pathways important to learned song production. Abbreviations: HVC (used here as a proper name); RA, robust nucleus of the acropallium; DM, dorsomedial nucleus of the intercollicularis (the avian equivalent to the PAG); PBvI, ventrolateral part of the parabrachial nucleus; IOS, superior infra-olivary nucleus; RVL, rostroventral lateral medulla; PAm, nucleus parambigualis; RAm, nucleus retroambigualis; Uva, nucleus uvaeformis; nXIIts, tracheosyringeal part of the hypoglossal motor nucleus; nTS, nucleus of the solitary tract; INSP, inspiratory motor neurons; EXP, expiratory motor neurons; AFP, anterior forebrain pathway. The AFP and auditory pathways are highly simplified in this diagram. Descending connectivity is emphasized on the left side, but is not depicted on the right to highlight the role of Uva in providing interhemispheric connections important to bilateral coordination of HVC and central control of song. Figure courtesy of Marc Schmidt. (Online version in colour.)

(b) Brainstem pattern-generating networks for innate vocalizations

An operational framework that enables the experimental identification of vocal pattern-generating circuits includes: (i) the capacity to trigger vocalizations when artificially stimulated with electrical, chemical, chemogenetic or optogenetic methods; (ii) the ability to alter, degrade or entirely suppress spontaneous vocalizations when stimulated, partially lesioned or pharmacologically inactivated; (iii) functionally direct (mono- or paucisynaptic) linkage to vocal and respiratory motor neurons important to phonation; and (iv) component neurons that are active before and during vocalization and that display firing patterns that correspond to acoustic features of the vocalization, such as duration, frequency or call type.

In mammals, brainstem structures that meet these criteria include portions of the lateral reticular formation (LRF) and the nucleus retroambiguus (NRA; or the caudal ventral respiratory group (VRG)) (reviewed in [75,102]). Both the LRF and the NRA contain neurons that project to laryngeal motor neurons, to the motor neurons that innervate the articulatory muscles, and to expiratory motor neurons, thus providing the highly divergent architecture needed to

coordinate vocal and expiratory motor activity during vocalization. Moreover, electrical stimulation in these structures is sufficient to elicit a range of vocalizations that resemble the innate vocalizations of the subject species. Conversely, inactivating the NRA abolishes spontaneous vocalizations as well as vocalizations that would otherwise be elicited by stimulating upstream regions, such as the periaqueductal grey (PAG; see §4). Finally, both the LRF and the NRA contain neurons that are active before and during vocalization and exhibit firing patterns that correspond to specific acoustic features of innate vocalizations. These features advance the LRF and NRA as two crucial components for the patterning of innate vocalizations in mammals.

Although the functional interrogation of the brainstem circuits for innate vocalizations has been less extensive in birds, a likely avian homologue to the mammalian NRA (or caudal VRG) is the nucleus retroambigualis (RAm), which projects to expiratory and syringeal vocal motor neurons (figure 2) (the avian counterpart to the rostral VRG is the nucleus parambigualis, which contains inspiratory premotor neurons and is reciprocally connected to the forebrain song system (discussed in further detail in §5)) [103]. Both the NRA in mammals and RAm in birds project bilaterally onto their respective vocal motor neuron pools [98,103].

This feature may facilitate bilateral coordination of the vocal muscles when upstream regions are highly lateralized in their vocal functions, as with speech cortical regions in humans [96], or are anatomically lateralized, as with descending projections from forebrain song nuclei the brainstem in certain songbirds [104–107]. Finally, because NRA and RAm receive highly convergent input from upstream regions in the brainstem and forebrain that are important to vocal gating, respiratory patterning and the volitional and learned control of vocalization, they likely represent a final common node in the mammalian and avian brainstem for vocal–respiratory patterning and coordination [98,103,108,109].

4. The contributions of the midbrain periaqueductal grey to vocalizations

A crucial question is how the vocal pattern-generating networks in the caudal brainstem are switched on and off to ensure that vocalizations are produced in a manner that is appropriate to the environment, social context and the internal (emotional, developmental or hormonal) state of the animal. In this light, the lateral part of the caudal PAG has emerged as an essential hub for producing innate vocalizations in mammals and birds [109–112]. In primates, the PAG receives inputs from various forebrain structures important to reproductive and social behaviours, including the hypothalamus, amygdala, anterior cingulate cortex, preoptic area and the bed nucleus of the stria terminalis [113] (figure 3). In songbirds, the dorsolateral PAG receives input from forebrain nuclei important to producing learned song and from hypothalamic regions, providing a link between hormonal state of the animal and courtship vocalizations [94,114,115]. Finally, the PAG projects to the NRA and to the rostral VRG (PAm in songbirds), which contain inspiratory premotor neurons and that are in turn reciprocally connected to respiratory pattern-generating circuits [109,113]. Consequently, even though the PAG in the primate provides at best only sparse input onto phonatory motor neurons [113], it is nonetheless well-situated to rapidly influence and coordinate vocal and respiratory activity.

(a) The mammalian periaqueductal grey ‘gates’ vocalization

Perhaps not surprisingly given its anatomical relationship to vocal patterning circuits, the mammalian PAG plays an obligatory role in vocal production. Indeed, bilateral lesions of the PAG induce mutism in a wide range of mammals and, in humans, abolish both speech and innate vocalizations [113,116]. In monkeys, the mutism induced by PAG lesions has been shown to be specific to the generation of vocal responses rather than a consequence of accessory deficits, because the vocal folds show normal respiratory movements (there is no paresis of the vocal folds) [108]. Interestingly, partial lesions of the caudolateral PAG lead to a loss of some vocalization types while others remain intact, suggesting that subsets of PAG neurons regulate specific vocalizations.

Further underscoring the important role of the PAG in vocalization, electrical stimulation of the PAG can elicit non-habituating vocalizations at short latencies (50 ms or less) [117] in apes, monkeys, cats, bats and rodents, with acoustic features that are typical of innate vocalizations

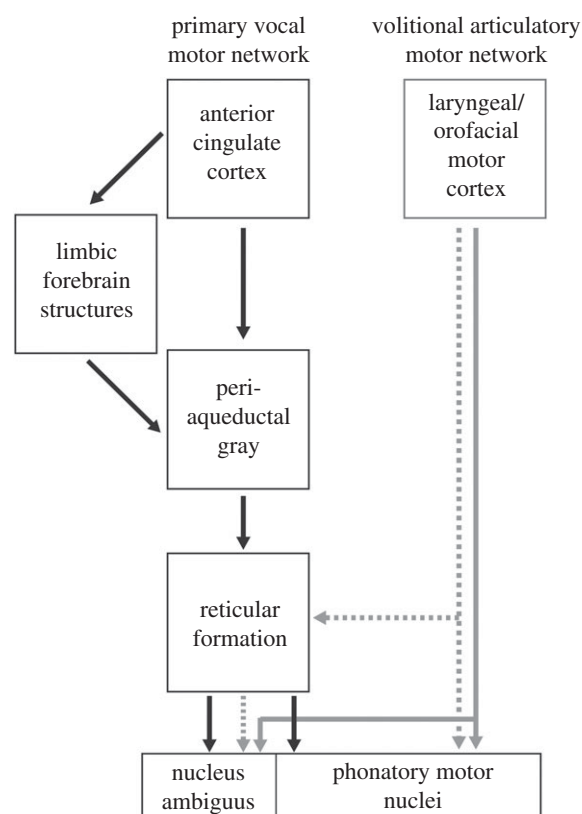


Figure 3. Primate vocal systems. Voice control in humans and nonhuman primates is accomplished by two hierarchically organized vocal systems. The phylogenetically ancient system (‘primary vocal motor network’, black colour) is responsible for innate and affective vocalizations. The phylogenetically new system (‘volitional articulatory motor network’, grey colour) is responsible for volitional vocalizations and learned speech in humans. The projections in nonhuman primates, in particular the indirect projection from the motor cortex to nucleus ambiguus, are indicated by dotted lines; solid lines represent connections in the human brain.

produced by that species (primates [118–121]). Furthermore, stimulation applied at different sites in the PAG can elicit different vocalization types, complementing the selective loss of certain vocal types following partial PAG lesions. And while electrical stimulation in some parts of the PAG can produce appetitive or aversive effects, self-stimulation experiments in monkeys support the idea that vocal effects of PAG stimulation are more direct and not attributable to stimulation-induced affective changes. Finally, the pattern of electrical stimulation in the PAG has little relationship to the pattern of vocalizations that are elicited, supporting a model in which the PAG gates downstream vocal pattern-generating networks, rather than contributing directly to the vocal patterning process itself [122].

(b) A genetic approach to understanding the structure and function of the periaqueductal grey

The functional and structural complexity of the PAG presents a particularly vexing challenge to understanding its role in context-dependent vocal gating. In addition to vocalization, the PAG helps to regulate respiration, nociception, and defensive and sexual behaviours, and vocalization-related PAG neurons are intermingled with the neurons that serve these various other functions [112,123,124]. This multifunctional

and interwoven organization of the PAG has provided something of a roadblock in extending our understanding of the PAG's function in vocalization beyond classical lesion- and stimulation-based approaches. Consequently, whether a distinct and specialized subset of PAG neurons plays a dedicated role in vocal gating has remained unclear. Resolving this issue is important for understanding how vocalizations are integrated into more holistic behaviours, such as courtship or territorial defence, and for more explicitly mapping the circuits for social communication and courtship.

In this regard, an important recent advance is the use of genetic methods to identify and manipulate vocalization-related neurons in the mouse PAG [125]. Tschida and her co-workers labelled PAG neurons that were active in male mice producing USVs (PAG-USV neurons) by using an intersectional strategy employing a knock-in mouse line in which highly active neurons express a specialized receptor that binds an engineered virus [126]. Injecting the virus into a region of the male mouse's PAG that expressed high levels of neuronal activity during vocalization genetically 'tagged' these 'vocal' neurons [125]. A distinct advantage of this approach is that it enables functionally specialized neurons to be genetically labelled and manipulated even when they are anatomically intermingled with neurons that serve other roles.

This approach shows that a subset of neurons in the caudolateral PAG are specialized for producing USVs: genetically silencing these 'PAG-USV' neurons with tetanus toxin prevented males from producing USVs in response to female mice without suppressing other aspects of the male's courtship behaviour; conversely, chemogenetic or optogenetic activation of PAG-USV neurons evoked abundant USVs in socially isolated male mice, which typically utter USVs only when a female is present [125]. Notably, in optogenetic experiments, the pattern of USVs that mice produced were similar with tonic or pulsed (10 Hz) light, further supporting the idea that the PAG gates rather than patterns vocalizations. The specialized and selective role of PAG-USV neurons is further underscored by the finding that optogenetic stimulation of PAG-USV neurons did not evoke the running or 'ballistic' escape movements that are routinely elicited by pan-neuronal stimulation in the PAG. Moreover, these experiments suggest that functionally distinct vocalizations are gated by anatomically distinct subsets of PAG neurons, because genetically silencing PAG-USV neurons did not prevent mice from producing audible 'squeaks' in response to foot shock. Lastly, anterograde tracing reveals that PAG-USV neurons make efferent projections to NRA and the ventral parabrachial complex, providing an experimental avenue for understanding how the PAG gates vocalization [125].

(c) Birds also have a vocal periaqueductal grey

Although the precise homology between the mammalian and avian midbrain PAG still awaits a full accounting [127], a dorsolateral part of the avian PAG just medial and ventral to the inferior colliculus (i.e. the dorsomedial nucleus of the intercollicular complex, or simply DM) serves an important role in producing innate calls [111] (figure 2). Similar to the behavioural effects of stimulating the mammalian PAG, electrical stimulation in DM elicits species-typical calls [72,109].

One potentially important distinction is that songbirds with bilateral DM lesions can still sing their learned songs [94], which contrasts with the loss of speech in humans with analogous brainstem injuries [116]; further experimental confirmation of this important distinction is warranted. Anterograde tracing from the avian DM reveals a pattern of descending projections remarkably similar to those described for the PAG in mammals, with DM axons terminating in RA_m (caudal VRG), the ventrolateral parabrachial nucleus, as well as the nucleus parambigualis (rostral VRG), a region that contains inspiratory premotor neurons and that is reciprocally connected to the respiratory pattern-generating network [109]. One distinction from the mammalian PAG is that DM axons also terminate in the syringeal part of the hypoglossal motor nucleus, providing the 'vocal' midbrain direct access to the phonatory motor neuron pool [109].

5. Forebrain inputs to brainstem vocal networks and their contributions to vocalization

The effects of cortical lesions on human speech can be so devastating that it is easy to forget that significant forebrain (telencephalic and diencephalic) contributions to vocalization are the exception rather than the rule in almost all other vertebrates. Given this exceptionalism, a reasonable assumption is that the ancestral circuits for vocalization reside entirely in the brainstem and the involvement of the forebrain in vocal behaviours has only recently evolved in just a few distantly related species. Nonetheless, numerous forebrain regions make connections with brainstem vocal-respiratory circuits, even in those species that produce only innate, affective vocalizations, presumably providing the scaffolding from which volitional and learned vocal control could be derived. A matter of current debate is whether the forebrain has largely usurped the role of pattern generation in those vocalizations highly dependent on cortical input, such as speech or bird-song. Alternatively, the forebrain might be engaged in a reciprocal interaction with the brainstem, mediated in part by the recurrent connections that brainstem respiratory-vocal regions make with the forebrain, to affect the patterning of these learned vocalizations.

(a) Forebrain vocal circuitry in primates: two parallel, segregated pathways

Two distinct and parallel vocal systems are recognized in primates [47,122,128]. The primordial brainstem vocal system called 'primary vocal motor network' [128] comprises projections from the anterior cingulate cortex (ACC) and other forebrain limbic structures to the PAG, as well as the reticular formation, and participates in stereotypic innate vocalizations driven by affect. The evolutionarily more recent cortical vocal system termed 'volitional articulation motor network' [128] includes projections from the laryngeal motor cortex (LMC) directly to phonatory and respiratory motoneuron pools and is the dominant system for human speech and language (figure 3).

(i) The primary vocal motor network

This system consists of two structurally and functionally distinct parts: the previously described midbrain PAG and

brainstem vocal pattern-generating system, and an upstream limbic vocal initiating network that switches vocal pattern generation on and off based on affective state. This limbic network includes the ACC as well as other telencephalic and diencephalic structures, as detailed in the following paragraphs.

The ACC is typically considered to sit atop the limbic vocal initiating hierarchy. Three cortical areas in the ACC are involved in vocal behaviour and all three of these areas also project to the LMC of the modern volitional articulation motor network [129]. These areas include the rostral (CMAr/M3/area 24c) and caudal (CMAc/M4/areas 24d, 23c) cingulate motor areas (CMA), located in the cingulate sulcus [130], and area 24b rostral to the genu of the corpus callosum [129].

Several lines of evidence demonstrate a role of these ACC regions in vocalization. First, electrical stimulation in the ACC elicits species-specific vocalizations in monkeys as well as other mammals [131–134]. Second, neurons in the macaque CMA modulate their discharge in association with vocalization [135,136]. Finally, bilateral ablation of the ACC vocalization region has mild effects on spontaneous vocalizations, usually characterized by a decrease in the vocalization rate [137–140], although call amplitudes and durations may also decrease [141,142]. Together, these findings suggest that the ACC plays a modulatory role in vocalization.

Several clinical studies implicate the ACC in the initiation of non-verbal vocal utterances and emotional speech in humans. In a type of frontal lobe epilepsy characterized by involuntary and stereotyped bursts of laughter ('gelastic seizures'; [143]), the cingulate gyrus appears to be the most commonly disrupted site [144]. Indeed, electrical stimulation of the rostral ACC can evoke uncontrollable but natural-sounding laughter [143,145,146]. Besides controlling non-verbal vocalizations, the ACC is also important for speech. In humans, bilateral infarction of the ACC near the rostrum of the corpus callosum results in akinetic mutism [147,148], which over a longer period can resolve to speech characterized by monotonous intonation [149]. These clinical findings together with functional imaging in vocalizing humans [150] suggest that the ACC is involved in the emotional intonation of human speech [149].

In addition to the ACC, focal electrical stimulation applied in several other telencephalic and diencephalic structures of the primate can elicit species-specific vocalizations. Such 'evocative' foci are found in the amygdala, the BNST (bed nucleus of the stria terminalis), the substantia innominata (including the basal nucleus of Meynert), nucleus accumbens, septum, preoptic area of the hypothalamus, hypothalamus, midline thalamus and zona incerta in the subthalamus [120,131]. Interestingly, stimulation in different limbic areas produced different call types, while lesions in the amygdala or hypothalamus suppressed distinct types of spontaneously uttered vocalizations [108,151].

The relatively long latency of the vocal responses elicited from these sites (greater than 1 s) and their fast habituation are consistent with stimulation-induced affective changes rather than primary motor responses [120]. In fact, studies of freely moving squirrel monkeys that could either seek or avoid electrical stimulation support the idea that the vocalizations elicited from many of these limbic areas result from stimulation-induced changes in affective state [120]. A notable exception was the ACC, for which elicited vocalization and specific affective states were not correlated.

Compared to the other limbic forebrain areas, the influence of the ACC on vocal behaviour thus seems to be more direct.

(ii) The volitional articulation motor network

The dominant vocal network of humans is a more recently evolved cortical vocal system that is largely distinct from the primordial vocal motor system. This modern cortical vocal system consists of cortical structures that are essential for human speech (figure 3). Although this system is also present in non-human primates, it is anatomically and functionally underdeveloped relative to the human cortical vocal system. The three major cortical components of this system are the laryngeal motor cortex, the supplementary motor area and prefrontal cortical regions that include Broca's area in the human brain.

In humans, the laryngeal motor cortex (LMC), which is located in the face region of the primary motor cortex (M1), provides direct cortical control over volitional vocalizations, most notably speech. This includes the phonatory motor neurons of the cranial nerve nuclei—trigeminal motor nucleus (V), facial nucleus (VII), nucleus ambiguus (IX, X, XI) and hypoglossal nucleus (XII) [47,152,153]. These projections are functionally significant, as the ability to produce speech depends on the LMC [47].

In humans, bilateral damage to the LMC causes complete loss of voluntary control over the speech apparatus, rendering patients unable to speak or sing [55,154]. Although such patients are occasionally able to utter non-verbal vocalizations, such as grunts, wails and laughs, presumably mediated by an intact primary brainstem vocal system, they cannot voluntarily modulate the pitch, intensity or the harmonious quality of their vocalizations [98]. In contrast, bilateral lesions of the LMC in nonhuman primates have no effect on calling behaviour, emphasizing that the pronounced role of the cortical vocal system in human speech is an evolutionarily recent enhancement [139,141,142,155].

The human LMC is located in the posterior part of the primary motor cortex (M1, BA 4) [156] and may consist of two distinct laryngeal representations (figure 4) [157,158]. The first laryngeal representation, which is the likely homologue of the non-human primate LMC [159], is found at the ventral extreme of the orofacial motor cortex, a region where electrical stimulation can elicit or disrupt the production of a range of human vocalizations [160,161]. A more dorsal laryngeal representation in the dorsal part of the orofacial M1 has recently been located that may be specific to humans [157,158]. Electrical stimulation applied in this dorsal laryngeal representation results in an involuntary forced exhalation and prolonged utterance of vowel-like sounds [158].

The different locations of the (ventral) LMC between primate species suggest an anterior-to-posterior shift during primate evolution. In apes, the LMC region is positioned more anteriorly at the border between the primary motor cortex (M1) and the ventral premotor cortex (PMv, BA 6). In macaque monkeys and squirrel monkeys, the LMC sits even more anteriorly in the premotor cortex (PMv, BA 4), between the inferior arcuate sulcus rostrally and the subcentral dimple caudally (figure 4). Unlike the reliable vocal effects of LMC stimulation in humans, electrical stimulation in the LMC of apes and monkeys results only in vocal cord movements, but not in vocalizations [152,162–166].

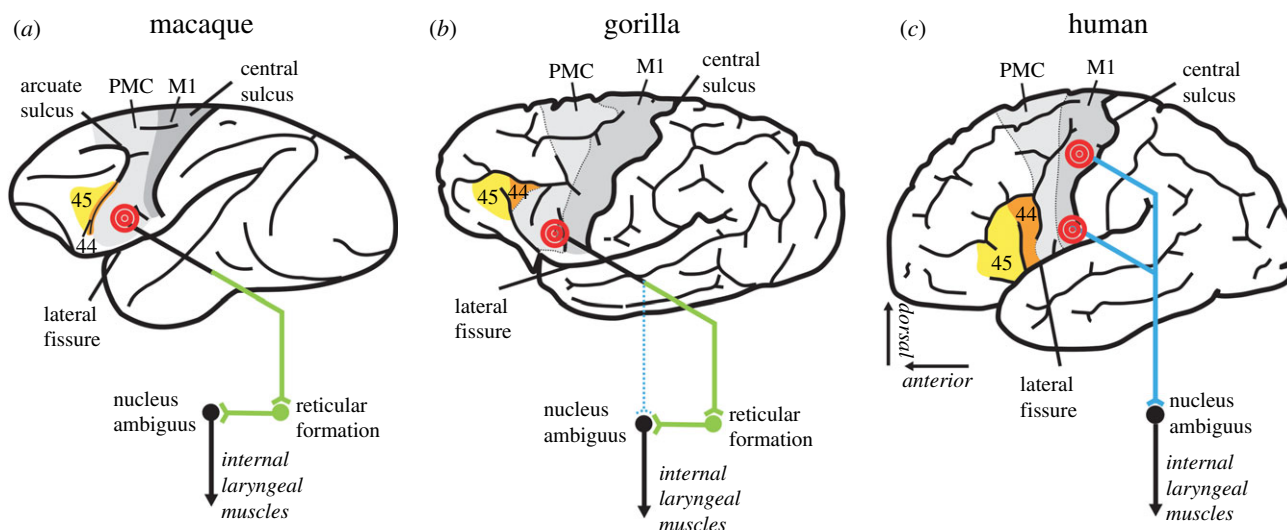


Figure 4. Putative evolution of the modern cortical vocal system ('volitional articulatory motor network') in primates with respect to the laryngeal motor cortex and its projections to phonatory motor nuclei. During the course of primate evolution (left to right), the laryngeal motor cortex (LMC; red concentric circles) shifted from the premotor cortex (PMC; area 6) in monkeys (a) and apes (b) to the primary motor cortex (area 4) in humans (c). This shift was accompanied by a gradual transition from an indirect (via the reticular formation; green colour) to a direct projection of the LMC to nucleus ambiguus (blue colour). In addition, the LMC of humans comprises not one, but two parts in the primary motor cortex. (Online version in colour.)

Important differences in connectivity also distinguish the human LMC from that of other primates. Only LMC neurons in humans make direct projections to nucleus ambiguus (figures 3, 4) [152,167–169]. In contrast, the monkey LMC is connected only indirectly with the nucleus ambiguus via the dorsal and parvocellular nuclei of the reticular formation of the brainstem [152,166,167]. Apes take an intermediate position and show a sparse monosynaptic pathway from the LMC to the nucleus ambiguus [169]. Another noteworthy distinction of the human relative to the macaque LMC is a nearly sevenfold stronger connectivity with somatosensory and inferior parietal cortices [170]. These enhanced LMC–parietal connections in humans likely allowed for more refined sensorimotor integration of information necessary for learned vocal production. Conversely, the monkey LMC has greater connectivity with the ACC, which may be important for the cortical initiation of innate vocalizations [98].

An intriguing idea is that over the course of hominid evolution, the LMC shifted caudally from an 'old' motor cortex present in all primates to a phylogenetically 'new' motor cortex found only in humans [171]. Presumably, the precise, volitional control of laryngeal movements necessary for human speech is the result of a bipartite LMC, the LMC's direct access to laryngeal motoneurons and the LMC's enhanced intracortical connectivity [156,172]. In contrast, the LMC in monkeys may primarily serve evolutionarily conserved non-vocal laryngeal functions, such as breathing, coughing and swallowing [129].

Of course, the LMC does not operate in isolation. Neuro-anatomical tract tracing in the rhesus monkey reveals that the LMC is connected with several cortical motor regions [129]. These cortical regions include surrounding (non-laryngeal) parts of M1, the ventrolateral and caudal ventrolateral premotor cortex (vLPM and vLPFC), which includes Broca's area in humans, the supplementary motor area (SMA) and the ACC.

The SMA in the frontal agranular cortex is an important source of input to the LMC in primates that has acquired a prominent role in human speech production [129]. Since the SMA does not project to the PAG [173], but does form a

direct projection to the LMC, it can be regarded as part of the volitional articulation motor network. Once defined as a single motor area, it is nowadays divided into SMA proper (also called F3) and the rostral part, now termed preSMA (F6) [174]. Only SMA proper is intensely connected with M1, contains cortico-spinal projection neurons [175–177] and readily elicits body movements after electrical stimulation. The preSMA, instead, is heavily interconnected with regions of prefrontal cortex (PFC) [178,179] and therefore not considered a premotor area [180].

Various functional studies support a role for the SMA in human speech. For example, electrical stimulation of the SMA in humans reliably elicits vocalizations [181], whereas lesions in SMA severely reduce the motivation to speak [182,183]. Moreover, functional imaging shows that the SMA is involved in singing, word selection and word production [184–186]. In contrast, the SMA appears to play a much more modest role in vocalizations of other primates, because electrical stimulation of the SMA in monkeys does not elicit vocal utterances [120]. However, bilateral lesions in the SMA decrease the number of spontaneous vocalizations in monkeys, while also increasing the response latencies of instructed vocalizations [139,187]. Moreover, vocalization-related neuronal activity is found in the SMA prior to call onset [136].

A key structure endowing humans with volitional speech control is Broca's area in the PFC. Broca's area classically comprises Brodmann areas 44 and 45 in the inferior frontal gyrus of the granular ventro-lateral PFC (vLPFC) [188]. Broca's pioneering work on the brains of aphasics revealed that areas 44 and 45, usually on the left side of the brain, are instrumental for producing speech and language [54], a finding that has been confirmed since then [189,190].

The primate PFC is regarded as the central executive of the brain [191] and hosts a variety of cognitive functions necessary for the evolution of semantic [192–194] and syntactical [195–197] language functions in humans [128]. In macaques, areas 44 and 45 in the rostral parts of the vLPFC have been identified as an anatomical homologue of Broca's area in the human brain [198,199]. Whereas bilateral ablation

of ventrolateral aspects of the frontal lobe were reported to have no significant impact on discriminatively conditioned vocal behaviour, the variation and lack of complete symmetry in lesion locations complicate interpretation of these negative results [141,142]. In contrast, it has been reported that bilateral removal of prefrontal and orbitofrontal regions (including areas 44 and 45) produced marked vocal deficits, often causing near total and permanent muteness [200]; this mutism may, however, partly be attributed to a disruption of social behaviour rather than a specific vocal deficit. Interestingly, anatomically precise electrical stimulation in area 44 of monkeys elicits orofacial and laryngeal movements, suggesting that this area might have originally enabled volitional control over orofacial actions, including those related to social communication [201], properties that one could readily imagine were exploited over hominid evolution to enable speech and verbal expression.

In macaques trained to vocalize, neurons in vLPFC respond specifically in preparation of volitional calls [202,203], and the pre-vocalization activity of many call-related vLPFC neurons correlates with the acoustic parameters of the ensuing vocalization. Several aspects of this vocalization-related activity indicate that the vLPFC plays a prominent role in volitional call initiation. First, call-related neurons in vLPFC showed the strongest and earliest pre-vocal modulation compared to neurons in the ACC and SMA [136]. Second, neuronal modulation in the vLPFC was absent whenever the monkeys missed a cued vocalization, or when they vocalized spontaneously between trials. Third, and most importantly, vLPFC neurons showed a strong correlation between the onset of neuronal activity and the timing of vocal output: irrespective of the monkeys' call reaction times, vLPFC neurons showed ramping-onset activity approximately 1.2 s prior to the call [136]. These various observations point to a direct involvement of the vLPFC in forming a decision signal for initiating vocalization. A realistic scenario is that the vLPFC, which does not directly project to M1, gains control over the vocal motor network via the premotor cortex (PM), which in turn projects directly to M1 [176,204,205]. In fact, the corticobulbar (for orofacial and laryngeal movements) and corticospinal (for thoracic and diaphragm movements [160]) pathways that originate from both M1 and PM are ideal substrates for volitional vocal control.

In new-world monkeys, the role of the frontal lobe in volitional vocal control has not yet been investigated, but there is emerging evidence for the general involvement of frontal cortical areas in vocal production [206]. Microelectrode recordings of prefrontal (PFC) and premotor (PMC) cortical areas during antiphonal calling have found vocalization-related activity that often precedes or is phase locked to the onset of vocal production [207,208]. In addition, immediate early gene studies in marmosets demonstrated vocal production-related expression particularly in dorsal PFC/PMC [209,210], while other studies have also found expression in ventral PFC [211]. In apes, functional imaging similarly showed activation in the chimpanzee left inferior frontal gyrus, a homologue to Broca's area in humans, during communicative gestural and vocal signalling [212]. Together, these studies suggest an involvement of the lateral frontal cortex in vocal production in non-human primates, with evidence for a role of the vLPFC in volitional vocal control in old-world monkeys (macaques).

Although the vLPFC, and more specifically Broca's area, clearly play a privileged role in human speech, this role may be more modulatory and preparatory in nature. A classical observation in this regard is the finding that electrical stimulation of Broca's area leads to speech arrest rather than speech production [189,213,214]. Moreover, direct cortical surface recordings in neurosurgical patients indicate that Broca's area is predominantly activated before the utterance of a speech element, but is silent during the corresponding articulation [215]. Thus, Broca's area may play a more indirect role in coordinating speech instead of a direct role in speech production [57,216]. In support of this hypothesis, cooling of Broca's area in awake neurosurgical patients slows speech without affecting articulation, whereas focal cooling in the LMC led to slurring [217], consistent with the idea that Broca's area may control aspects of vocal timing without directly regulating speech articulation (figure 5).

One idea is that Broca's area controls articulation by disinhibition of articulatory motor activity briefly before vocal output [219], which would predict that non-verbal, emotional vocalizations might emerge once the modulatory (and/or inhibitory) influence of the voluntary articulation network vanishes [220]. Indeed, after damage to Broca's area, non-verbal vocal utterances remain intact despite devastating impairments in speech and language production [54–57]. Moreover, patients with a clinical diagnosis of primary progressive aphasia develop abnormal laughter-like vocalizations that increasingly replace speech in the context of progressive speech output impairment leading to mutism, until ultimately laughter-like vocalizations are the only extended utterance produced by these patients [58]. Finally, some non-verbal vocal utterances are more common during conversational speech of some aphasic patients [221]. Together, these observations suggest that the volitional articulation motor network and primary vocal motor network may compete to some extent for access to the brainstem vocal–respiratory network.

(b) Forebrain contributions to vocalizations in rodents

(i) Descending inputs to the nucleus ambiguus

In contrast to the primate, the consensus from numerous anatomical studies is that laryngeal motor neurons in rodents receive no direct (monosynaptic) input from the forebrain, and instead receive direct inputs only from medullary, pontine and midbrain regions. For example, injections of horseradish peroxidase into the nucleus ambiguus of the rat retrogradely label afferent neurons extensively in the medulla and pons [222]. These afferents to the nucleus ambiguus include the lateral aspect of the medial parabrachial nucleus, the Kolliker-Fuse nucleus, the reticular formation, the nucleus of the solitary tract, as well as in midbrain structures such as the red nucleus and the superior colliculus, but not any neurons rostral of the midbrain [222]. More recently, transsynaptic viral tracers injected into the laryngeal muscles of rats and mice have helped to map central regions that may be involved in vocalization [10,223–225]. Specifically, pseudorabies virus (PRV, an alpha-herpesvirus) injected into the laryngeal muscles first infects laryngeal motor neurons (i.e. first-order neurons) and then jumps retrogradely into neurons that are directly (second-order) or indirectly (third-order) connected to these motor neurons [223]. Although highly sensitive, PRV and other herpes simplex virus

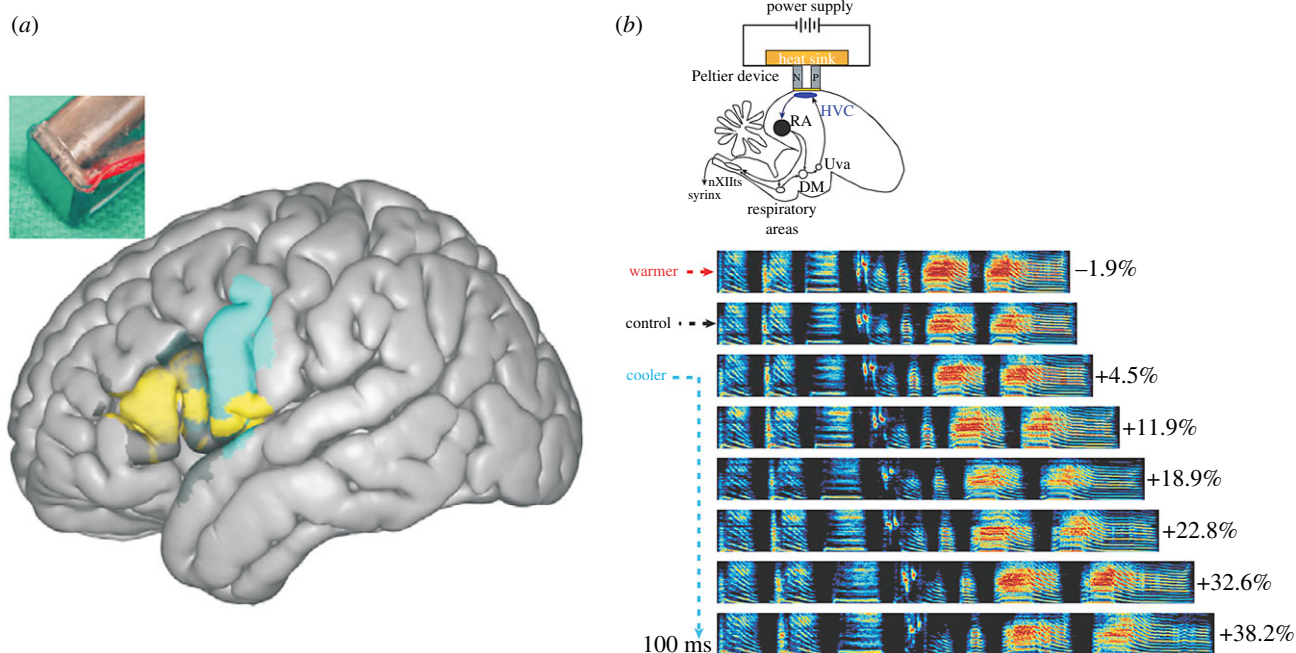


Figure 5. Cooling telencephalic vocal structures slows down vocal output in humans and songbirds. (a) Cortical sites on the left hemisphere that showed effects on speech timing (yellow colour) or deterioration of speech quality (blue colour) when a cooling probe (inset) was placed on Broca's area or primary motor cortex on the precentral gyrus. Cooling of Broca's area primarily affected speech timing, whereas speech quality was mostly affected during cooling of the primary motor cortex. Adapted from [217]. (b) Top panel: a Peltier device was used to cool down neurons in HVC of singing zebra finches. Bottom panel: sonograms recorded from a bird during the cooling experiment show that the song became progressively dilated (indicated by the percentage value to the right) with cooler temperatures relative to the control. Slightly increasing the temperature even mildly shortened the song (top sonogram). Adapted from [218]. (Online version in colour.)

(HSV)-based retrograde synaptic tracers 'jump' synapses rapidly yet variably, and thus this method must be carefully monitored and accompanied by internal controls to establish which labelled neurons synapse directly onto the motor neurons [223,225–227].

Indeed, at the shortest times at which central labelling can be detected after PRV injection into the laryngeal muscles, central labelling is only found in the nucleus ambiguus, a first-order structure, and in second-order structures including the nucleus retroambiguus and the nucleus of solitary tract [223,224]. Notably, third-order structures, such as the midbrain PAG, are only labelled at longer latencies [223], consistent with anterograde tracing studies that fail to reveal a direct projection from the PAG to the laryngeal motor neurons [113]. And although various forebrain neurons, including hypothalamic, amygdalar and cortical neurons are also labelled at longer latencies following PRV injections, these most certainly reflect higher-order and thus relatively indirect projections to the phonatory motor neurons [10,223]. A further caveat is that these retrograde tracing studies only reveal pathways that ultimately access laryngeal motor neurons, but cannot distinguish whether these pathways contribute to vocalization or other behaviours that engage the intrinsic laryngeal muscles, such as respiration or swallowing.

(ii) Forebrain afferents to PAG-USV neurons in mice

As previously discussed, the multifunctional nature of the PAG presents a challenge to gaining further insight into how the forebrain modulates the vocal 'gate' in the PAG. A major recent advance in this regard has been to combine genetic tagging of PAG-USV neurons in the mouse with modified rabies virus tracing methods (K. Tschida 2019, unpublished). This combinatorial approach enables the

identification of neurons that make monosynaptic inputs onto PAG neurons that gate USVs, rather than afferents to the PAG more broadly, effectively focusing the hunt for forebrain neurons that are most likely important to vocal communication. Consistent with prior analysis of vocalization-related forebrain afferents to the PAG, this approach emphasizes the high degree of forebrain convergence onto PAG-USV neurons, including from neurons in the cingulate cortex, motor (M1 and M2) cortex, insular cortex, the BNST, nucleus accumbens, central amygdala, ventral pallidum, the preoptic area and the lateral hypothalamus.

Such highly convergent architecture presumably reflects the extensive and often competing demands to either generate or suppress vocalizations as a function of specific social, reproductive and environmental cues. These results also elevate the mouse as a suitable model for understanding how these various cues are integrated in the forebrain to ultimately engage the vocal machinery of the brainstem in the service of social and sexual signalling, as well as territorial defence. Thus, a profitable future line of enquiry will be to monitor and manipulate the activity of various forebrain neurons afferent to PAG-USV neurons during naturalistic encounters in which USVs figure prominently, including male–female courtship, and same-sex as well as adult–pup interactions.

(iii) The mouse cortex plays little or no role in vocal patterning

Although there are many forebrain inputs to the vocal PAG, their vocal functions are either less obvious or simply not yet known. Notably, mice with bilateral lesions of the motor cortex or that have been genetically engineered to lack a cortex altogether still produce a normal repertoire of USVs [10,228]. And while slight differences in the relative

abundance of certain syllables can be detected in genetically decorticate mice [228], and syllable variability in mice with motor cortical lesions may increase very slightly [10], the cortex of the mouse, as in monkeys and most other mammals outside of humans, simply does not appear to be necessary for vocal patterning. And while the vocal role of the many other forebrain inputs to the PAG-USV neurons has yet to be tested, we speculate here that in rodents, as in most mammals, the brainstem contains all of the neural machinery necessary to produce a normal and complete repertoire of innate vocalizations.

(iv) The mouse cortex may play a role in the contextual control of vocalization

Independent of a role in vocal patterning, the forebrain of the mouse may still play an important role in motivational and emotional aspects of vocalization, for example by regulating vocal output as a function of social or reproductive context. In fact, support for a motor cortical role in the social control of vocalization comes from studies of a wild species of muroid rodent, the short-tailed singing mouse (*Scotinomys teguina*) [229,230]. Male *Scotinomys* produce long strings (tens) of voiced frequency-modulated syllables, typically as part of a ‘call and response’ behaviour with other males. When two males first encounter each other, they take turns at singing, minimizing temporal overlap between their songs. Although male *Scotinomys* can still produce songs with normal patterns following motor cortical lesions, they lose the ability to sing antiphonally in response to hearing another male’s song, suggesting that the role of the motor cortex is to help control vocalization as a function of socially salient auditory cues [230].

Although male *Mus* do not engage in such antiphonal calling with other males, they do produce USVs in the presence of females or female odorants [125]. This ‘sexual’ gating of vocalization implicates reproductive structures, such as the preoptic area (POA) and the lateral hypothalamus, both of which project to the PAG and where electrical stimulation evokes vocalizations in many mammalian species [98]. Moreover, in socially isolated male mice, optogenetically stimulating POA axons in the caudolateral PAG is sufficient to evoke USVs (V. Michael and K. Tschida 2019, unpublished observations), suggesting that POA activity during social encounters provides a contextual signal to the PAG that is sufficient to trigger USV production. In the future, a similar approach can be applied to other forebrain afferents to the PAG to better understand their role in vocalization, whether different afferents recruit different types of vocalizations, or conversely act to suppress vocalization in response to stress or threat.

(c) Forebrain contributions to birdsong: the song system

Other than the essential cortical involvement in human speech, the executive role of the telencephalon in producing learned birdsong has few parallels. Following the revolutionary insights of Nottebohm and his colleagues four decades ago [94,231], a slew of functional and anatomical studies have established that birdsong results from the interaction between a specialized network of forebrain nuclei and the more ancestral vertebrate brainstem vocal-respiratory network described in previous sections (for extensive reviews, see [41,43,45,46,76,94,231–233]). The forebrain components

of this ‘song system’ can be further divided into a song motor pathway, which is obligatory for singing and song learning, and an anterior forebrain pathway that is necessary to song learning but may be largely dispensable for producing songs that a bird has already learned (figure 2). For the purposes of this review, we focus largely on the role of the song motor pathway in adult song production.

Although precise homologies between the avian and mammalian forebrain are still a matter of fascinating debate, the song motor pathway consists of a serially connected chain of cortical nuclei (nucleus interface of the nidopallium (Nif), HVC and RA), the latter of which projects directly onto the syringeal motor neurons, the respiratory premotor complex (nucleus retroambiguus and parambiguus), as well as DM (figure 2) [94,103,234]. This direct projection of RA onto the vocal motor neurons may distinguish songbirds from other bird species that only produce innate vocalizations as well as all non-human mammals, where forebrain projections to the vocal motor neuron pool are typically di- or polysynaptic. In fact, bilateral lesions of HVC and RA abolish learned vocalizations in songbirds, pointing to heavy investment of telencephalic control in birdsong [94,95]. Further, songbirds with HVC lesions can still ‘babble’ and produce other innate calls, paralleling the specialized role of the human cortex in the utterance of speech but not innate vocalizations [94,95]. However, bilateral lesions placed in RA do disrupt antiphonal calling in zebra finches [235], an expressive deficit similar to that reported in short-tailed singing mice, a parallel property that may point to a generalized role for motor cortex in regulating innate vocalizations as a function of social context.

(i) Top-down versus recurrent models of song control

The telencephalic control of birdsong has proven to be an extremely rich system for experimentalists to mine, resulting in wide-ranging studies employing correlative physiology and imaging, as well as causal manipulations such as electrical stimulation, focal cooling and reversible inactivation [95,218,236–244]. Taken together, these various studies underscore that HVC and RA directly contribute to the temporal and spectral patterning of birdsong. Of great current interest is whether song results from the ability of HVC and RA to effectively override or supplant the brainstem vocal and respiratory pattern generators [218,245], or instead arises from recurrent interactions between the forebrain and the brainstem [76,246]. In this latter model, song patterning results not only from the descending projections RA makes to the vocal-respiratory brainstem, but also from the ascending projections that certain neurons in nucleus parambiguus make with Nif and HVC via a thalamic intermediary, the nucleus Uva. As such recurrent architecture is present in most mammalian forebrain-brainstem circuits, resolution of this debate is likely to have ramifications that extend beyond the neural control of birdsong.

(ii) Temporally precise bursts of activity underlie song motor codes

Single-unit extracellular recordings and, in some cases, intracellular recordings, made in the HVC of singing birds reveal that neural activity at these sites can precede song onset by tens to hundreds of milliseconds and is insensitive to auditory feedback perturbations, both of which are features

consistent with premotor activity [236,239,242,247,248]. Moreover, during song, HVC neurons that project to RA (i.e. HVC_{RA} neurons), as well as RA neurons that innervate the brainstem, show clock-like precision and stereotypy of firing patterns that are time-locked to individual notes and syllables [239,241]. In the most closely studied species, the zebra finch, individual HVC_{RA} neurons typically fire only one brief (approx. 10 ms) burst of high frequency (approx. 400 Hz) action potentials during an entire polysyllabic motif, with different HVC_{RA} neurons firing at different times [239,247], possibly tiling the entire motif [245]. While some debate remains as to the uniformity of this neural tiling of song [249], the most recent evidence involving tour-de-force *in vivo* multiphoton imaging in singing finches indicates that the temporal coverage of the motif provided by large populations of HVC_{RA} neurons is both even and complete [250]. One interpretation of such even tiling of the song is that HVC_{RA} neurons are serially connected in a feed-forward excitatory architecture characteristic of a synfire chain [243,245].

In contrast to HVC_{RA} neurons, RA neurons in the zebra finch fire multiple (8–10) action potential bursts per motif [241,245], presumably reflecting a highly convergent input they receive from HVC_{RA} neurons that fire at different times in the motif [245]. Nonetheless, the singing-related activity patterns of RA neurons are also remarkably reproducible across motifs and locked with sub-millisecond precision to acoustic features of individual syllables [251]. A noteworthy feature of this transformation in how song is encoded by HVC_{RA} and RA neurons is its efficiency for song learning [252], as error signals arising from a ‘mistake’ made at one time in the song only affect HVC>RA synapses at that single time point, and do not propagate inadvertently to other times elsewhere in the song where the performance is ‘correct’.

The stereotyped structure and sequence of syllables sung by adult male zebra finches is a boon to perturbation experiments in singing birds. Thus, electrical microstimulation in HVC can interrupt and sometimes entirely reset a bird’s motif [237,240], consistent with a role for HVC in pattern generation. In a pioneering study using this approach, similar stimulation applied in RA could interrupt an ongoing note, but not reset the motif, supporting a model in which pattern generators are hierarchically organized, with HVC specifying global patterning of the motif and RA neurons encoding the structure of individual notes or syllables [237]. However, subsequent experiments have elicited motif resetting not only by focal stimulation applied in RA, as well as HVC, but also by stimulating in regions of the rostroventral medulla that contain nucleus parambiguus [240]. Along with the finding that bilateral lesions placed in Uva severely and permanently disrupt song, these more recent microstimulation studies lend support to a recurrent model of song production.

(iii) Focal cooling of HVC slows down song

Yet another highly inventive approach has been to focally cool HVC in singing zebra finches with a miniature Peltier device [218] (figure 5). A quite extraordinary set of behavioural observations show that cooling HVC bilaterally over a moderate temperature range (less than 10°C) slows song timing without affecting the spectral or amplitude features of the song [218]. The stretching effect is close to uniform across timescales ranging from milliseconds to seconds, with single syllables and hence the entire motif stretching

approximately 3% per degree of cooling [218,244]. Whereas cooling stretches syllable gaps less than syllables, the remarkably selective effects of temperature on timing indicate that a major function of HVC is to encode—quite selectively—the temporal aspects of birdsong. As qualitatively similar effects on human speech result from focal cooling of Broca’s region [217] (figure 5), this regionalized control of vocal timing in humans and songbirds may reflect a common strategy for encoding temporal features of learned vocalizations.

While the effects of HVC cooling on song timing have been interpreted as experimental confirmation that HVC_{RA} neurons are linked together to form feedforward excitatory (synfire) chains, some important caveats remain. First, the effects of temperature on song timing are quite modest, in that the neural machinery comprising a synfire chain (i.e. action potentials and chemical synapses) show a much steeper temperature dependence (10% per degree cooling) [244]. In other words, if synapses and action potentials in HVC constitute a synfire chain for song timing, cooling HVC 10°C should double the length of a motif, rather than stretching it slightly less than a third, as actually observed. Second, cooling RA [218] or Uva [244] with implanted Peltier probes also stretches the bird’s song. Whereas the effects of RA cooling on song timing were largely attributed to indirect cooling of HVC by the implanted probe, song still stretched when Uva was cooled and the radiative cooling effect on HVC from the implanted probes was compensated for by simultaneously warming HVC with a surface Peltier [218]. Finally, while modest cooling stretches song more or less uniformly, more extreme cooling of HVC applied in canaries causes individual syllables to break into smaller fragments [244]. Notably, breaks form when a single expiratory pulse of air underlying the original syllable is interrupted by a short inhalation [253], a process that presumably reflects the interaction of respiratory premotor networks in the brainstem with HVC.

(d) Duetting

Duetting is a special form of antiphonal singing between two individuals, most typically mated partners. Duetting has been described for many species of birds and several species of gibbons [254,255]. Duets comprise a vocalization initiated by one individual that is answered by its partner such that their vocalizations tightly overlap or alternate. Duetting therefore requires a special coordination of vocal timing and song type between partners [256]. Unlike most other primates, including other great apes, duetting occurs in most species of gibbons, where mated pairs characteristically combine their songs in a relatively rigid pattern to produce coordinated duet songs [255].

The production of such highly coordinated antiphonal singing involves fine-scale adjustments of the vocal elements that the partners produce through the course of the duet. Since these adjustments require integration of vocal production and perception on a scale of milliseconds, highly developed pathways linking the auditory and vocal pathways can be expected. While the neuronal correlates of duet singing in gibbons are unknown, recent studies provide a first glimpse into brain mechanisms for antiphonal calling in songbirds. Perhaps not surprisingly, given the precise interplay between hearing and vocalization, the song production system is directly involved in antiphonal calling. In a particularly clever experiment, male and female zebra

finches learn to precisely time their calls to avoid the ‘jamming’ calls of a vocal ‘robot,’ but were no longer able to avoid jamming after RA lesions [235]. This indicates that the descending song-production pathway functions as a general-purpose sensorimotor communication system for both antiphonal calls as well as the acquisition and production of learned songs. Interestingly, calls in males and females are innate, raising the possibility that the song system initially evolved to facilitate antiphonal calling, an auditory–vocal capacity that may have led eventually to a more flexible and profound capacity for vocal imitation.

An arguably more sophisticated form of duetting involves the elaborate song duets produced by tropical wrens. A pioneering neurophysiological recording study performed in anesthetized plain-tailed wrens (*Pheugopedius euophrys*), a songbird species in which male and female breeding pairs sing well-coordinated duets, found that HVC neurons were not only responsive to playback of the bird’s own part of the duet but also to the partner’s vocalizations [257]. Most importantly, neurons in the wren’s HVC responded most strongly to presentations of the complete duet sequence, suggesting that auditory responses in HVC to the partner’s half of the duet might be important for the precise coordination of vocal premotor activity during duetting [257]. That HVC in duetting songbirds might rapidly switch between premotor and auditory roles was a surprising idea, because HVC activity in awake (nonduetting) songbirds is mainly premotor, and responses of HVC neurons to auditory stimulation are largely or entirely suppressed during singing [242,248,258,259]. Neural recordings in the HVC of actively duetting birds were needed to test this ‘multiplex’ model.

Indeed, recent multi-unit recordings in freely-moving and duetting white-browed sparrow-weaver (*Plocepasser mahali*), another duetting songbird, challenge this multiplex model [260]. In actively duetting sparrow-weavers, the neural activity in HVC was exclusively premotor during singing, just as in non-duetting birds. However, the auditory information generated by the duet partner somehow alters the temporal parameters of HVC activity in the duet-initiating bird, thus enabling the birds to alternate their vocalizations [260]. Such studies underscore that simultaneous measurements of brain activity in pairs or groups of individuals during natural vocal communication are indispensable to understand the neural mechanisms that underlie duetting and other forms of antiphonal vocalizations, including call and response singing and conversational speech in humans.

6. Concluding remarks

A comparative approach using primates, rodents, other mammals and birds is highly advantageous for understanding the central and peripheral control of vocalization. Each group brings distinct advantages to such an understanding, ranging

from the close homology to human vocal circuits provided by vocal pathways in non-human primates, to genetic tractability of the mouse and the close parallels to human speech learning afforded by studies of birdsong learning. The increasing ease of applying genetic, physiological and optical methods across different animal models should increase the relevance and power of these comparative approaches.

Human speech is a tremendously complex behaviour but has clear antecedents in the vocal behaviours of other vertebrates. The conservation of the vertebrate brainstem and indeed much of the subcortical forebrain means that human speech is built on a general platform for vocalization that solved the fundamental problem of how to integrate vocal motor and respiratory activity. Speech and birdsong both require cooperation and coordination between motor cortical elements specialized for learned vocal control and these more ancestral brainstem vocal–respiratory networks. While much work remains, studies in songbirds are beginning to provide the clearest insights into how such forebrain–brainstem coordination is achieved.

An important goal will be to extend dynamical manipulations, such as focal cooling of cortical regions, to studies of vocalizations in non-human primates and rodents. Such perturbation experiments are necessary to move beyond lesion- and inactivation-based ‘necessity’ experiments. The ability to produce normal vocal patterns in the absence of a cortex does not rule out cortical contributions to vocal modulation, contextual engagement of vocalization and associative processes in which animals may learn to vocalize in response to conditioning.

The rudimentary capacities that gave rise to our highly flexible vocal abilities may be found in context-dependent vocalizations that can be produced by songbirds, non-human primates and certain rodents. That is, the transition from a purely affective vocal gating mechanism to the production of antiphonal vocalizations triggered by the sound of another conspecific’s vocalizations may represent a bridge to more complex auditory–motor interactions that ultimately form the foundation for vocal learning.

In a similar vein, the ability to learn through conditioning to vocalize in response to an arbitrary sensory stimulus, as has been achieved in non-human primates, may provide a stepping stone for understanding how more complex associations between context and vocal behaviour are achieved. Therefore, a useful goal will be to systematically explore how widespread the capacity for conditioned vocalizations may be in other mammals and birds.

Data accessibility. This article has no additional data.

Authors’ contributions. Both A.N. and R.M. wrote this review.

Competing interests. We declare we have no competing interests.

Funding. R.M. was funded by grants NIH R01 DC013826, MH117778, NS099288 and NSF 1354962. A.N. was funded by grants DFG NI 618/4.1 and NI 618/6.1.

References

1. Ficken MS, Ficken RW, Witkin SR. 1978 Vocal repertoire of the Black-capped Chickadee. *Auk* **95**, 34–48. (doi:10.2307/4085493)
2. 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. (doi:10.1126/science.7433999)
3. Price T, Wadewitz P, Cheney D, Seyfarth R, Hammerschmidt K, Fischer J. 2015 Vervets revisited: a quantitative analysis of alarm call structure and context specificity. *Sci. Rep.* **5**, 13220. (doi:10.1038/srep13220)
4. Winter P, Handley P, Ploog D, Schott D. 1973 Ontogeny of squirrel monkey calls under normal

- conditions and under acoustic isolation. *Behaviour* **47**, 230–239. (doi:10.1163/156853973X00085)
5. Brockelman WY, Schilling D. 1984 Inheritance of stereotyped gibbon calls. *Nature* **312**, 634–636. (doi:10.1038/312634a0)
 6. Owren MJ, Dieter JA, Seyfarth RM, Cheney DL. 1992 Food calls produced by adult female rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques, their normally-raised offspring, and offspring cross-fostered between species. *Behaviour* **120**, 218–231. (doi:10.1163/156853992X00615)
 7. Kikusui T, Nakanishi K, Nakagawa R, Nagasawa M, Mogi K, Okanoya K. 2011 Cross fostering experiments suggest that mice songs are innate. *PLoS ONE* **6**, e17721. (doi:10.1371/journal.pone.0017721)
 8. Hammerschmidt K, Reisinger E, Westekemper K, Ehrenreich L, Strenzke N, Fischer J. 2012 Mice do not require auditory input for the normal development of their ultrasonic vocalizations. *BMC Neurosci.* **13**, 40. (doi:10.1186/1471-2202-13-40)
 9. Mahrt EJ, Perkel DJ, Tong L, Rubel EW, Portfors CV. 2013 Engineered deafness reveals that mouse courtship vocalizations do not require auditory experience. *J. Neurosci.* **33**, 5573–5583. (doi:10.1523/JNEUROSCI.5054-12.2013)
 10. Arriaga G, Zhou EP, Jarvis ED. 2012 Of mice, birds, and men: the mouse ultrasonic song system has some features similar to humans and song-learning birds. *PLoS ONE* **7**, e46610. (doi:10.1371/journal.pone.0046610)
 11. Eibl-Eibesfeldt I. 1973 The expressive behaviour of the deaf-and-blind-born. In *Social communication and movement* (eds M Von Cranach, I Vine), pp. 163–194. London, UK: Academic Press.
 12. Takahashi DY, Fenley AR, Teramoto Y, Narayanan DZ, Borjon JJ, Holmes P, Ghazanfar AA. 2015 LANGUAGE DEVELOPMENT. The developmental dynamics of marmoset monkey vocal production. *Science* **349**, 734–738. (doi:10.1126/science.aab1058)
 13. Gultekin YB, Hage SR. 2017 Limiting parental feedback disrupts vocal development in marmoset monkeys. *Nat. Commun.* **16**, 14046. (doi:10.1038/ncomms14046)
 14. Gultekin YB, Hage SR. 2018 Limiting parental interaction during vocal development affects acoustic call structure in marmoset monkeys. *Sci. Adv.* **4**, eaar4012. (doi:10.1126/sciadv.aar4012)
 15. Takahashi DY, Liao DA, Ghazanfar AA. 2017 Vocal learning via social reinforcement by infant marmoset monkeys. *Curr. Biol.* **27**, 1844–1852. (doi:10.1016/j.cub.2017.05.004)
 16. Janik VM, Slater PJB. 1997 Vocal learning in mammals. *Adv. Study Behav.* **26**, 59–99. (doi:10.1016/S0065-3454(08)60377-0)
 17. Takahashi DY, Liao DA, Ghazanfar AA. 2019 Volition and learning in primate vocal behaviour. *Anim. Behav.* **151**, 239–247. (doi:10.1016/j.anbehav.2019.01.021)
 18. Tyack PL. 2016 Vocal learning and auditory–vocal feedback. In *Vertebrate sound production and acoustic communication* (eds RA Suthers, WT Fitch, RR Fay, AN Popper), Springer Handbook of Auditory Research **53**, pp. 261–298. Heidelberg, Germany: Springer. (doi:10.1007/978-3-319-27721-9_9)
 19. Goldstein MH, King AP, West MJ. 2003 Social interaction shapes babbling: testing parallels between birdsong and speech. *Proc. Natl Acad. Sci. USA* **100**, 8030–8035. (doi:10.1073/pnas.1332441100)
 20. Goldstein MH, Schwade JA. 2008 Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol. Sci.* **19**, 515–523. (doi:10.1111/j.1467-9280.2008.02117.x)
 21. Chen Y, Matheson LE, Sakata JT. 2016 Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proc. Natl Acad. Sci. USA* **113**, 6641–6646. (doi:10.1073/pnas.1522306113)
 22. Carouso-Peck S, Goldstein MH. 2019 Female social feedback reveals non-imitative mechanisms of vocal learning in zebra finches. *Curr. Biol.* **29**, 631–636. (doi:10.1016/j.cub.2018.12.026)
 23. Prat Y, Taub M, Yovel Y. 2015 Vocal learning in a social mammal: demonstrated by isolation and playback experiments in bats. *Sci. Adv.* **1**, e1500019. (doi:10.1126/sciadv.1500019)
 24. Prat Y, Azoulay L, Dor R, Yovel Y. 2017 Crowd vocal learning induces vocal dialects in bats: playback of conspecifics shapes fundamental frequency usage by pups. *PLoS Biol.* **15**, e2002556. (doi:10.1371/journal.pbio.2002556)
 25. Snowdon CT. 2017 Learning from monkey 'talk'. *Science* **355**, 1120–1122. (doi:10.1126/science.aam7443)
 26. Petkov CI, Jarvis ED. 2012 Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Front. Evol. Neurosci.* **4**, 12. (doi:10.3389/fnevo.2012.00012)
 27. Thorpe W. 1958 The learning of song patterns by birds, with especial reference to the song of the chaffinch. *Ibis* **100**, 535–570. (doi:10.1111/j.1474-919X.1958.tb07960.x)
 28. Marler P, Tamura M. 1964 Culturally transmitted patterns of vocal behaviour in sparrows. *Science* **146**, 1483–1486. (doi:10.1126/science.146.3650.1483)
 29. Immelmann K. 1969 Song development in zebra finch and other estrildid finches. In *Bird vocalisations* (ed. RA Hinde), pp. 61–74. London, UK: Cambridge University Press.
 30. Gramza A. 1970 Vocal mimicry in captive budgerigars (*Melopsittacus undulatus*). *Z. Tierpsychol.* **27**, 971–983. (doi:10.1111/j.1439-0310.1970.tb01913.x)
 31. Pepperberg IM. 1981 Functional vocalizations by an African grey parrot (*Psittacus erithacus*). *Z. Tierpsychol.* **55**, 139–160. (doi:10.1111/j.1439-0310.1981.tb01265.x)
 32. Farabaugh SM, Linzenbold A, Dooling RJ. 1994 Vocal plasticity in budgerigars (*Melopsittacus undulatus*): evidence for social factors in the learning of contact calls. *J. Comp. Psychol.* **108**, 81–92. (doi:10.1037/0735-7036.108.1.81)
 33. Baptista LF, Schuchmann KL. 1990 Song learning in the Anna hummingbird (*Calypte anna*). *Ethology* **84**, 15–26. (doi:10.1111/j.1439-0310.1990.tb00781.x)
 34. Reichmuth C, Casey C. 2014 Vocal learning in seals, sea lions, and walrus. *Curr. Opin. Neurobiol.* **28**, 66–71. (doi:10.1016/j.conb.2014.06.011)
 35. Janik VM. 2014 Cetacean vocal learning and communication. *Curr. Opin. Neurobiol.* **28**, 60–65. (doi:10.1016/j.conb.2014.06.010)
 36. Knörnschild M. 2014 Vocal production learning in bats. *Curr. Opin. Neurobiol.* **28**, 80–85. (doi:10.1016/j.conb.2014.06.014)
 37. Stoeger AS, Manger P. 2014 Vocal learning in elephants: neural bases and adaptive context. *Curr. Opin. Neurobiol.* **28**, 101–107. (doi:10.1016/j.conb.2014.07.001)
 38. Fischer J, Wheeler BC, Higham JP. 2015 Is there any evidence for vocal learning in chimpanzee food calls? *Curr. Biol.* **25**, R1028–R1029. (doi:10.1016/j.cub.2015.09.010)
 39. Marler P. 1970 Birdsong and speech development: could there be parallels? *Am. Sci.* **58**, 669–673.
 40. Locke JL. 1993 *The child's path to spoken language*. Cambridge, MA: Harvard University Press.
 41. Doupe AJ, Kuhl PK. 1999 Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* **22**, 567–631. (doi:10.1146/annurev.neuro.22.1.567)
 42. Tschida K, Mooney R. 2012 The role of auditory feedback in vocal learning and maintenance. *Curr. Opin. Neurobiol.* **22**, 320–327. (doi:10.1016/j.conb.2011.11.006)
 43. Brainard MS, Doupe AJ. 2013 Translating birdsong: songbirds as a model for basic and applied medical research. *Annu. Rev. Neurosci.* **36**, 489–517. (doi:10.1146/annurev-neuro-060909-152826)
 44. Wild JM. 1997 Neural pathways for the control of birdsong production. *J. Neurobiol.* **33**, 653–670. (doi:10.1002/(SICI)1097-4695(19971105)33:5<653::AID-NEU11>3.0.CO;2-A)
 45. Mooney R. 2009 Neural mechanisms for learned birdsong. *Learn. Mem.* **16**, 655–669. (doi:10.1101/lm.1065209)
 46. Fee MS, Goldberg JH. 2011 A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience* **198**, 152–170. (doi:10.1016/j.neuroscience.2011.09.069)
 47. Simonyan K, Horwitz B. 2011 Laryngeal motor cortex and control of speech in humans. *Neuroscientist* **17**, 197–208. (doi:10.1177/1073858410386727)
 48. Todt D. 1986 Hinweis-Charakter und Mittler-Funktion von Verhalten. *Zeitschrift für Semiotik* **8**, 183–232.
 49. Fischer J, Price T. 2017 Meaning, intention, and inference in primate vocal communication. *Neurosci. Biobehav. Rev.* **82**, 22–31. (doi:10.1016/j.neubiorev.2016.10.014)
 50. Jürgens U. 1979 Vocalization as an emotional indicator—a neuroethological study in the squirrel monkey. *Behaviour* **69**, 88–117. (doi:10.1163/156853979X00412)
 51. Liao DA, Zhang YS, Cai LX, Ghazanfar AA. 2018 Internal states and extrinsic factors both determine monkey vocal production. *Proc. Natl Acad. Sci. USA* **115**, 3978–3983. (doi:10.1073/pnas.1722426115)

52. Hopf HC, Müller-Forell W, Hopf NJ. 1992 Localization of emotional and volitional facial paresis. *Neurology* **42**, 1918–1923. (doi:10.1212/WNL.42.10.1918)
53. Cattaneo L, Pavesi G. 2014 The facial motor system. *Neurosci. Biobehav. Rev.* **38**, 135–159. (doi:10.1016/j.neubiorev.2013.11.002)
54. Broca P. 1861 Remarques sur le siege de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole) [Remarks on the seat of the faculty of articulated language, following an observation of aphemia (loss of speech)]. *Bull. Mem. Soc. Anat. Paris* **36**, 330–357. (doi:10.3406/bmsap.1865.9495)
55. Groswasser Z, Korn C, Groswasser-Reider I, Solzi P. 1988 Mutism associated with buccofacial apraxia and bihemispheric lesions. *Brain Lang.* **34**, 157–168. (doi:10.1016/0093-934X(88)90129-0)
56. Caplan D, Hildebrandt N, Makris N. 1996 Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain* **119**, 933–994. (doi:10.1093/brain/119.3.933)
57. Lazar RM, Mohr JP. 2011 Revisiting the contributions of Paul Broca to the study of aphasia. *Neuropsychol. Rev.* **21**, 236–239. (doi:10.1007/s11065-011-9176-8)
58. Rohrer JD, Warren JD, Rossor MN. 2009 Abnormal laughter-like vocalisations replacing speech in primary progressive aphasia. *J. Neurol. Sci.* **284**, 120–123. (doi:10.1016/j.jns.2009.04.021)
59. Townsend SW *et al.* 2017 Exorcising Grice's ghost: an empirical approach to studying intentional communication in animals. *Biol. Rev. Camb. Philos. Soc.* **92**, 1427–1433. (doi:10.1111/brv.12289)
60. Hage SR, Gavrilov N, Nieder A. 2013 Cognitive control of distinct vocalizations in rhesus monkeys. *J. Cogn. Neurosci.* **25**, 1692–1701. (doi:10.1162/jocn_a_00428)
61. Brecht KF, Hage SR, Gavrilov N, Nieder A. 2019 Volitional control of vocalizations in corvid songbirds. *PLoS Biol.* **17**, e3000375. (doi:10.1371/journal.pbio.3000375)
62. Hage SR, Gavrilov N, Nieder A. 2016 Developmental changes of cognitive vocal control in monkeys. *J. Exp. Biol.* **219**, 1744–1749. (doi:10.1242/jeb.137653)
63. Rakic P, Bourgeois J-P, Eckenhoff MF, Zecevic N, Goldman-Rakic PS. 1986 Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex. *Science* **232**, 232–235. (doi:10.1126/science.3952506)
64. Titze I. 1994 *Principles of voice production*. Englewood Cliffs, NJ: Prentice Hall.
65. Suthers RA, Goller F, Pytte C. 1999 The neuromuscular control of birdsong. *Phil. Trans. R. Soc. Lond. B* **354**, 927–939. (doi:10.1098/rstb.1999.0444)
66. Goller F, Larsen ON. 2002 New perspectives on mechanisms of sound generation in songbirds. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **188**, 841–850. (doi:10.1007/s00359-002-0350-6)
67. Riede T, Goller F. 2010 Peripheral mechanisms for vocal production in birds—differences and similarities to human speech and singing. *Brain Lang.* **115**, 69–80. (doi:10.1016/j.bandl.2009.11.003)
68. Goller F, Riede T. 2013 Integrative physiology of fundamental frequency control in birds. *J. Physiol. Paris* **107**, 230–242. (doi:10.1016/j.jphysparis.2012.11.001)
69. Elemans CP *et al.* 2015 Universal mechanisms of sound production and control in birds and mammals. *Nat. Commun.* **6**, 8978. (doi:10.1038/ncomms9978)
70. Nakamura F, Uyeda Y, Sonoda Y. 1958 Electromyographic study on respiratory movements of the intrinsic laryngeal muscles. *Laryngoscope* **68**, 109–119. (doi:10.1288/00005537-195802000-00002)
71. Megirian D, Sherrey JH. 1980 Respiratory functions of the laryngeal muscle during sleep. *Sleep* **3**, 289–298. (doi:10.1093/sleep/3.3-4.289)
72. Vicario DS, Simpson HB. 1995 Electrical stimulation in forebrain nuclei elicits learned vocal patterns in songbirds. *J. Neurophysiol.* **73**, 2602–2607. (doi:10.1152/jn.1995.73.6.2602)
73. Goller F, Suthers RA. 1996 Role of syringeal muscles in gating airflow and sound production in singing brown thrashers. *J. Neurophysiol.* **75**, 867–876. (doi:10.1152/jn.1996.75.2.867)
74. Hage SR, Jurgens U. 2006 On the role of the pontine brainstem in vocal pattern generation: a telemetric single-unit recording study in the squirrel monkey. *J. Neurosci.* **26**, 7105–7115. (doi:10.1523/JNEUROSCI.1024-06.2006)
75. Jurgens U, Hage SR. 2007 On the role of the reticular formation in vocal pattern generation. *Behav. Brain Res.* **182**, 308–314. (doi:10.1016/j.bbr.2006.11.027)
76. Schmidt MF, McLean J, Goller F. 2012 Breathing and vocal control: the respiratory system as both a driver and a target of telencephalic vocal motor circuits in songbirds. *Exp. Physiol.* **97**, 455–461. (doi:10.1113/expphysiol.2011.058669)
77. Ramirez JM, Baertsch NA. 2018 The dynamic basis of respiratory rhythm generation: one breath at a time. *Annu. Rev. Neurosci.* **41**, 475–499. (doi:10.1146/annurev-neuro-080317-061756)
78. Greenwalt C. 1968 *Birdsong: acoustics and physiology*. Washington, DC: Smithsonian University Press.
79. Suthers RA, Goller F, Hartley RS. 1994 Motor dynamics of song production by mimic thrushes. *J. Neurobiol.* **25**, 917–936. (doi:10.1002/neu.480250803)
80. Klingholz, F. 1993 Overtone singing: productive mechanisms and acoustic data. *J. Voice* **7**, 118–122. (doi:10.1016/S0892-1997(05)80340-6)
81. Riede T. 2013 Stereotypic laryngeal and respiratory motor patterns generate different call types in rat ultrasound vocalization. *J. Exp. Zool. A Ecol. Genet. Physiol.* **319**, 213–224. (doi:10.1002/jez.1785)
82. Speakman JD. 1968 Physical analysis of speech in helium environments. *Aerosp. Med.* **39**, 48–53.
83. Nowicki S. 1987 Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. *Nature* **325**, 53–55. (doi:10.1038/325053a0)
84. Fant G. 1970 *Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations*. The Hague, The Netherlands: Mouton.
85. Riede T, Suthers RA, Fletcher NH, Blevins WE. 2006 Songbirds tune their vocal tract to the fundamental frequency of their song. *Proc. Natl. Acad. Sci. USA* **103**, 5543–5548. (doi:10.1073/pnas.0601262103)
86. Nelson BS, Beckers GJ, Suthers RA. 2005 Vocal tract filtering and sound radiation in a songbird. *J. Exp. Biol.* **208**, 297–308. (doi:10.1242/jeb.01378)
87. Sewell GD. 1967 Ultrasound in adult rodents. *Nature* **215**, 512. (doi:10.1038/215512a0)
88. Sewell GD. 1970 Ultrasonic signals from rodents. *Ultrasonics* **8**, 26–30. (doi:10.1016/0041-624X(70)90795-X)
89. Roberts LH. 1975 The rodent ultrasound production mechanism. *Ultrasonics* **13**, 83–88. (doi:10.1016/0041-624X(75)90052-9)
90. Knutson N, Burgdorf J, Panskepp J. 2002 Ultrasonic vocalizations as indices of affective states in rats. *Psychol. Bull.* **128**, 961–977. (doi:10.1037/0033-2909.128.6.961)
91. Riede T, Borgard HL, Pasch B. 2017 Laryngeal airway reconstruction indicates that rodent ultrasonic vocalizations are produced by an edge-tone mechanism. *R. Soc. open sci.* **4**, 170976. (doi:10.1098/rsos.170976)
92. Riede T. 2018 Peripheral vocal motor dynamics and combinatory call complexity of ultrasonic vocal production in rats. In *Handbook of ultrasonic vocalizations: a window into the emotional brain*, vol. 25 (ed. SM Brudzynski), pp. 45–60. London, England: Elsevier. (doi:10.1016/b978-0-12-809600-0.00005-6)
93. Busnel RG, Moles A, Gilbert M. 1962 A case of the whistle language used in the French Pyrenees. *Logos* **5**, 76–91.
94. Nottebohm F, Stokes TM, Leonard CM. 1976 Central control of song in the canary, *Serinus canarius*. *J. Comp. Neurol.* **165**, 457–486. (doi:10.1002/cne.901650405)
95. Aronov D, Andalman AS, Fee MS. 2008 A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science* **320**, 630–634. (doi:10.1126/science.1155140)
96. Alexander MP, Naeser MA, Palumbo C. 1990 Broca's area aphasia: aphasia after lesions including the frontal operculum. *Neurology* **40**, 353–362. (doi:10.1212/WNL.40.2.353)
97. Wild JM. 2004 Functional neuroanatomy of the sensorimotor control of singing. *Ann. N. Y. Acad. Sci.* **1016**, 438–462. (doi:10.1196/annals.1298.016)
98. Jurgens U. 2002 Neural pathways underlying vocal control. *Neurosci. Biobehav. Rev.* **26**, 235–258. (doi:10.1016/S0149-7634(01)00068-9)
99. Ludlow CL. 2005 Central nervous system control of the laryngeal muscles in humans. *Respir. Physiol. Neurobiol.* **147**, 205–222. (doi:10.1016/j.resp.2005.04.015)
100. Wild JM, Zeigler HP. 1980 Central representation and somatotopic organization of the jaw muscles within the facial and trigeminal nuclei of the

- pigeon (*Columba livia*). *J. Comp. Neurol.* **192**, 175–201. (doi:10.1002/cne.901920112)
101. Wild JM, Krutzfeldt NE. 2012 Trigeminal and telencephalic projections to jaw and other upper vocal tract premotor neurons in songbirds: sensorimotor circuitry for beak movements during singing. *J. Comp. Neurol.* **520**, 590–605. (doi:10.1002/cne.22752)
 102. Holstege G, Subramanian HH. 2016 Two different motor systems are needed to generate human speech. *J. Comp. Neurol.* **524**, 1558–1577. (doi:10.1002/cne.23898)
 103. Wild JM. 1993 The avian nucleus retroambigialis: a nucleus for breathing, singing and calling. *Brain Res.* **606**, 319–324. (doi:10.1016/0006-8993(93)91001-9)
 104. Nottebohm F. 1971 Neural lateralization of vocal control in a passerine bird. I. Song. *J. Exp. Zool.* **177**, 229–261. (doi:10.1002/jez.1401770210)
 105. Nottebohm F, Arnold AP. 1976 Sexual dimorphism in vocal control areas of the songbird brain. *Science* **194**, 211–213. (doi:10.1126/science.959852)
 106. Williams H, Crane LA, Hale TK, Esposito MA, Nottebohm F. 1992 Right-side dominance for song control in the zebra finch. *J. Neurobiol.* **23**, 1006–1020. (doi:10.1002/neu.480230807)
 107. Wild J, Williams M, Suthers R. 2000 Neural pathways for bilateral vocal control in songbirds. *J. Comp. Neurol.* **423**, 413–426. (doi:10.1002/1096-9861(20000731)423:3<413::AID-CNE5>3.0.CO;2-7)
 108. Jürgens U, Pratt R. 1979 Role of the periaqueductal grey in vocal expression of emotion. *Brain Res.* **167**, 367–378. (doi:10.1016/0006-8993(79)90830-8)
 109. Wild JM, Li D, Eagleton C. 1997 Projections of the dorsomedial nucleus of the intercollicular complex (DM) in relation to respiratory-vocal nuclei in the brainstem of pigeon (*Columba livia*) and zebra finch (*Taeniopygia guttata*). *J. Comp. Neurol.* **377**, 392–413. (doi:10.1002/(SICI)1096-9861(19970120)377:3<392::AID-CNE7>3.0.CO;2-Y)
 110. Armitage SE, Seller TJ. 1978 Low threshold sites for vocalizations in the mid-brain of Japanese quail [proceedings]. *J. Physiol.* **278**, 24–25.
 111. Seller TJ. 1980 Midbrain regions involved in call production in Java sparrows. *Behav. Brain Res.* **1**, 257–265. (doi:10.1016/0166-4328(80)90033-9)
 112. Holstege G. 2014 The periaqueductal gray controls brainstem emotional motor systems including respiration. *Prog. Brain Res.* **209**, 379–405. (doi:10.1016/B978-0-444-63274-6.00020-5)
 113. Jürgens U. 1994 The role of the periaqueductal grey in vocal behaviour. *Behav. Brain Res.* **62**, 107–117. (doi:10.1016/0166-4328(94)90017-5)
 114. Gurney M. 1981 Hormonal control of cell form and number in the zebra finch song system. *J. Neurosci.* **1**, 658–673. (doi:10.1523/JNEUROSCI.01-06-00658.1981)
 115. Ball GF, Balthazart J. 2004 Hormonal regulation of brain circuits mediating male sexual behavior in birds. *Physiol. Behav.* **83**, 329–346. (doi:10.1016/j.physbeh.2004.08.020)
 116. Esposito A, Demeurisse G, Alberti B, Fabbro F. 1999 Complete mutism after midbrain periaqueductal gray lesion. *Neuroreport* **10**, 681–685. (doi:10.1097/00001756-199903170-00004)
 117. Ortega JD, DeRosier E, Park S, Larson CR. 1988 Brainstem mechanisms of laryngeal control as revealed by microstimulation studies. In *Vocal physiology: voice production, mechanisms and functions* (ed. O Fujimura), pp. 19–28. New York, NY: Raven.
 118. Brown TG. 1915 Note on physiology of basal ganglia and midbrain of anthropoid ape especially in reference to act of laughter. *J. Physiol.* **49**, 195–207. (doi:10.1113/jphysiol.1915.sp001703)
 119. Magoun HW, Atlas D, Ingersoll EH, Ranson SW. 1937 Associated facial, vocal and respiratory components of emotional expression: an experimental study. *J. Neurol. Psychopathol.* **17**, 241–255. (doi:10.1136/jnnp.s1-17.67.241)
 120. Jürgens U, Ploog D. 1970 Cerebral representation of vocalization in the squirrel monkey. *Exp. Brain Res.* **10**, 532–554. (doi:10.1007/BF00234269)
 121. Apfelbach R. 1972 Electrically elicited vocalization in the gibbon *Hylobates lar* (*Hylobatidae*), and their behavioral significance. *Z. Tierpsychol.* **30**, 420–430. (doi:10.1111/j.1439-0310.1972.tb00868.x)
 122. Jürgens U. 2009 The neural control of vocalization in mammals: a review. *J. Voice* **23**, 1–10. (doi:10.1016/j.jvoice.2007.07.005)
 123. Carrive P. 1993 The periaqueductal gray and defensive behavior: functional representation and neuronal organization. *Behav. Brain Res.* **58**, 27–47. (doi:10.1016/0166-4328(93)90088-8)
 124. Bandler R, Shipley MT. 1994 Columnar organization in the midbrain periaqueductal gray: modules for emotional expression? *Trends Neurosci.* **17**, 379–389. (doi:10.1016/0166-2236(94)90047-7)
 125. Tschida K, Michael V, Takato J, Han B-X, Zhao S, Sakurai K, Mooney R, Wang F. 2019 A specialized neural circuit gates social vocalizations in the mouse. *Neuron* **103**, 459–472. (doi:10.1016/j.neuron.2019.05.025)
 126. Sakurai K, Zhao S, Takato J, Rodriguez E, Lu J, Leavitt AD, Fu M, Han BX, Wang F. 2016 Capturing and manipulating activated neuronal ensembles with CANE delineates a hypothalamic social-fear circuit. *Neuron* **92**, 739–753. (doi:10.1016/j.neuron.2016.10.015)
 127. Kingsbury MA, Kelly AM, Schrock SE, Goodson JL. 2011 Mammal-like organization of the avian midbrain central gray and a reappraisal of the intercollicular nucleus. *PLoS ONE* **6**, e20720. (doi:10.1371/journal.pone.0020720)
 128. Hage SR, Nieder A. 2016 Dual neural network model for the evolution of speech and language. *Trends Neurosci.* **39**, 813–829. (doi:10.1016/j.tins.2016.10.006)
 129. Simonyan K, Jürgens U. 2005 Afferent cortical connections of the motor cortical larynx area in the rhesus monkey. *Neuroscience* **130**, 133–149. (doi:10.1016/j.neuroscience.2004.08.031)
 130. Morecraft R, Tanji J. 2009 Cingulofrontal interactions and the cingulate motor areas. In *Cingulate neurobiology and disease* (ed. B Vogt), pp. 113–144. New York, NY: Oxford University Press.
 131. Robinson BW. 1967 Vocalization evoked from forebrain in *Macaca mulatta*. *Physiol. Behav.* **2**, 345–346. (doi:10.1016/0031-9384(67)90050-9)
 132. Jürgens U. 1976 Reinforcing concomitants of electrically elicited vocalizations. *Exp. Brain Res.* **26**, 203–214. (doi:10.1007/BF00238284)
 133. Müller-Preuss P, Newman JD, Jürgens U. 1980 Anatomical and physiological evidence for a relationship between the ‘cingular’ vocalization area and the auditory cortex in the squirrel monkey. *Brain Res.* **202**, 307–315. (doi:10.1016/0006-8993(80)90143-2)
 134. Newman JD. (ed). 1988 *The physiological control of mammalian vocalization*. New York, NY: Plenum Press.
 135. West RA, Larson CR. 1995 Neurons of the anterior mesial cortex related to faciovocal activity in the awake monkey. *J. Neurophysiol.* **74**, 1856–1869. (doi:10.1152/jn.1995.74.5.1856)
 136. Gavrilov N, Hage SR, Nieder A. 2017 Functional specialization of the primate frontal lobe during cognitive control of vocalizations. *Cell Rep.* **21**, 2393–2406. (doi:10.1016/j.celrep.2017.10.107)
 137. Franzen EA, Myers RE. 1973 Neural control of social behavior: prefrontal and anterior temporal cortex. *Neuropsychology* **11**, 141–157. (doi:10.1016/0028-3932(73)90002-X)
 138. Trachy RE, Sutton D, Lindeman RC. 1981 Primate phonation: anterior cingulate lesion effects on response rate and acoustical structure. *Am. J. Primatol.* **1**, 43–55. (doi:10.1002/ajp.1350010106)
 139. Kirzinger A, Jürgens U. 1982 Cortical lesion effects and vocalization in the squirrel monkey. *Brain Res.* **233**, 299–315. (doi:10.1016/0006-8993(82)91204-5)
 140. MacLean PD, Newman JD. 1988 Role of midline frontolimbic cortex in production of the isolation call in squirrel monkeys. *Brain Res.* **450**, 111–123. (doi:10.1016/0006-8993(88)91550-8)
 141. Sutton D, Larson C, Lindeman RC. 1974 Neocortical and limbic lesion effects on primate phonation. *Brain Res.* **71**, 61–75. (doi:10.1016/0006-8993(74)90191-7)
 142. Aitken PG. 1981 Cortical control of conditioned and spontaneous vocal behavior in rhesus monkeys. *Brain Lang.* **13**, 171–184. (doi:10.1016/0093-934X(81)90137-1)
 143. Wild B, Rodden FA, Grodd W, Ruch W. 2003 Neural correlates of laughter and humour. *Brain* **126**, 2121–2138. (doi:10.1093/brain/awg226)
 144. Kovac S, Deppe M, Mohammadi S, Schifffbauer H, Schwindt W, Möddel G, Dogan M, Evers S. 2009 Gelastic seizures: a case of lateral frontal lobe epilepsy and review of the literature. *Epilepsy Behav.* **15**, 249–253. (doi:10.1016/j.yebeh.2009.03.004)
 145. Kuzniecky R, Guthrie B, Mountz J, Bebin M, Faught E, Gilliam F, Liu HG. 1997 Intrinsic epileptogenesis of hypothalamic hamartomas in gelastic epilepsy. *Ann. Neurol.* **42**, 60–67. (doi:10.1002/ana.410420111)
 146. Speri F, Spinelli L, Pollo C, Seeck M. 2006 Contralateral smile and laughter, but no mirth, induced by electrical stimulation of the cingulate

- cortex. *Epilepsia* **47**, 440–443. (doi:10.1111/j.1528-1167.2006.00442.x)
147. Nielsen JM, Jacobs LL. 1951 Bilateral lesions of the anterior cingulate gyri; report of case. *Bull. Los Angel. Neuro. Soc.* **16**, 231–234.
148. Barris RW, Schuman HR. 1953 Bilateral anterior cingulate gyrus lesions; syndrome of the anterior cingulate gyri. *Neurology* **3**, 44–52. (doi:10.1212/WNL.3.1.44)
149. Jürgens U, Von Cramon D. 1982 On the role of the anterior cingulate cortex in phonation: a case report. *Brain Lang.* **15**, 234–248. (doi:10.1016/0093-934X(82)90058-X)
150. Frühholz S, Klass HS, Patel S, Grandjean D. 2015 Talking in fury: the cortico-subcortical network underlying angry vocalizations. *Cereb. Cortex* **25**, 2752–2762. (doi:10.1093/cercor/bhu074)
151. Jürgens U, Kirzinger A, von Cramon D. 1982 The effects of deep-reaching lesions in the cortical face area on phonation. A combined case report and experimental monkey study. *Cortex* **18**, 125–139. (doi:10.1016/S0010-9452(82)80024-5)
152. Simonyan K, Jürgens U. 2003 Efferent subcortical projections of the laryngeal motor cortex in the rhesus monkey. *Brain Res.* **974**, 43–59. (doi:10.1016/S0006-8993(03)02548-4)
153. Simonyan K, Ostuni J, Ludlow CL, Horwitz B. 2009 Functional but not structural networks of the human laryngeal motor cortex show left hemispheric lateralization during syllable but not breathing production. *J. Neurosci.* **29**, 14 912–14 923. (doi:10.1523/JNEUROSCI.4897-09.2009)
154. Alajouanine T, Thurel R. 1932 La diplégie faciale (forme corticale de la paralysie pseudobulbaire). *Rev. Neurol.* **38**, 516–517.
155. Green HD, Walker AE. 1938 The effects of ablation of the cortical motor face area in monkeys. *J. Neurophysiol.* **1**, 262–280. (doi:10.1152/jn.1938.1.3.262)
156. Simonyan K. 2014 The laryngeal motor cortex: its organization and connectivity. *Curr. Opin. Neurobiol.* **28**, 15–21. (doi:10.1016/j.conb.2014.05.006)
157. Bouchard KE, Mesgarani N, Johnson K, Chang EF. 2013 Functional organization of human sensorimotor cortex for speech articulation. *Nature* **495**, 327–332. (doi:10.1038/nature11911)
158. Breshears JD, Molinaro AM, Chang EF. 2015 A probabilistic map of the human ventral sensorimotor cortex using electrical stimulation. *J. Neurosurg.* **123**, 340–349. (doi:10.3171/2014.11.JNS14889)
159. Belyk M, Brown S. 2017 The origins of the vocal brain in humans. *Neurosci. Biobehav. Rev.* **77**, 177–193. (doi:10.1016/j.neubiorev.2017.03.014)
160. Foerster, O. 1931 The cerebral cortex in man. *Lancet* **218**, 309–312. (doi:10.1016/S0140-6736(00)47063-7)
161. Penfield W, Rasmussen T. 1950 *The cerebral cortex of man: a clinical study of localization of function*. New York, NY: Macmillan.
162. Leyton ASF, Sherrington CS. 1917 Observations on the excitable cortex of the chimpanzee, orang-utan and gorilla. *Q. J. Exp. Physiol.* **11**, 135–222. (doi:10.1113/expphysiol.1917.sp000240)
163. Hines M. 1940 Movements elicited from precentral gyrus of adult chimpanzees by stimulation with sine wave currents. *J. Neurophysiol.* **3**, 442–466. (doi:10.1152/jn.1940.3.5.442)
164. Dusser de Barenne J et al. 1941 The ‘motor’ cortex of the chimpanzee. *J. Neurophysiol.* **4**, 287–303. (doi:10.1152/jn.1941.4.4.287)
165. Hast MH, Fischer JM, Wetzell AB, Thompson VE. 1974 Cortical motor representation of the laryngeal muscles in *Macaca mulatta*. *Brain Res.* **73**, 229–240. (doi:10.1016/0006-8993(74)91046-4)
166. Jürgens U, Ehrenreich L. 2007 The descending motorcortical pathway to the laryngeal motoneurons in the squirrel monkey. *Brain Res.* **1148**, 90–95. (doi:10.1016/j.brainres.2007.02.020)
167. Jürgens U. 1976 Projections from the cortical larynx area in the squirrel monkey. *Exp. Brain Res.* **25**, 401–411. (doi:10.1007/BF00241730)
168. Iwatsubo T, Kuzuhara S, Kanemitsu A, Shimada H, Toyokura Y. 1990 Corticofugal projections to the motor nuclei of the brainstem and spinal cord in humans. *Neurology* **40**, 309–312. (doi:10.1212/WNL.40.2.309)
169. Kuypers HG. 1958 Corticobulbar connexions to the pons and lower brain-stem in man: an anatomical study. *Brain* **81**, 364–388. (doi:10.1093/brain/81.3.364)
170. Kumar V, Crosson PL, Simonyan K. 2016 Structural organization of the laryngeal motor cortical network and its implication for evolution of speech production. *J. Neurosci.* **36**, 4170–4181. (doi:10.1523/JNEUROSCI.3914-15.2016)
171. Rathelot JA, Strick PL. 2009 Subdivisions of primary motor cortex based on cortico-motoneuronal cells. *Proc. Natl Acad. Sci. USA* **106**, 918–923. (doi:10.1073/pnas.0808362106)
172. Brown S, Ngan E, Liotti M. 2008 A larynx area in the human motor cortex. *Cereb. Cortex* **18**, 837–845. (doi:10.1093/cercor/bhm131)
173. Jürgens U. 1984 The efferent and afferent connections of the supplementary motor area. *Brain Res.* **300**, 63–81. (doi:10.1016/0006-8993(84)91341-6)
174. Tanji J. 1994 The supplementary motor area in the cerebral cortex. *Neurosci. Res.* **19**, 251–268. (doi:10.1016/0168-0102(94)90038-8)
175. Matelli M, Luppino G, Rizzolatti G. 1991 Architecture of superior and mesial area 6 and the adjacent cingulate cortex in the macaque monkey. *J. Comp. Neurol.* **311**, 445–462. (doi:10.1002/cne.903110402)
176. Dum RP, Strick PL. 1991 The origin of corticospinal projections from the premotor areas in the frontal lobe. *J. Neurosci.* **11**, 667–689. (doi:10.1523/JNEUROSCI.11-03-00667.1991)
177. Luppino G, Matelli M, Camarda R, Rizzolatti G. 1993 Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. *J. Comp. Neurol.* **338**, 114–140. (doi:10.1002/cne.903380109)
178. Bates JF, Goldman-Rakic PS. 1993 Prefrontal connections of medial motor areas in the rhesus monkey. *J. Comp. Neurol.* **336**, 211–228. (doi:10.1002/cne.903360205)
179. Wang Y, Isoda M, Matsuzaka Y, Shima K, Tanji J. 2005 Prefrontal cortical cells projecting to the supplementary eye field and presupplementary motor area in the monkey. *Neurosci. Res.* **53**, 1–7. (doi:10.1016/j.neures.2005.05.005)
180. Picard N, Strick PL. 2001 Imaging the premotor areas. *Curr. Opin. Neurobiol.* **11**, 663–672. (doi:10.1016/S0959-4388(01)00266-5)
181. Penfield W, Welch K. 1951 The supplementary motor area of the cerebral cortex. *Arch. Neurol. Psychiatry* **66**, 289–317. (doi:10.1001/archneurpsyc.1951.02320090038004)
182. Erickson TC, Woolsey CN. 1951 Observations on the supplementary motor area of man. *Trans. Am. Neurol. Assoc.* **76**, 50–56.
183. Gelmers HJ. 1983 Non-paralytic motor disturbances and speech disorders: the role of the supplementary motor area. *J. Neurol. Neurosurg. Psychiatry* **46**, 1052–1054. (doi:10.1136/jnnp.46.11.1052)
184. Perry DW, Zatorre RJ, Petrides M, Alivisatos B, Meyer E, Evans AC. 1999 Localization of cerebral activity during simple singing. *Neuroreport* **10**, 3453–3458. (doi:10.1097/00001756-199911080-00035)
185. Bookheimer SY, Zeffiro TA, Blaxton TA, Gaillard PW, Theodore WH. 2000 Activation of language cortex with automatic speech tasks. *Neurology* **55**, 1151–1157. (doi:10.1212/WNL.55.8.1151)
186. Alario FX, Chainay H, Lehericy S, Cohen L. 2006 The role of the supplementary motor area (SMA) in word production. *Brain Res.* **1076**, 129–143. (doi:10.1016/j.brainres.2005.11.104)
187. Sutton D, Trachy RE, Lindeman RC. 1985 Discriminative phonation in macaques: effects of anterior mesial cortex damage. *Exp. Brain Res.* **59**, 410–413. (doi:10.1007/BF00230922)
188. Brodmann K. 1909 Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues. Leipzig, Germany: Barth Verlag.
189. Penfield W, Roberts L. 1959 *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.
190. Friederici A. 2017 *Language in our brain: the origins of a uniquely human capacity*. Cambridge, MA: MIT Press. (doi:10.7551/mitpress/11173.001.0001)
191. Miller EK, Cohen JD. 2001 An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202. (doi:10.1146/annurev.neuro.24.1.167)
192. Freedman DJ, Riesenhuber M, Poggio T, Miller EK. 2001 Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* **291**, 312–316. (doi:10.1126/science.291.5502.312)
193. Viswanathan P, Nieder A. 2015 Differential impact of behavioral relevance on quantity coding in primate frontal and parietal neurons. *Curr. Biol.* **25**, 1259–1269. (doi:10.1016/j.cub.2015.03.025)
194. Ramirez-Cardenas A, Moskaleva M, Nieder A. 2016 Neuronal representation of numerosity zero in the primate parieto-frontal number network. *Curr. Biol.* **26**, 1285–1294. (doi:10.1016/j.cub.2016.03.052)
195. Wallis JD, Anderson KC, Miller EK. 2001 Single neurons in prefrontal cortex encode abstract rules. *Nature* **411**, 953–956. (doi:10.1038/35082081)

196. Vallentin D, Bongard S, Nieder A. 2012 Numerical rule coding in the prefrontal, premotor, and posterior parietal cortices of macaques. *J. Neurosci.* **32**, 6621–6630. (doi:10.1523/JNEUROSCI.5071-11.2012)
197. Eisele AK, Nieder A. 2013 Representation of abstract quantitative rules applied to spatial and numerical magnitudes in primate prefrontal cortex. *J. Neurosci.* **33**, 7526–7534. (doi:10.1523/JNEUROSCI.5827-12.2013)
198. Petrides M, Pandya DN. 1999 Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *Eur. J. Neurosci.* **11**, 1011–1036. (doi:10.1046/j.1460-9568.1999.00518.x)
199. Petrides M, Pandya DN. 2002 Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *Eur. J. Neurosci.* **16**, 291–310. (doi:10.1046/j.1460-9568.2001.02090.x)
200. Myers RE. 1976 Comparative neurology of vocalization and speech: proof of a dichotomy. *Ann. N. Y. Acad. Sci.* **280**, 745–760. (doi:10.1111/j.1749-6632.1976.tb25537.x)
201. Petrides M, Cadoret G, Mackey S. 2005 Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. *Nature* **435**, 1235–1238. (doi:10.1038/nature03628)
202. Hage SR, Nieder A. 2013 Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nat. Commun.* **4**, 2409. (doi:10.1038/ncomms3409)
203. Hage SR, Nieder A. 2015 Audio-vocal interaction in single neurons of the monkey ventrolateral prefrontal cortex. *J. Neurosci.* **35**, 7030–7040. (doi:10.1523/JNEUROSCI.2371-14.2015)
204. Simonyan K, Jürgens U. 2002 Cortico-cortical projections of the motorcortical larynx area in the rhesus monkey. *Brain Res.* **949**, 23–31. (doi:10.1016/S0006-8993(02)02960-8)
205. Yeterian EH *et al.* 2012 The cortical connectivity of the prefrontal cortex in the monkey brain. *Cortex* **48**, 58–81. (doi:10.1016/j.cortex.2011.03.004)
206. Eliades SJ, Miller CT. 2017 Marmoset vocal communication: behavior and neurobiology. *Dev. Neurobiol.* **77**, 286–299. (doi:10.1002/dneu.22464)
207. Miller CT, Thomas AW, Nummela SU, de la Mothe LA. 2015 Responses of primate frontal cortex neurons during natural vocal communication. *J. Neurophysiol.* **114**, 1158–1171. (doi:10.1152/jn.01003.2014)
208. Roy S, Zhao L, Wang X. 2016 Distinct neural activities in premotor cortex during natural vocal behaviors in a New World primate, the common marmoset (*Callithrix jacchus*). *J. Neurosci.* **36**, 12 168–12 179. (doi:10.1523/JNEUROSCI.1646-16.2016)
209. Miller CT, Dimauro A, Pistorio A, Hendry S, Wang X. 2010 Vocalization induced cFos expression in marmoset cortex. *Front. Integr. Neurosci.* **4**, 128. (doi:10.3389/fnint.2010.00128)
210. Miller CT, Mandel K, Wang X. 2010 The communicative content of the common marmoset phee call during antiphonal calling. *Am. J. Primatol.* **72**, 974–980. (doi:10.1002/ajp.20854)
211. Simoes CS *et al.* 2010 Activation of frontal neocortical areas by vocal production in marmosets. *Front. Integr. Neurosci.* **4**, 1–12. (doi:10.3389/fnint.2010.00123)
212. Tagliabata JP, Russell JL, Schaeffer JA, Hopkins WD. 2008 Communicative signaling activates 'Broca's' homolog in chimpanzees. *Curr. Biol.* **18**, 343–348. (doi:10.1016/j.cub.2008.01.049)
213. Epstein CM *et al.* 1999 Localization and characterization of speech arrest during transcranial magnetic stimulation. *Clin. Neurophysiol.* **110**, 1073–1079. (doi:10.1016/S1388-2457(99)00047-4)
214. Axelson HW *et al.* 2009 Successful localization of the Broca area with short-train pulses instead of 'Penfield' stimulation. *Seizure* **18**, 374–375. (doi:10.1016/j.seizure.2009.01.005)
215. Flinker A, Korzeniewska A, Shestuyk AY, Francaszczuk PJ, Dronkers NF, Knight RT, Crone NE. 2015 Redefining the role of Broca's area in speech. *Proc. Natl Acad. Sci. USA* **112**, 2871–2875. (doi:10.1073/pnas.1414491112)
216. Dronkers NF, Baldo JV. 2010 Broca's area. In *The Cambridge encyclopedia of the language sciences* (ed. PC Hogan), pp. 139–142. Cambridge, UK: Cambridge University Press.
217. Long MA, Katlowitz KA, Svirsky MA, Clary RC, Byun TM, Majaj N, Oya H, Howard III MA, Greenlee DW. 2016 Functional segregation of cortical regions underlying speech timing and articulation. *Neuron* **89**, 1187–1193. (doi:10.1016/j.neuron.2016.01.032)
218. Long MA, Fee MS. 2008 Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* **456**, 189–194. (doi:10.1038/nature07448)
219. McCairn KW *et al.* 2016 A primary role for nucleus accumbens and related limbic network in vocal tics. *Neuron* **89**, 300–307. (doi:10.1016/j.neuron.2015.12.025)
220. Lauterbach EC *et al.* 2013 Toward a more precise, clinically-informed pathophysiology of pathological laughing and crying. *Neurosci. Biobehav. Rev.* **37**, 1893–1916. (doi:10.1016/j.neubiorev.2013.03.002)
221. Norris MR, Drummond SS. 1998 Communicative functions of laughter in aphasia. *J. Neurolinguistics* **11**, 391–402. (doi:10.1016/S0911-6044(98)00026-8)
222. Travers JB, Norgren R. 1983 Afferent projections to the oral motor nuclei in the rat. *J. Comp. Neurol.* **220**, 280–298. (doi:10.1002/cne.902200303)
223. Van Daele DJ, Cassell MD. 2009 Multiple forebrain systems converge on motor neurons innervating the thyroarytenoid muscle. *Neuroscience* **162**, 501–524. (doi:10.1016/j.neuroscience.2009.05.005)
224. Mukudai S, Sugiyama Y, Hisa Y. 2016 Central projections to the nucleus ambiguus. In *Neuroanatomy and neurophysiology of the larynx* (ed. Y Hisa), pp. 104–106. Tokyo, Japan: Springer. (doi:10.1007/978-4-431-55750-0_13)
225. Fay R, Gilbert KA, Lydic R. 1993 Pontomedullary neurons transsynaptically labeled by laryngeal pseudorabies virus. *Neuroreport* **5**, 141–144. (doi:10.1097/00001756-199311180-00011)
226. Barrett RT, Bao X, Miselis RR, Altschuler SM. 1994 Brain stem localization of rodent esophageal premotor neurons revealed by transneuronal passage of pseudorabies virus. *Gastroenterology* **107**, 728–737. (doi:10.1016/0016-5085(94)90120-1)
227. Ekstrand MI, Enquist LW, Pomeranz LE. 2008 The alpha-herpesviruses: molecular pathfinders in nervous system circuits. *Trends Mol. Med.* **14**, 134–140. (doi:10.1016/j.molmed.2007.12.008)
228. Hammerschmidt K, Whelan G, Eichele G, Fischer J. 2015 Mice lacking the cerebral cortex develop normal song: insights into the foundations of vocal learning. *Sci. Rep.* **5**, 8808. (doi:10.1038/srep08808)
229. Banerjee A, Pehlps SM, Long MA. 2019 Singing mice. *Curr. Biol.* **29**, PR190–PR191. (doi:10.1016/j.cub.2018.11.048)
230. Okobi DE, Banerjee A, Matheson AMM, Pehlps SM, Long MA. 2019 Motor cortical control of vocal interaction in neotropical singing mice. *Science* **363**, 983–988. (doi:10.1126/science.aau9480)
231. Nottebohm F, Kelley DB, Paton JA. 1982 Connections of vocal control nuclei in the canary telencephalon. *J. Comp. Neurol.* **207**, 344–357. (doi:10.1002/cne.902070406)
232. Doupe AJ, Perkel DJ, Reiner A, Stern EA. 2005 Birdbrains could teach basal ganglia research a new song. *Trends Neurosci.* **28**, 353–363. (doi:10.1016/j.tins.2005.05.005)
233. Mooney R, Prather J, Roberts T. 2008 Neurophysiology of birdsong learning. In *Memory systems*, vol. 3 (ed. H Eichenbaum), pp. 441–474. Oxford, UK: Elsevier. (doi:10.1016/b978-012370509-9.00116-9)
234. 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**, 118–138. (doi:10.1002/(SICI)1096-9861(19980720)397:1<118::AID-CNE9>3.0.CO;2-3)
235. Benichov JI, Benezra SE, Vallentin D, Globerson E, Long MA, Tchernichovski O. 2016 The forebrain song system mediates predictive call timing in female and male zebra finches. *Curr. Biol.* **26**, 309–318. (doi:10.1016/j.cub.2015.12.037)
236. McCasland JS. 1987 Neuronal control of birdsong production. *J. Neurosci.* **7**, 23–39. (doi:10.1523/JNEUROSCI.07-01-00023.1987)
237. Vu ET, Mazurek ME, Kuo YC. 1994 Identification of a forebrain motor programming network for the learned song of zebra finches. *J. Neurosci.* **14**, 6924–6934. (doi:10.1523/JNEUROSCI.14-11-06924.1994)
238. Yu AC, Margoliash D. 1996 Temporal hierarchical control of singing in birds. *Science* **273**, 1871–1875. (doi:10.1126/science.273.5283.1871)
239. Hahnloser RH, Kozevnikov AA, Fee MS. 2002 An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* **419**, 65–70. (doi:10.1038/nature00974)
240. Ashmore RC, Wild JM, Schmidt MF. 2005 Brainstem and forebrain contributions to the generation of

- learned motor behaviors for song. *J. Neurosci.* **25**, 8543–8554. (doi:10.1523/JNEUROSCI.1668-05.2005)
241. Leonardo A, Fee MS. 2005 Ensemble coding of vocal control in birdsong. *J. Neurosci.* **25**, 652–661. (doi:10.1523/JNEUROSCI.3036-04.2005)
242. 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)
243. Long MA, Jin DZ, Fee MS. 2010 Support for a synaptic chain model of neuronal sequence generation. *Nature* **468**, 394–399. (doi:10.1038/nature09514)
244. Hamaguchi K, Tanaka M, Mooney R. 2016 A distributed recurrent network contributes to temporally precise vocalizations. *Neuron* **91**, 680–693. (doi:10.1016/j.neuron.2016.06.019)
245. Fee MS, Kozhevnikov AA, Hahnloser RH. 2004 Neural mechanisms of vocal sequence generation in the songbird. *Ann. N. Y. Acad. Sci.* **1016**, 153–170. (doi:10.1196/annals.1298.022)
246. Alonso RG, Trevisan MA, Amador A, Goller F, Mindlin GB. 2015 A circular model for song motor control in *Serinus canaria*. *Front. Comput. Neurosci.* **9**, 41. (doi:10.3389/fncom.2015.00041)
247. Kozhevnikov AA, Fee MS. 2007 Singing-related activity of identified HVC neurons in the zebra finch. *J. Neurophysiol.* **97**, 4271–4283. (doi:10.1152/jn.00952.2006)
248. Hamaguchi K, Tschida KA, Yoon I, Donald BR, Mooney R. 2014 Auditory synapses to song premotor neurons are gated off during vocalization in zebra finches. *Elife* **3**, e01833. (doi:10.7554/eLife.01833)
249. Amador A, Perl YS, Mindlin GB, Margoliash D. 2013 Elemental gesture dynamics are encoded by song premotor cortical neurons. *Nature* **495**, 59–64. (doi:10.1038/nature11967)
250. Picardo MA *et al.* 2016 Population-level representation of a temporal sequence underlying song production in the zebra finch. *Neuron* **90**, 866–876. (doi:10.1016/j.neuron.2016.02.016)
251. Chi Z, Margoliash D. 2001 Temporal precision and temporal drift in brain and behavior of zebra finch song. *Neuron* **32**, 899–910. (doi:10.1016/S0896-6273(01)00524-4)
252. Fiete IR, Hahnloser RH, Fee MS, Seung HS. 2004 Temporal sparseness of the premotor drive is important for rapid learning in a neural network model of birdsong. *J. Neurophysiol.* **92**, 2274–2282. (doi:10.1152/jn.01133.2003)
253. Goldin MA, Alonso LM, Alliende JA, Goller F, Mindlin GB. 2013 Temperature induced syllable breaking unveils nonlinearly interacting timescales in birdsong motor pathway. *PLoS ONE* **8**, e67814. (doi:10.1371/journal.pone.0067814)
254. Farabough SM. 1982 The ecological and social significance of duetting. In *Acoustic communication in birds* (eds DE Kroodsma, EH Miller), pp. 85–124. New York, NY: Academic Press.
255. Geissmann T. 2002 Duet-splitting and the evolution of gibbon songs. *Biol. Rev. Camb. Philos. Soc.* **77**, 57–76. (doi:10.1017/S1464793101005826)
256. Hall ML. 2009 A review of vocal duetting in birds. *Adv. Study Behav.* **40**, 67–121. (doi:10.1016/S0065-3454(09)40003-2)
257. Fortune ES, Rodríguez C, Li D, Ball GF, Coleman MJ. 2011 Neural mechanisms for the coordination of duet singing in wrens. *Science* **334**, 666–670. (doi:10.1126/science.1209867)
258. Schmidt MF, Konishi M. 1998 Gating of auditory responses in the vocal control system of awake songbirds. *Nat. Neurosci.* **1**, 513–518. (doi:10.1038/2232)
259. Coleman MJ, Roy A, Wild JM, Mooney R. 2007 Thalamic gating of auditory responses in telencephalic song control nuclei. *J. Neurosci.* **27**, 10 024–10 036. (doi:10.1523/JNEUROSCI.2215-07.2007)
260. Hoffmann S *et al.* 2019 Duets recorded in the wild reveal that interindividually coordinated motor control enables cooperative behavior. *Nat. Commun.* **10**, 2577. (doi:10.1038/s41467-019-10593-3)