

(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

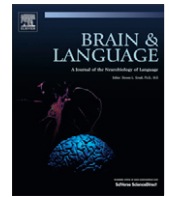
Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>

Contents lists available at [SciVerse ScienceDirect](http://www.sciencedirect.com)

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Review

Mouse vocal communication system: Are ultrasounds learned or innate?

Gustavo Arriaga, Erich D. Jarvis*

Department of Neurobiology, Howard Hughes Medical Institute, Duke University Medical Center, Durham, NC 27710, United States

ARTICLE INFO

Article history:

Accepted 8 October 2012

Keywords:

Ultrasonic vocalization
Vocal learning
Song system
Mouse communication
Motor cortex
Deafening
call convergence
Nucleus ambiguus

ABSTRACT

Mouse ultrasonic vocalizations (USVs) are often used as behavioral readouts of internal states, to measure effects of social and pharmacological manipulations, and for behavioral phenotyping of mouse models for neuropsychiatric and neurodegenerative disorders. However, little is known about the neurobiological mechanisms of rodent USV production. Here we discuss the available data to assess whether male mouse song behavior and the supporting brain circuits resemble those of known vocal non-learning or vocal learning species. Recent neurobiology studies have demonstrated that the mouse USV brain system includes motor cortex and striatal regions, and that the vocal motor cortex sends a direct sparse projection to the brainstem vocal motor nucleus ambiguus, a projection previously thought to be unique to humans among mammals. Recent behavioral studies have reported opposing conclusions on mouse vocal plasticity, including vocal ontogeny changes in USVs over early development that might not be explained by innate maturation processes, evidence for and against a role for auditory feedback in developing and maintaining normal mouse USVs, and evidence for and against limited vocal imitation of song pitch. To reconcile these findings, we suggest that the trait of vocal learning may not be dichotomous but encompass a broad spectrum of behavioral and neural traits we call the continuum hypothesis, and that mice possess some of the traits associated with a capacity for limited vocal learning.

© 2012 Elsevier Inc. All rights reserved.

Contents

1. Introduction	97
2. Vocal communication	97
2.1. Vocalizations and the vocal organ	97
2.2. Types of vocalizations	99
2.2.1. Notes, calls and syllables	99
2.2.2. Songs	100
2.3. Vocal Learning	102
2.3.1. Auditory comprehension learning	102
2.3.2. Vocal usage learning	102
2.3.3. Vocal production learning	102
3. Brain pathways for vocal communication	103
3.1. Programming innate vocalizations	103
3.2. Programming learned vocalizations	103
3.2.1. Vocal motor forebrain pathway in birds and mammals	103
3.2.2. Cortico-basal ganglia-thalamic loops	105
3.3. Identifying vocal communication pathways in mice	105
3.3.1. Mice have a forebrain vocal pathway with some similarities to humans and vocal learning birds	106
4. Innate and learned features of mouse vocalizations	107
4.1. Effects of deafening on innate and learned vocalizations	107
4.2. Evidence for and against a requirement of auditory feedback to maintain specific features of mouse songs	108
4.3. Evidence that mouse songs are innate	109
4.4. Evidence that mouse songs have some learned features	109

* Corresponding author. Address: Box 3209, Department of Neurobiology, Durham, NC 27710, United States. Fax: +1 919 681 0877.

E-mail addresses: arriaga@neuro.duke.edu (G. Arriaga), jarvis@neuro.duke.edu (E.D. Jarvis).

4.4.1.	Ontogeny of mouse USVs.....	109
4.4.2.	Song pitch convergence in mice	111
5.	Conclusions and future directions.....	111
5.1.	Functional connections of the mouse song system.....	112
5.2.	Vocal mimicry.....	112
5.3.	Clearly define vocal learning and categories.....	112
5.4.	Genetically manipulating vocal learning pathways.....	113
	Appendix A. Supplementary material.....	113
	References	113

1. Introduction

Laboratory mice (*Mus musculus*) and rats (*Rattus norvegicus*) participate in a significant amount of communication using ultrasonic vocalizations (USVs) produced at frequencies ranging from 30 to 110 kHz (Constantini & D'Amato, 2006; Portfors, 2007). Traditionally, two types of USVs have been studied in laboratory rodents as measures of internal states: pup isolation calls (Branchi, Santucci, & Alleva, 2001; Brudzynski, Kehoe, & Callahan, 1999; D'Amato, Scalera, Sarli, & Moles, 2005; Elwood & Keeling, 1982; Hahn, Hewitt, Adams, & Trully, 1987; Hofer & Shair, 1992; Ise & Ohta, 2009; Noirod & Pye, 1969; Sales & Smith, 1978; Wöhr, Dalhoff, et al., 2008) and adult USVs in aversive or rewarding conditions (Brudzynski, 2007, 2009; Burgdorf, Wood, Kroes, Moskal, & Panksepp, 2007; Knutson, Burgdorf, & Panksepp, 2002; Wöhr, Houx, Schwarting, & Spruijt, 2008). Reliable elicitation of isolation calls by quantifiable stimuli and a well characterized developmental trajectory have made pup USVs a useful tool for testing the effects of anxiogenic or anxiolytic compounds (Dirks et al., 2002; Fish, Faccidomo, Gupta, & Miczek, 2004; Fish, Sekinda, Ferrari, Dirks, & Miczek, 2000) and for phenotyping mouse models of neuropsychiatric disorders associated with deficits in vocal communication (Scattoni, Crawley, & Ricceri, 2009).

Adult mouse USVs appear to both signal internal emotional states and facilitate social communication during non-aggressive encounters (Gourbal, Barthelemy, Petit, & Gabrion, 2004; Moles, Costantini, Garbugino, Zanettini, & D'Amato, 2007; Portfors, 2007). The most well characterized adult mouse USVs are those produced by males in a mating context. Males of many strains produce long bouts of USVs during courtship of a female and after copulation (Constantini & D'Amato, 2006; Gourbal et al., 2004; Nyby, 1983; Portfors, 2007). Male courtship USVs are sexually selective, and pheromones present in female urine are a strong and sufficient trigger (Guo & Holy, 2007). In two-choice experiments females responded with approach behavior preferentially to adult male USVs over pup isolation calls (Hammerschmidt, Radyushkin, Ehrenreich, & Fischer, 2009; Musolf, Hoffmann, & Penn, 2010), and spent more time with vocalizing males (Pomerantz, Nunez, & Bean, 1983).

Although the general occurrence of male mouse USVs has been known for decades, the spectro-temporal and syntactic features of male courtship USVs were only recently analyzed in depth. Holy and Guo showed that courtship USVs from males contain identifiable syllable types produced in regular temporal patterns that differed between individuals (Holy & Guo, 2005). Moreover, the long strings of syllables they recorded sounded remarkably similar to some bird songs when the pitch of the USVs was shifted to the human audible frequency range and played in real time (Supplementary Audio 1). After observing the complexity of mouse USVs, individual differences, and their similarity to some birdsongs, many researchers wondered what is the neural substrate for USV production, whether mice might share central control mechanisms for vocalization with vocal learning species like songbirds and humans, and whether mouse vocalizations are innate or learned.

The generally accepted list of vocal learning species includes three lineages of birds (songbirds, parrots, hummingbirds) and up to four lineages of mammals (humans, cetaceans [dolphins and whales], bats, elephants, and pinnipeds [sea lions and seals]) (Janik & Slater, 1997; Jarvis, 2004; Schusterman, 2008; Schusterman & Reichmuth, 2008). This vocal learning ability, which includes the ability to modify the spectral and syntactic composition of vocalizations, is a rare trait that serves as a critical substrate for human speech (Doupe & Kuhl, 1999; Jarvis, 2004; Marler, 1970a). It has been well studied in humans and songbirds because songbirds display a capacity for vocal mimicry using a process similar to human speech acquisition (Doupe & Kuhl, 1999; Marler, 1970a) and some species are easy to breed and study in the laboratory. Underlying the vocal learning process in both humans and song learning birds are specialized forebrain circuits so far not found in species that produce only innate vocalizations, despite decades of searching for them (Jarvis, 2004; Jürgens, 2009). Even closely related non-human primate species reportedly lack the behavioral and neural elements classically associated with a capacity for vocal learning (Hammerschmidt, Freudenstein, & Jürgens, 2001; Janik & Slater, 1997; Jürgens, 2009). Like non-human primates, mice have been assumed to be vocal non-learners (Enard et al., 2009; Fischer & Hammerschmidt, 2010; Jarvis, 2004), but this had not been tested. Here we discuss the concepts of innate versus learned vocal communication, give an overview of the neural pathways involved, critically review recent studies that have approached the issue of vocal plasticity mice (Arriaga, Zhou, & Jarvis, 2012; Chabout et al., 2012; Grimsley, Monaghan, & Wenstrup, 2011; Hammerschmidt et al., 2012; Kikusui et al., 2011), address some conflicting views, and propose avenues for reconciliation. The views we propose will be relevant to all studies on innate and learned vocal communication in vertebrates.

2. Vocal communication

2.1. Vocalizations and the vocal organ

Many animals communicate by broadcasting species-typical acoustic signals including insects, frogs, birds, and mammals. However, not all of these sounds are classically defined vocalizations, which are produced by a vocal organ. The vocal organ in birds is the syrinx, and it is the larynx in frogs and most mammals. Dolphins, a marine mammal, are believed to vocalize using specialized nasal sacs in addition to the larynx (Madsen, Jensen, Carder, & Ridgway, 2012). Gross laryngeal anatomy is well conserved among mammals, including between mouse and human, and most of the cartilages and muscles are similarly positioned in both species (Harrison, 1995; Thomas, Stemple, Andreatta, & Andrade, 2009). Premotor signals to the larynx are transmitted via the superior and recurrent laryngeal nerves, and their shared root is the brainstem nucleus ambiguus (Amb). Mouse USVs are most likely generated by the larynx, as revealed in laryngeal nerve transection

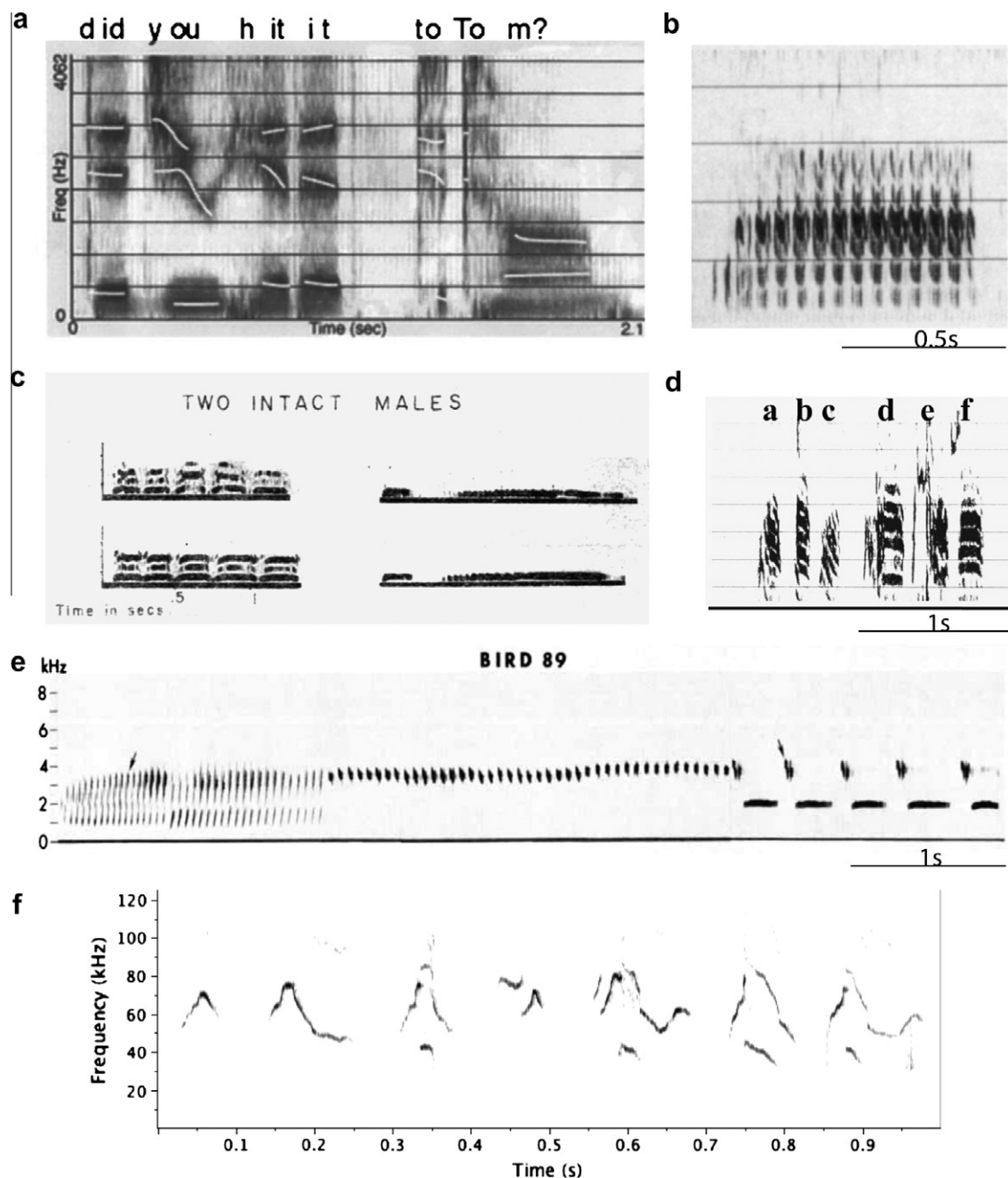


Fig. 1. Sonograms of species-typical vocalizations produced by (a) humans (Doupe & Kuhl, 1999), (b) vervet monkeys (Seyfarth & Cheney, 1986) (the cited study uses the older species name of *Cercopithecus aethiops*), (c) ringdoves (Nottebohm & Nottebohm, 1971), (d) zebra finches (Scharff & Nottebohm, 1991), (e) canaries (Nottebohm et al., 1976), and (f) mice (Arriaga et al., 2012). The mouse song sonogram was generated from Supplementary Audio 2. All images used with permission.

and electrophysiology studies. Bilaterally severing the recurrent laryngeal nerve abolishes pup and adult USVs (Nunez, Pomerantz, Bean, & Youngstrom, 1985; Roberts, 1975). Electrical recordings in anesthetized rats show that a majority of the Amb motoneurons recorded display tonic bursts tightly coupled to and preceding sound production by 46 ms (Yajima, Hayashi, & Yoshii, 1982). Similar results were obtained for extracellular recordings in awake Southern pigtailed macaques (*Macaca nemestrina*), with bursts in Amb associated with variations in vocal output preceding vocalization by 100–200 ms (Yajima & Larson, 1993). Preliminary observations indicate that the explanted mouse larynx is capable of producing sounds displaying the non-linear dynamics characteristic of natural USVs (Berquist, Ho, & Metzner, 2010). However, these

sounds were in the human audible spectrum and it remains unclear if they depend on vibrations of the vocal folds or a whistle mechanism. Other body parts can be used to produce sounds for communication, such as the lips for lip smacking or whistling, or the wings or legs for courtship songs in insects. However, only the larynx and syrinx are known to have the capacity to produce the complex imitated vocalization repertoire observed in humans and song learning birds (Hauser & Konishi, 1999).

Vocalizations can take many forms, the parameters of which are often heavily determined by the production and perceptual mechanisms of the sender and receiver of the acoustic signals. Example spectrograms of a spoken human sentence, songs of a male zebra finch (*Taeniopygia guttata*) and canary (*Serinus canaria*), call of a

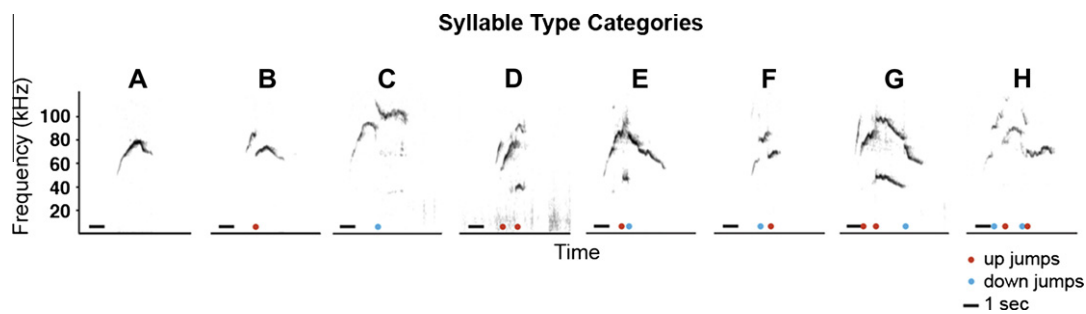


Fig. 2. Examples of syllables categories from courtship vocalizations of adult male BxD mice. Eight major syllable classes (A–H) and several minor (I–K) can be distinguished by the series of notes (boundaries marked by colored dots) and the corresponding sequence of upward or downward instantaneous jumps (>10 kHz) in the dominant pitch (Holy & Guo, 2005). The simplest syllables (Type A) are characterized by a single note with no pitch jumps. Two-note syllables (Types B and C) can be classified by a single 'Down' or 'Up' pitch jump. Common three-note syllables (Types D–F) follow one of the following sequences: 'Down–Down', 'Down–Up', or 'Up–Down'; the fourth possible pitch jump combination 'Up–Up' is rarely observed. Commonly observed four- and five-note syllables (Types G and H) follow 'Down, Down, Up' and 'Up, Down, Up, Down' pitch jump sequences, respectively. Since a jump is defined based on the instantaneous peak frequency, the harmonics present in some notes are not considered for classification purposes. Blue dots mark 'Up' jumps, and red lines mark 'Down' jumps. Scale bar: 20 ms.

ringdove (*Streptopelia risoria*), predator alarm call of a vervet monkey (*Chlorocebus pygerythrus*), and courtship USV of a male mouse reveal the diversity of sounds generated by laryngeal and syringeal mechanisms (Fig. 1). An example recording of a male mouse song shifted into frequencies audible to humans and slowed to highlight the pitch transitions can be heard in Supplementary Audio 2. A sonogram representing 1 s from the same USV bout is shown below (Fig. 1e). These USVs are typically composed of whistle-like syllables that are more similar to the vocalizations of dolphins, some songbirds like canaries, and several primate species, like marmosets. Spectrally, these USVs are unlike the typical vocalizations of zebra finches, parrots, and humans; however, such differences do not preclude them from being used to model mechanisms of vocal production across species.

2.2. Types of vocalizations

Many species produce a diverse repertoire of vocalizations that can include calls, songs, "laughter", and cries. We review some important classifications and describe how they may relate to mouse USVs.

2.2.1. Notes, calls and syllables

Notes are the most basic acoustic unit, and are formed by a single continuous sound with gradual variations in fundamental frequency. One or more notes can be combined to form **Calls** and **Syllables**, which are reproducible single acoustic units separated by periods of silence. Although syllables are structurally similar to calls, we distinguish them from calls by patterns of usage. Calls are typically produced in isolation or in short bursts and may obtain semantic content on their own (Seyfarth, Cheney, & Marler, 1980). Syllables, however, derive their classification from being included in a larger unit representing a longer series of rapidly produced vocalizations of varying types. A reproducible series of syllables with a relatively fixed order is labeled a 'motif'. By clustering units into motifs, an animal with a repertoire of only a few syllables can generate a wide variety of larger communication units. In this classification scheme, syllables can be void of specific meaning themselves, and they would not necessarily serve a communication function if produced in isolation. This distinction is not always entirely clear. For example, the long call of male zebra finches can function alone as a contact call or be incorporated into a motif that is reproduced in song bouts (Zann, 1990). In this case, the same unit could be labeled a call or a syllable depending on the context of production.

Adult mouse USVs feature reproducible sound units that different groups have categorized by their spectral morphology (Fig. 2) (Arriaga et al., 2012; Grimsley et al., 2011; Holy & Guo, 2005; Scattoni, Gandhi, Ricceri, & Crawley, 2008). Most of these units are frequently produced in long sequences containing different types of sound units, and some simple motifs (Holy & Guo, 2005). We will call these recurring units of adult male USVs 'syllables' because they are grouped into non-random series, rarely produced in isolation, and there is no evidence that they serve a communication function individually.

In a study by our laboratory on mouse USV produced in response to female urine, we used a modified version of the Holy and Guo categorization method to identify 8 common and 3–4 rare (<1% of repertoire) syllable types produced by adult males of the B6D2F1/J (BxD) and C57BL6/J (B6) strains (Fig. 2) (Arriaga et al., 2012). The first major morphological distinction between syllable types under this classification scheme is the presence or absence of an instantaneous 'pitch jump' separating notes within a syllable. Thus, the morphologically simplest note type does not contain any pitch jumps (Type A in Fig. 2). For syllables containing pitch jumps, each jump marks the end of one note and the beginning of the next note. Two-note syllables are identified by a single upward or downward pitch-jump (Types B and C in Fig. 2, respectively). Similarly, more complex syllables are identified by the series of upward and downward pitch jumps occurring as the fundamental frequency varies between notes of higher and lower pitch (Types D–H in Fig. 2).

Other researchers have categorized syllables differently, including grouping some of these types and splitting others into subtypes according to the pitch trajectory or note duration (Fischer & Hammerschmidt, 2010; Grimsley et al., 2011; Kikusui et al., 2011; Scattoni et al., 2008). For example, the single note contained in our Type A syllables can have short or long duration, and long notes can be further split based on a downward, upward, chevron-shaped, complex, or flat trajectory (Fig. 3). Our Types B and C syllables have been lumped by others into a two-note super-category (Fischer & Hammerschmidt, 2010; Grimsley et al., 2011; Kikusui et al., 2011; Scattoni et al., 2008) despite having clearly distinct morphologies. Similarly, our syllable Types D through H have been grouped into a 'Frequency Steps' super-category (Scattoni et al., 2008), or into a more than one jump category (Kikusui et al., 2011), and one study grouped all syllable types containing pitch jumps (Types B through H) into a 'whistles with pitch jumps' category (Fischer & Hammerschmidt, 2010). A combination of pitch jump sequences and frequency contours may be necessary to accurately capture the variability of mouse vocal behavior. The number

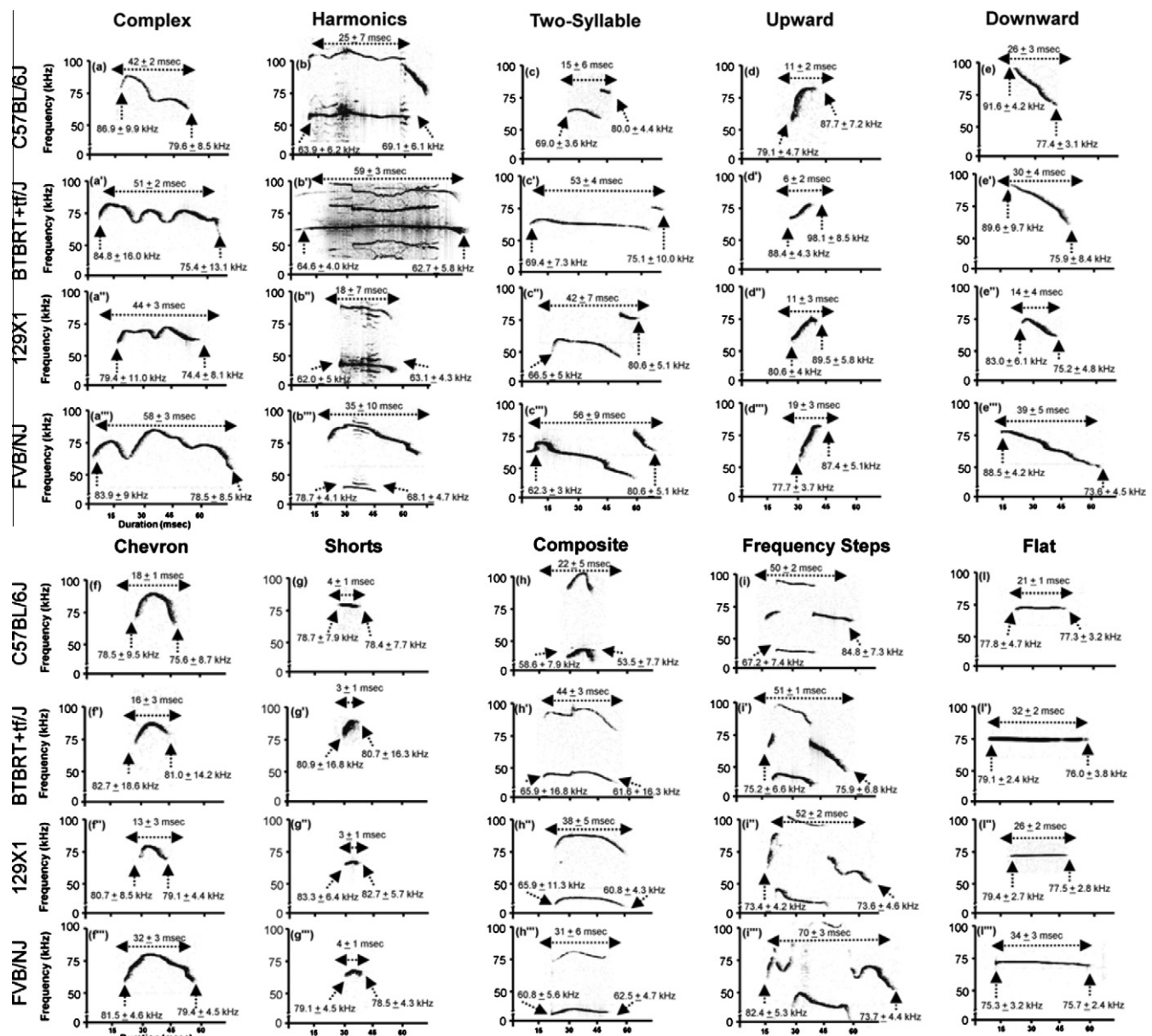


Fig. 3. Examples of syllables from courtship vocalizations of adult male mice as classified by Scattoni et al. (2008). This alternative 10 syllable classification splits syllable Type A from Fig. 2 into six different sub-types (Complex, Upward, Downward, Chevron, Shorts, Flat) and groups syllable Types B and C into a super-category (Two-Syllable).

and sequence of pitch jumps can serve as an initial discriminator, followed by a refined categorization based on frequency contours and duration, as described for syllable Type A. However, arbitrarily grouping syllables with measurably different numbers and sequences of pitch jumps obscures real variability in vocal behavior and may complicate subsequent analysis of heterogeneous syllable categories. The issue of classification is an active area of investigation that has not yet reached consensus. A conference was held at the Institut Pasteur in Paris, France in April 2012 to address problems of syllable/note classification (http://www.ura2182.cnrs-bellevue.fr/workshop_usv/). Until a robust classification scheme is developed, negative results must be interpreted with caution due to the possibility of improper classification (grouping very different syllables) masking real effects.

2.2.2. Songs

A song is set of vocalizations, often elaborate, delivered periodically and sometimes with a rhythm. Songs may be produced spon-

taneously or in response to an external stimulus such as the presence of a conspecific. Songs typically contain multiple syllable types, or categories of reproducible vocalizations distinct from other vocalizations comprising the song. To distinguish a series of syllables in a song from a succession of calls we will apply the *sensu strictissimo* definition used previously (Holy & Guo, 2005) and borrowed from Broughton (Broughton, 1963):

'a sound of animal origin which is not both accidental and meaningless'

containing,

'a series of notes, generally of more than one type, uttered in succession and so related as to form a recognizable sequence or pattern in time,'

produced in,

'a complete succession of periods or phrases'

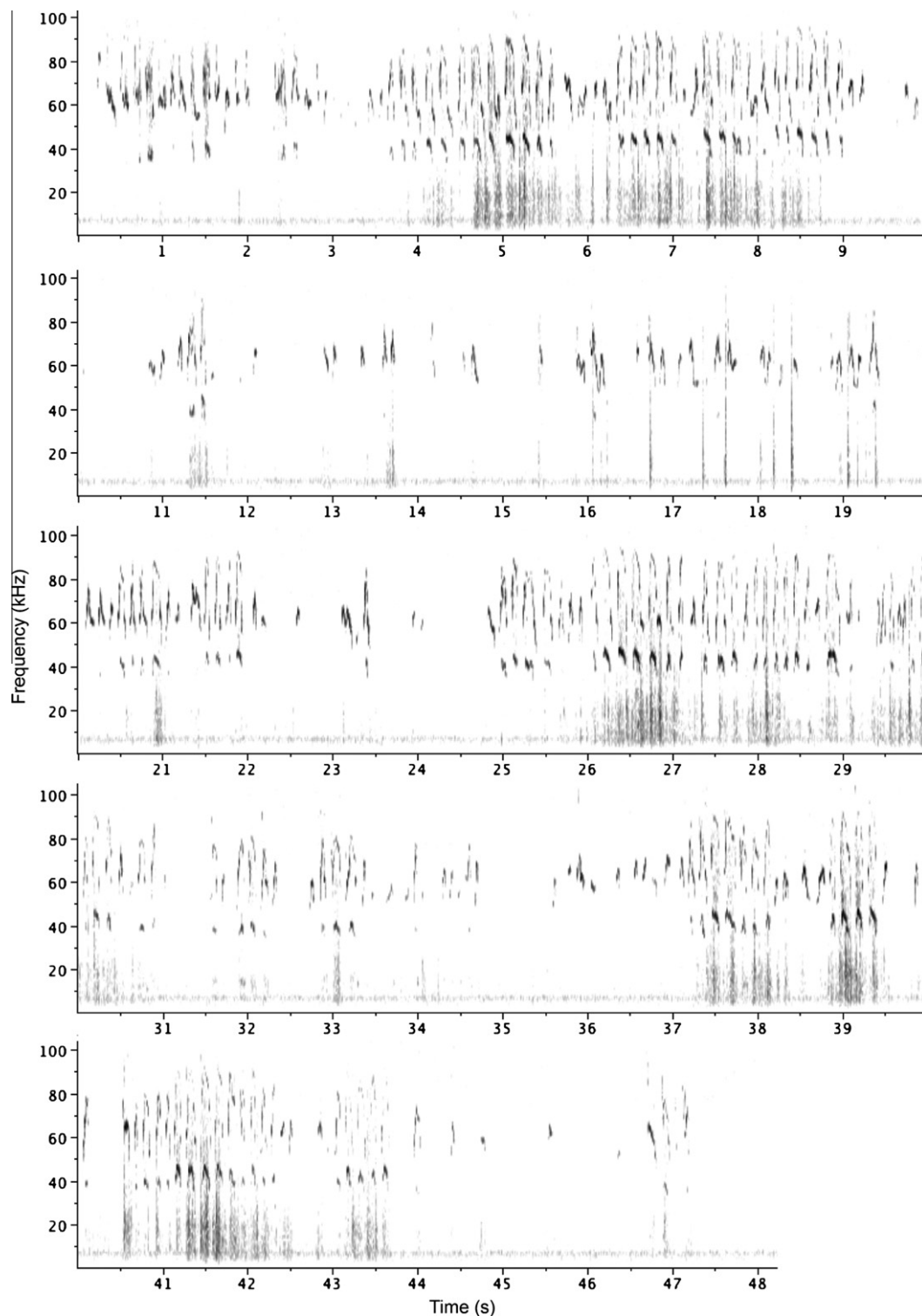


Fig. 4. Song bout of an adult BxD male lasting 47 s and containing 264 syllables.

Holy and Guo's analysis of the spectro-temporal features of male courtship USVs demonstrated that these vocalizations satisfy all conditions required for classification as song (Holy & Guo, 2005). Visually, the song-like quality of male mouse courtship USVs can be appreciated in spectrograms of longer sequences (Fig. 4). Acoustically, when the pitch of courtship USVs is shifted to the audible spectrum they sound very similar to some birdsongs in both temporal and melodic structure (Supplementary Audio 1).

The behavioral responses of conspecifics also provide clues that male mouse songs are distinct from calls. Males often do not sing in isolation or to other males, but are triggered to sing by the presence of a female or female urine (Guo & Holy, 2007; Musolf et al., 2010; Nyby, 1983). Female mice can distinguish male songs from pup isolation calls (Hammerschmidt et al., 2009). Given the choice, females selectively approach the source of the songs instead of the source of the isolation calls. Preference for male songs is striking

given that pup calls are considered a very strong and reliable stimulus, and the frequency ranges of the two signals overlap significantly. Moreover, a separate choice experiment reported a slight tendency for females to prefer the songs of non-kin males (Musolf et al., 2010), further suggesting that individual songs are distinguishable and could serve an important social and reproductive function.

Despite the structural and behavioral evidence that they meet the *sensu strictissimo* definition of song, some researchers still prefer to refer to them as long sequences of calls. We leave it to the reader to decide based on the evidence, but for the purpose of this review we will simply refer to these vocalizations as “mouse songs”. This designation does not necessarily imply learning. For example, the songs of some suboscine passerine birds are innate, although they share structural and behavioral characteristics with the learned songs of oscine songbirds (Kroodsma & Konishi, 1991). Likewise some calls of oscine songbirds are learned (Simpson & Vicario, 1990). Given the various learning strategies described, the multiple functions of vocal signals, and the existence of innate songs, our working definition of song deliberately excludes ontogeny and focuses primarily on phenotype.

2.3. Vocal Learning

Many types of learning are possible within the framework of vocal communication systems. Thus, it is important not only to determine what learning capabilities are present in the mouse vocal system, but also to distinguish which types of learning are most relevant to studies of speech learning in humans. Three types of learning are related to vocal communication systems: auditory comprehension learning, vocal usage learning, and vocal production learning (Egnor & Hauser, 2004; Janik & Slater, 1997, 2000; Jarvis, 2004; Schusterman, 2008).

2.3.1. Auditory comprehension learning

Auditory comprehension learning is an auditory learning strategy characterized by the ability to associate a particular sound with an appropriate behavioral response or objects in the environment (Janik & Slater, 1997; Jarvis, 2004). Comprehension learning capabilities are broadly distributed among vertebrates. For example, dogs (*Canis lupus familiaris*) can be trained to respond to the human word ‘sit’; however, the vocal part of this training process is restricted to the act of correctly identifying the word through auditory learning. Learning in this case does not extend to the vocal production (i.e. vocal motor) component. Dogs do not learn to produce the word ‘sit’ by adaptively modifying motor commands to achieve the required sequence of laryngeal and respiratory patterns. However, some motor behaviors can be associated with a learned auditory cue. For example, the dog’s typical learned response to the verbal command ‘sit’ is the motor act of sitting on the hindquarters.

2.3.2. Vocal usage learning

Vocal usage learning is characterized by the ability to learn when and where, but not how, to produce vocalizations in a specific social or environmental context. Usage learning does not require acoustic vocal imitation. A well-studied example of usage learning is the alarm call repertoire of vervet monkeys produced in response to specific predator threats. An eagle (*Polemaetus bellicosus*) in the sky, a leopard (*Panthera pardus*) in the trees, and a python (*Python sebae*) on the ground will elicit different species-typical calls, and a young vervet monkey must learn through experience when it is appropriate to produce each call (Seyfarth et al., 1980). However, the spectral content of alarm calls is thought to be innately determined (Seyfarth & Cheney, 1986). Learning is restricted to the context or ‘when’ of production, but the ‘how’ is

inflexible. Usage learning and comprehension learning are often intimately linked. For example, it is critical that a young vervet monkey learn not only which call to produce in response to each predator, but also learn the appropriate predator-specific defensive behavior to produce upon hearing each call. The leopard-specific call triggers retreat into the trees, and the eagle-specific call causes listeners to hide in the dense bush (Seyfarth et al., 1980). The learned association of auditory cues with effective predator defense strategies is similar to the training of a dog’s behavior to verbal commands.

2.3.3. Vocal production learning

In contrast, vocal **production learning** is the ability to generate experience-dependent modifications of acoustic signals, and is considered the most relevant to the study of human speech (Janik & Slater, 1997; Jarvis, 2004). Strictly defined, production learning excludes changes in the amplitude and duration of vocalizations because they rely on control of respiratory patterns rather than control of the musculature of the vocal organ (Janik & Slater, 1997). In this context, the most dramatic and well-studied examples of vocal production learning are song learning in birds and speech learning in humans. Birdsong and speech share many features: auditory acquisition of learning templates, dependence on auditory feedback for learning and maintenance of learned vocalizations, temporally restrictive critical periods for learning, and specialized forebrain networks for vocal control (Doupe & Kuhl, 1999; Jarvis, 2004; Marler, 1970a). Because of these important similarities, songbirds have become the dominant neuro-ethological animal models for vocal learning studies.

One consequence of the intense focus on the songbird model is a situation where the meaning of the term ‘vocal learning’ has been restricted to refer exclusively to learning vocalizations *de novo* with reference to an externally acquired model, as occurs for bird-song and speech learning. Certainly, this type of vocal mimicry is the most relevant to study for modeling and understanding the process of human speech acquisition. However, we believe this represents an overly restrictive definition of vocal learning that ignores many other factors and strategies that can be used to adaptively modify the spectral content of vocalizations. For example, white-crowned sparrows (*Zonotrichia leucophrys*) that normally learn songs from a tutor will still produce novel songs despite having been raised in social isolation (Konishi, 1985). This process of generating an isolate song without previous instruction, or adding novel parts to a tutored song has been called improvisation (Janik & Slater, 1997; Konishi, 1964; Kroodsma, Houlihan, Fallon, & Wells, 1997; Marler, 1997).

Improvisation is one of the simplest ways that animals may change their vocalizations without explicit need for a tutored model. Using improvisation, an animal could rely on internal preference or the response of conspecifics to guide the learning process. Therefore, it is important to evaluate the relative roles of improvisation and imitation in any vocal learning species. In some experiments, gray catbirds (*Dumetella carolinensis*), which are a type of songbird, often failed to copy song models and routinely generated normal songs with novel elements not present in the template (Kroodsma et al., 1997). More strikingly, when the abnormal song of a socially isolated adult zebra finch was used as the tutor template, the tutored juveniles modified the song to more closely match a more typical finch song (Fehér, Wang, Saar, Mitra, & Tchernichovski, 2009). Accumulation of corrective improvisations over five generations was sufficient to transform the isolate song to a normal-sounding zebra finch song. Preferential learning by improvisation was performed even though all the birds should be perfectly capable of mechanically reproducing the isolate songs heard.

In some vocal learning species determination of what is worth learning is shaped by individuals other than the one learning. For

example, non-singing female cowbirds (*Molothrus ater*) exert a strong sexual selection on male song development by selectively reinforcing song variants with their wing displays (West & King, 1988). The effect of female selection is so strong that both tutored and untutored males develop different songs depending on the preferences of co-housed females from different sub-species (King, 1983). Experiments with Pacific walrus (*Odobenus rosmarus divergens*) demonstrated that the preferences of human trainers could also reinforce novel vocal behavior (Schusterman & Reichmuth, 2008). Using a contingency learning paradigm, walrus were rewarded with fish when a vocalization was judged by the human trainer to be significantly different than the preceding vocalization. Under stimulus control, sounds in the existing repertoire were elaborated with pitch and contour changes, and several novel vocalizations emerged that had not been heard before.

It is clear that mimicry is not the only viable strategy for vocal production learning. Indeed, different strategies could have been necessary for different species to transition from generating exclusively innate sounds to generating novel sounds. For these reasons, we subscribe to the view proposed by Konishi (1985) by accepting as production learning the development of any vocalizations that depend on auditory feedback for the development or maintenance of spectral content. Under this definition it is the reliance on auditory feedback to control the vocal organ and guide the trajectory of sound development that is most important. Of secondary importance is whether the trajectory results in convergence toward or divergence from an external model, the emergence of internal preferences, or acquisition of a social or food reward.

3. Brain pathways for vocal communication

It has been proposed that two different, but converging pathways are involved in the production of learned and innate vocalizations (Jarvis, 2004; Jürgens, 2009; Simonyan & Horwitz, 2011; Wild, 1994, 1997). According to this division of labor, innate calls are programmed by a phylogenetically older brainstem pathway, and the forebrain influences the context (i.e. usage) of calling but not acoustic structure. In contrast, control of the spectral content of learned calls would be given over to a phylogenetically more recent vocal pathway driven directly by forebrain premotor structures—the so-called Kuypers/Jürgens hypothesis (Fitch, Huber, & Bugnyar, 2010).

3.1. Programming innate vocalizations

The brain pathway for programming acoustically innate vocalizations includes midbrain premotor structures and medullary motoneuron pools for motor control of phonation and respiration. This pathway has been found in all vocalizing avian and mammalian species studied to date, and homologous pathways can even be found in vocalizing fish (Bass & McKibben, 2003; Jürgens, 2009; Kittelberger, Land, & Bass, 2006; Wild, 1997). In both vocal learning and non-learning birds, this innate vocal circuit comprises the dorsomedial nucleus (DM) in the midbrain that projects to multiple medullary nuclei including the parabrachial region (PBr), the expiratory premotor nucleus retroambigualis (RAm), and the tracheosyringeal part of the hypoglossal nucleus (XII_{ts}) that innervates the syrinx (Fig. 5) (Wild, 1997). The analogous vocal circuit in mammalian brains comprises the caudal periaqueductal gray (PAG) in the midbrain which projects to brainstem respiratory premotor nuclei including RAm for control of respiration, and cranial nerve nuclei including Amb that directly innervates the larynx (Fig. 5) (Ennis, Xu, & Rizvi, 1997; Jürgens, 1998, 2002a, 2009; Mantyh, 1983).

These pathways have been identified in two well-studied non-human primate models of vocalizations, the squirrel monkey (*Saimiri sciureus*) and rhesus macaque (*Macaca mulatta*). Decades of work by Uwe Jürgens and colleagues using anatomical tracing (Dujardin & Jürgens, 2005; Hannig & Jürgens, 2005; Jürgens, 1982, 1983, 1984; Jürgens & Alipour, 2002; Müller-Preuss & Jürgens, 1976; Müller-Preuss, Newman, & Jürgens, 1980; Simonyan & Jürgens, 2002, 2003, 2004, 2005; Thoms & Jürgens, 1987), brain imaging (Jürgens, Ehrenreich, & de Lanerolle, 2002), electrophysiology (Düsterhöft, Häusler, & Jürgens, 2003; Hage & Jürgens, 2006a, 2006b; Jürgens, 2002a; Lütke, Häusler, & Jürgens, 2000), electrical (Jürgens & Ploog, 1970) and chemical (Lu & Jürgens, 1993) brain activation, lesions (Jürgens & Pratt, 1979; Jürgens, Kirzinger, & von Cramon, 1982; Kirzinger, 1985; Kirzinger & Jürgens, 1982, 1985), and reversible inactivations (Jürgens & Ehrenreich, 2007; Siebert & Jürgens, 2003) has produced a detailed description of the pathways involved in controlling innate primate vocalizations (Jürgens, 2009). The general conclusions drawn from this body of work are as follows: (1) limbic regions regulating arousal and the drive to vocalize including the amygdala and anterior cingulate cortex converge on the PAG; (2) the PAG serves a gating function to activate motor programs for specific calls associated with different arousal states; and (3) the spectral structure of calls is primarily determined at the level of medullary premotor circuits that coordinate the activity of phonatory motoneuron pools in various cranial nerve nuclei (Jürgens, 1998, 2002b, 2009; Jürgens & Alipour, 2002). Lesions of the anterior cingulate cortex or amygdala do not eliminate the ability to produce the innate vocalizations, but reduce the motivation to vocalize and to do so in the appropriate context. However, lesioning or blocking the PAG or Amb eliminates production of innate vocalizations (Floody & DeBold, 2004; Jürgens & Ehrenreich, 2007; Jürgens & Pratt, 1979; Kirzinger & Jürgens, 1985; Siebert & Jürgens, 2003). These findings suggest that what is truly indispensable for vocalization is the PAG and downstream circuits of the brainstem.

3.2. Programming learned vocalizations

Although the gross anatomy of avian and mammalian forebrains is remarkably different (nucleated in birds and layered in mammals) there are some general principles shared among all vocal learning systems (Jarvis, 2004; Jarvis et al., 2005). In addition to the limbic-midbrain-hindbrain pathway for innate vocal production, vocal-learning avian species and humans have evolved cortico-bulbar pathways and cortico-basal ganglia-thalamic loops for generating and learning novel vocalizations, respectively.

3.2.1. Vocal motor forebrain pathway in birds and mammals

Learned song in birds is controlled by a hierarchically organized pre-motor control pathway contained within two nuclei of the caudal telencephalon that sends direct and indirect output to the vocal motoneurons of the brainstem located in XII_{ts} (Wild, 1997). In songbirds, this premotor pathway begins with the nucleus HVC (used as the proper name), from which a specific subset of projection neurons innervates the robust nucleus of the arcopallium (RA) (Foster & Bottjer, 1998; Nottebohm, Stokes, & Leonard, 1976). These RA-projecting neurons appear to encode the timing of song via a sparse code that coordinates the bursting activity of neuron ensembles in RA (Fee, Kozhevnikov, & Hahnloser, 2004; Hahnloser, Kozhevnikov, & Fee, 2002; Leonardo & Fee, 2005; Yu & Margoliash, 1996). RA projects to various midbrain and brainstem nuclei including DM of the innate call generating pathway, the respiratory premotor nucleus RAm, Amb, and the motoneurons of XII_{ts} that control the vocal organ (Nottebohm et al., 1976; Wild, 1993). These direct downstream targets of RA make it well positioned to allow forebrain control over the activity of respiratory,

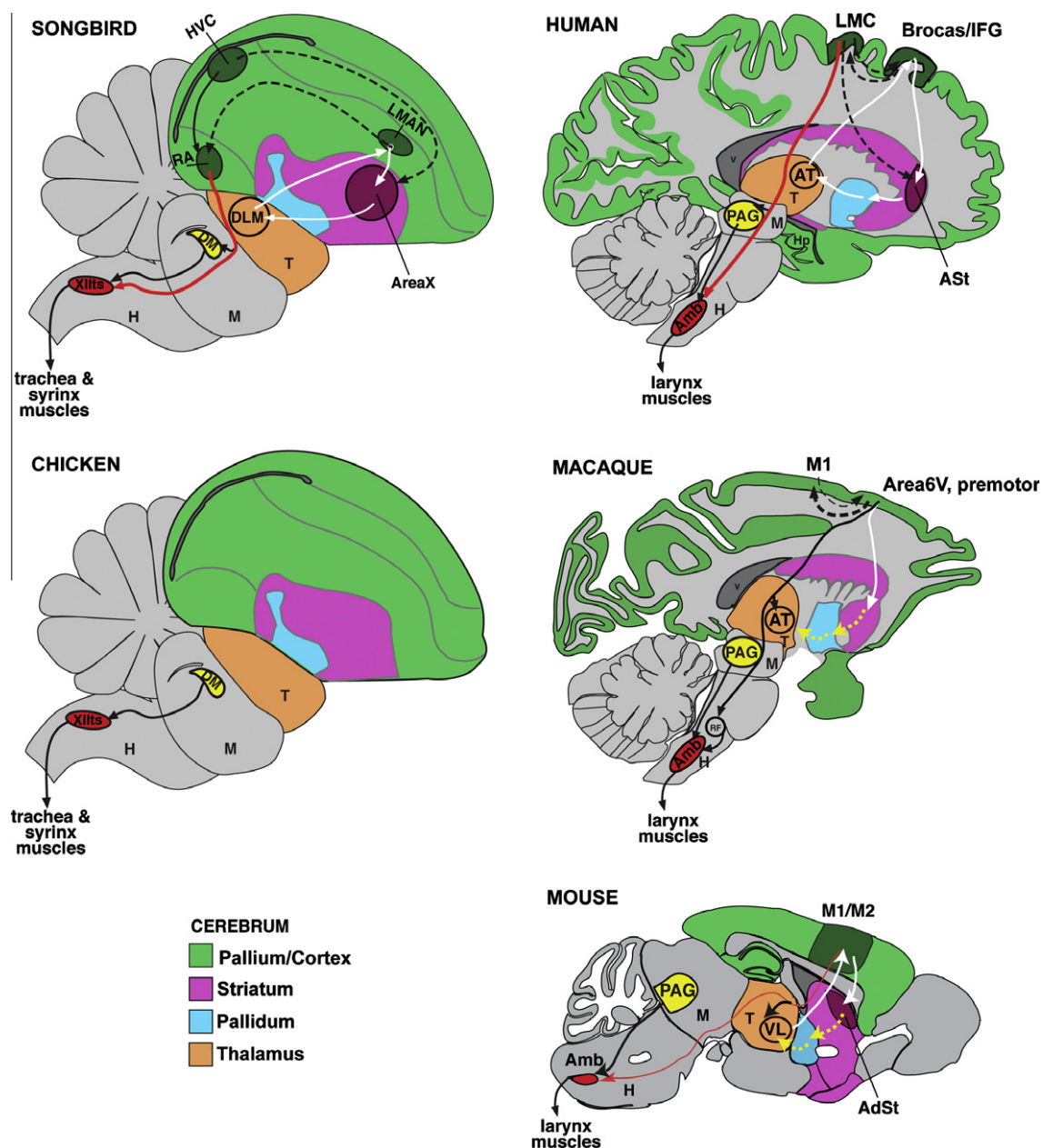


Fig. 5. Summary diagrams of brain systems for vocalization in mice, and classical vocal learning and vocal non-learning species for comparison. All vocalizing species including monkeys and chickens have a midbrain/brainstem vocal motor pathway. Monkeys have a premotor cortex region in Area 6V that makes an indirect projection to vocal motor neurons, but is not required for vocalizing. The vocal learning species (Human and Songbird) possess additional forebrain premotor circuits that are critical for producing and learning vocalizations, including cortico-striatal-thalamic loops (dotted lines) and a direct primary motor cortical projection to vocal motoneurons in the brainstem (red arrows: RA to Xlts in songbirds; Laryngeal motor cortex (LMC) to Amb in humans) (Jarvis, 2004; Jürgens, 2009; Kuypers, 1958c; Simonyan & Horwitz, 2011). Mice have a similar direct cortico-bulbar projection (Arriaga et al., 2012). Red arrows, the direct forebrain projection to vocal motor neurons in the brainstem (RA to Xlts in song learning birds; Laryngeal motor cortex [LMC] to Amb in human and mouse) (Jarvis, 2004; Jürgens, 2002b; Kuypers, 1958b; Wild, 1997). White lines, anterior forebrain premotor circuits, including cortico-striatal-thalamic loops. Dashed lines, connections between the anterior forebrain and posterior vocal motor circuits. Yellow lines, proposed connections for cortico-striatal-thalamic loop that need to be tested. Auditory input is not shown. All diagrams show the sagittal view. Figure used with permission from Arriaga et al. (2012).

laryngeal, and syringeal muscle groups during vocalization. A similarly connected hierarchical vocal premotor pathway was found in the forebrain of parrots (Durand, Heaton, Amateau, & Brauth, 1997; Jarvis, 2004; Jarvis & Mello, 2000; Paton, Manogue, & Nottebohm, 1981; Striedter, 1994) and hummingbirds (Gahr, 2000; Jarvis et al., 2000). In parrots the pathway involves analogous projections from the central nucleus of the lateral nidopallium (NCL) to the central nucleus of the anterior arcopallium (AAc), which projects in turn to midbrain and brainstem vocal nuclei (Durand et al., 1997; Striedter, 1994). In hummingbirds, a nucleus similar in location and cytoarchitecture to songbird HVC was found called the

vocal nucleus of the lateral nidopallium (VLN) or HB-HVC (Gahr, 2000; Jarvis et al., 2000). HB-HVC sends descending projections to the vocal nucleus of the arcopallium (VA) also called HB-RA, which resembles songbird RA and innervates Xlts (Gahr, 2000; Jarvis et al., 2000). In contrast, no such forebrain nuclei or direct projections from the arcopallium have been found in vocal non-learning birds, such as pigeons and chickens (Wada, Sakaguchi, Jarvis, & Hagiwara, 2004; Wild, 1997).

Among mammals, projections from primary motor cortex to phonatory brainstem nuclei have only been found in primates. In a comparative study of projections from the motor cortical tongue

area to the hypoglossal nucleus (XII) that innervates the tongue muscles, it was observed that the density of the projection varies between primate species (Jürgens & Alipour, 2002). Rhesus macaques have a relatively denser projection than squirrel monkeys, and saddle-back tamarins (*Saguinus fuscicollis*) have putative fibers of passage but no terminals in XII. In chimpanzees (*Pan troglodytes*) (Kuypers, 1958a) and humans (Kuypers, 1958b) projections to XII are dense. By contrast, no motor cortical projection to XII was observed in tree shrews (*Tupaia belangeri*) (Jürgens & Alipour, 2002), cats (*Felis catus*) (Kuypers, 1958c), or rats (Travers & Norgren, 1983). A direct motor cortical vocal pathway, consisting of a direct cortical projection to the laryngeal motoneurons in Amb had only been found in humans among mammals (Iwatsubo, Kuzuhara, Kanemitsu, Shimada, & Toyokura, 1990; Kuypers, 1958d, 1958b; Simonyan & Jürgens, 2003). This distribution of cortico-bulbar projections to XII and Amb has been interpreted as a progressive increase in cortical innervation in phylogenetically newer primate species leading to improved vocal abilities (Jürgens & Alipour, 2002). This interpretation reflects the general assumption that presence of direct cortical input to phonatory motor nuclei determines the level of vocal abilities. Indeed, the presence of a direct motor cortical/pallial vocal pathway in vocal learning birds and humans has been proposed by many researchers as one of the key neural transformations in the evolution of spoken-language and learned song (Deacon, 2007; Fischer & Hammerschmidt, 2010; Fitch et al., 2010; Jarvis, 2004; Jürgens et al., 1982; Kirzinger & Jürgens, 1982; Okanoya, 2004; Simonyan & Horwitz, 2011; Simonyan & Jürgens, 2003).

3.2.2. Cortico-basal ganglia-thalamic loops

In songbirds, there is a cortico-basal ganglia-thalamic loop dedicated to vocalization called the anterior forebrain pathway (AFP). Premotor input to the AFP comes from a distinct subset of HVC projection neurons that innervate a region of the anteromedial striatum specialized for vocal learning called Area X (Foster & Bottjer, 1998; Nottebohm et al., 1976). Area X sends a GABAergic projection to the dorsolateral anterior thalamic nucleus (DLM), which projects in turn to the lateral magnocellular nucleus of the anterior nidopallium (LMAN) (Bottjer, Halsema, Brown, & Miesner, 1989; Okuhata & Saito, 1987; Person, Gale, Farries, & Perkel, 2008). LMAN then projects back to Area X forming a cortico-striatal-thalamic loop specialized for vocalization (Okuhata & Saito, 1987). A similar second medial AFP loop has been proposed, which comprises a projection from medial Area X to the dorsomedial nucleus of the posterior thalamus (DMP), then to the medial magnocellular nucleus of the anterior nidopallium (MMAN) (Kubikova, Turner, & Jarvis, 2007). LMAN and MMAN are the output nuclei of the AFP, projecting to RA (Nottebohm, Paton, & Kelley, 1982) and HVC (Foster & Bottjer, 1998), respectively. These outputs allow the AFP to modulate the ongoing activity of the direct HVC-RA premotor circuit (Kao, Doupe, & Brainard, 2005). Lesions and chemical inactivation of MAN nuclei and Area X revealed that the AFP is not required for singing, but is critical for generating the acoustic variability necessary for vocal exploration in normal song learning (Bottjer, Miesner, & Arnold, 1984; Foster & Bottjer, 2001; Nottebohm et al., 1976; Olveczky, Andelman, & Fee, 2005; Scharff & Nottebohm, 1991), social context-dependent modulation of song (Kao et al., 2005; Kao & Brainard, 2006), experimentally-induced song deterioration (Brainard & Doupe, 2000; Williams & Mehta, 1999), and modulation of activity and singing-driven gene regulation of HVC and RA (Kubikova et al., 2007; Olveczky et al., 2005).

A similar recurrent cortico-basal ganglia-thalamic pathway was found in the forebrain of parrots, except that NLC (HVC analog) does not project to the basal ganglia song nucleus (MMSt) (Durand et al., 1997; Jarvis & Mello, 2000); instead the ventral portion of the

RA analog (AACv) projects to the LMAN analog (Durand et al., 1997). In hummingbirds, analogous basal ganglia and cortical regions have been found to be active during song production (Jarvis et al., 2000). The connectivity between these AFP-like regions has not been established in hummingbirds except for the projection from the proposed LMAN analog to the RA analog, which is similar to the oscine and parrot song systems (Gahr, 2000). Thus, the general design of several similarly arranged discrete forebrain nuclei forming a direct forebrain premotor pathway modulated by a recurrent basal ganglia loop seems to be a universal feature among independently derived lineages of avian vocal learners (Jarvis, 2004).

In humans, cortical, basal ganglia, and thalamic vocalization-related brain regions have typically been identified with functional neuroimaging techniques during speech production or brain lesion case studies (Jürgens, 2002b; Ludlow, 2005). In contrast, vocalization-specific neural activity in vocal non-learning mammalian species had been demonstrated only in limbic, midbrain and brainstem circuits (Hage & Jürgens, 2006a, 2006b; Jürgens, 2002a, 2009; Wild, 1997). In non-human primates, electrical micro-stimulation of a specific premotor cortical region in area 6 produced movement of the vocal folds (Hast, Fischer, Wetzel, & Thompson, 1974). Tract tracing studies of this putative laryngeal premotor region revealed extensive subcortical projections to the basal ganglia, thalamus, pons and medulla (Simonyan & Jürgens, 2003). However, chemically inactivating these connecting structures does not abolish vocal fold movements elicited by motor cortical stimulation (Jürgens & Ehrenreich, 2007). Moreover, lesions to prefrontal and primary motor cortex (Aitken, 1981; Kirzinger & Jürgens, 1982; Sutton, Larson, & Lindeman, 1974) or globus pallidus (MacLean, 1978) do not produce changes in the structure of vocalizations in monkeys, but abolish learned volitional vocalizations in humans (Jürgens, 2002b). Therefore, it is questionable that these structures play a role in the programming of monkey vocalization, but they may serve other laryngeal functions in non-vocal behaviors like swallowing.

3.3. Identifying vocal communication pathways in mice

We were unaware of any previous studies attempting to define vocal premotor forebrain circuits in mice, so we addressed this issue first (Arriaga et al., 2012). We looked for motor-driven singing-regulated expression of activity-dependent immediate early genes using a similar experimental design as previous studies that identified seven similar forebrain song nuclei among the three lineages of song learning birds (Jarvis & Mello, 2000; Jarvis & Nottebohm, 1997; Jarvis et al., 2000). We found that relative to the non-singing treatment groups, male mice that produced USVs expressed higher levels of mRNA for two immediate early genes (IEGs), *egr-1* and *arc*, bilaterally in restricted regions of the primary motor (M1) and premotor (M2) cortices, adjacent anterior cingulate cortex (Cg), and subjacent anterodorsal striatum (ADSt) (Fig. 6a and b). Importantly, similar amounts of *egr-1* and *arc* expression were observed for mice singing with intact hearing and mice singing after deafening. Moreover, playback of mouse songs in the absence of active singing did not induce similar IEG expression in these forebrain regions. These results indicate that the greater levels of mRNA expression in these regions were not caused by auditory processing during singing. Instead, the results show that singing-induced expression of activity-dependent IEGs in motor cortical, limbic, and striatal regions of the mouse brain is motor-driven. This pattern of vocal motor specific activity is similar to what is observed in the songbird song system during singing (Jarvis & Nottebohm, 1997), but had not been previously shown in the forebrain of a non-human mammal.

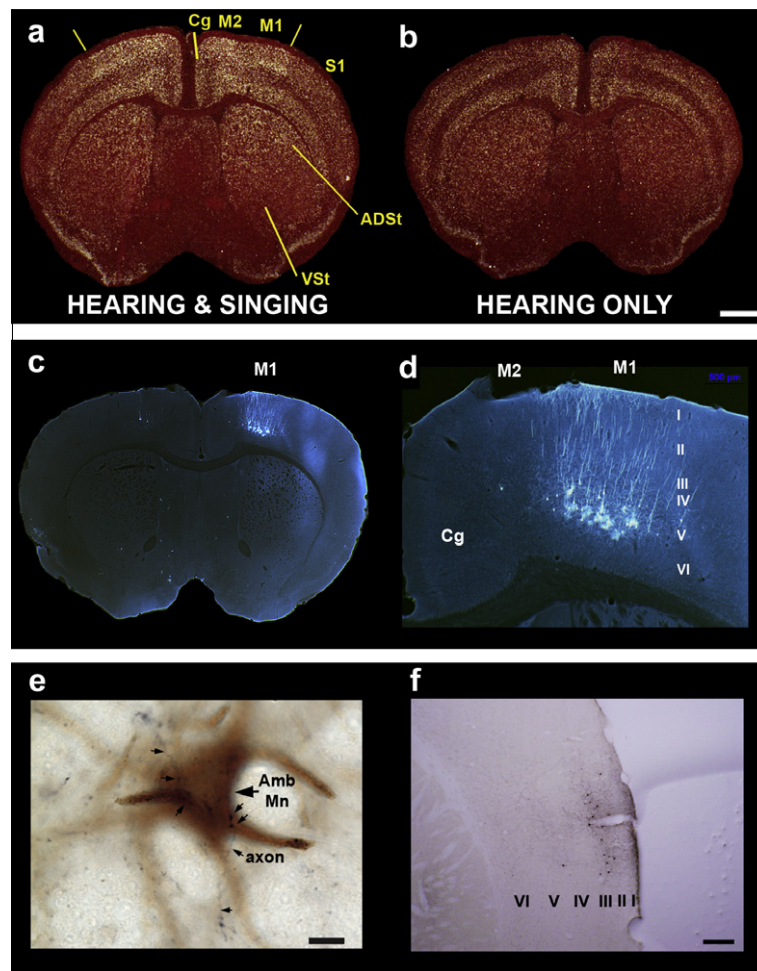


Fig. 6. Molecular mapping and some connectivity of mouse song system forebrain areas. (a and b) Dark-field images of cresyl violet stained (red) coronal brain sections at the level of motor cortex, approximately 0.2 mm rostral to Bregma, showing singing-induced *egr1* expression (white) in the forebrain of a male mouse (a) relative to a non-vocalizing control that moved around the cage in a similar amount (b). (c) Pyramidal neurons expressing enhanced green fluorescent protein in cortical layer V of the singing activated region of M1 following injection of pseudorabies virus (PRV-Bartha) into the cricothyroid and lateral cricoarytenoid laryngeal muscles. Labeled cells were not observed in the adjacent M2 and cingulate cortex (Cg) or subjacent anterodorsal striatum (adSt). (d) Higher magnification of the labeled cells in (c). (e) Fine caliber M1 axons (black arrows) contact CTb-labeled laryngeal Amb motor neurons (MN; brown) from an injection in the M1 singing activated region of cortex. (f) Backfilled layer III cells in secondary auditory cortex (A2) from the same animal. Scale bar = 1 mm for a–c; 0.5 mm for d and f; 10 μ m for e. Figures used with permission from Arriaga et al. (2012). Figure panels c and d are from an additional animal not shown in that paper.

Two recent studies claimed to find cortical activation during vocalization in marmosets (*Callithrix jacchus*) by examining brain expression patterns of *egr-1* (Simões et al., 2010) and *c-fos* (Miller, DiMauro, Pistorio, Hendry, & Wang, 2010). In the first study, expression levels of *egr-1* were measured in prefrontal cortex of two groups of animals that heard playbacks of conspecific calls and either vocalized or remained silent (Simões et al., 2010). Higher numbers of *egr-1* immunopositive cells were observed in ventral and dorsal prefrontal cortex when animals vocalized than when they remained silent. However, given the audio-motor nature of the task it is difficult to separate the relative effects of sensory processing and preparation of the motor program for vocalization. The second study attempted to distinguish between sensory, motor, and sensorimotor integration effects by including a treatment group that vocalized without hearing any conspecific playbacks (Miller et al., 2010). Interestingly, this production-only group showed the lowest amount of *c-fos* induction for the majority of prefrontal sites tested. The animals that showed the highest levels of induction overall were those that only heard playbacks of calls. There was one area in the dorsal prefrontal cortex where the expression levels for the vocal production group matched the levels seen in other adjacent areas for the vocal perception group;

however, it is still not possible to eliminate auditory feedback induced activation of this region.

Another recent study used PET imaging to identify activation of the inferior frontal gyrus (Broca's area analog) in chimpanzees while simultaneously producing vocalizations and hand gestures. The level of activation was greater than when the animals gestured without vocalizing (Taglialetela, Russell, Schaeffer, & Hopkins, 2011). Another study that recorded neuronal activity in macaques suggests that when the monkeys produce conditioned innate vocalizations, some neurons are activated in the ventral premotor cortex (Coudé et al., 2011). However, these neurons did not fire when the animals vocalized spontaneously, indicating that they do not encode motor commands for the vocalizations.

The authors of these studies concluded that this is the first time vocalizing-driven activity has been found in the non-cingulate cortex of a non-human primate. However, it is still possible that activity observed in vocalizing groups was largely due to sensory processing of conspecific calls, the animals hearing themselves vocalize, or other features of the vocalizing setting. A control group vocalizing after deafening, like the one included in our study on mice, is required to exclude the first two alternatives. Such studies may not be feasible due to ethical concerns regarding deafening

experiments in primates. Neurophysiology experiments also need to demonstrate whether there is premotor neural firing for spontaneous vocalization, and if the recorded regions are analogous to the motor cortical areas that are critical for production of learned vocalizations in humans and songbirds. Therefore, until another approach is developed, it remains to be determined if cortical regions associated with vocal production in humans also control natural vocal production in non-human primates.

3.3.1. Mice have a forebrain vocal pathway with some similarities to humans and vocal learning birds

Mice have been assumed to lack a direct cortico-bulbar projection to Amb (Fischer & Hammerschmidt, 2010; Jarvis, 2004); however, this assumption had also never been experimentally tested until our recent study (Arriaga et al., 2012). To test the possibility of M1 input to the vocal premotor system, we performed neural tracing experiments in mice using the retrograde trans-synaptic tracer pseudorabies virus (PRV-Bartha) expressing enhanced green fluorescent protein (eGFP) injected into the cricothyroid and lateral cricoarytenoid laryngeal muscles in order to trace premotor brain pathways that converge on Amb. By approximately 4 days post-injection, a pattern of labeling was observed consistent with known connectivity in mammals (Jürgens, 2002b), including rodents (van Daele & Cassell, 2009). The PRV spread to a set of regions in the midbrain and limbic system with known roles in the control of innate species-specific calls and respiration (Jürgens, 2002b): the medullary reticular formation, spinal trigeminal nucleus, and solitary nucleus of the brainstem; PAG and ventral tegmentum of the midbrain; throughout the hypothalamus; and the amigdalopyriform transition area, and central amygdala in the telencephalon (Arriaga et al., 2012; Arriaga & Jarvis, in preparation). At the same survival time, only two neocortical regions were reliably labeled: (1) a population of layer V pyramidal neurons in M1 within the motor cortex region that exhibited robust singing-driven IEG expression (Fig. 6c and d); and (2) a small number of layer III neurons in the insular cortex (IC).

The relatively short latency at which PRV label was observed in M1 suggested that perhaps it projects directly to Amb. To test this hypothesis, we injected BDA into the M1 region identified by PRV tracing, and injected cholera toxin subunit b (CTb) into the cricothyroid and lateral cricoarytenoid laryngeal muscles (Arriaga et al., 2012). This dual tracing technique permitted visualization of motor cortical axons as well as laryngeal motoneuron somata and dendrites from the same animals. We found that the singing-activated portion of M1 projects directly to Amb. There were fine caliber M1 axons that exited the pyramidal tract, extended laterally to the zone where Amb motoneuronal cell bodies were located, and terminated on labeled motoneurons (Fig. 6e). Compared to songbirds (Wild, 1993) and the limited data on humans (Iwatsubo et al., 1990), the mouse M1 connections was much more sparse; there appeared to be no more than one or two axons per connected motor neuron.

This region of M1 also projects densely to the region of ADSt that displayed a singing-related increase of IEG expression, and connects reciprocally to the ipsilateral ventral lateral nucleus of the thalamus (VL). These two projections are likely to form part of a cortico-striatal-thalamic loop for vocalization similar to those reported in humans and song learning birds; however, the striatal projection to globus pallidus or the pallidal projection to thalamus have not been confirmed for this circuit in mice. The tracer injections in M1 also showed that this region receives a projection from neurons of the ipsilateral secondary auditory cortex (Fig. 6f). The cell bodies for the secondary auditory cortex were in layer III. This projection still needs to be confirmed in the anterograde direction.

The combined retrograde and anterograde tracing patterns show that mice have a cortical vocal premotor circuit that projects

directly to vocal motoneurons in the brainstem, the anterior striatum and thalamus, and it may receive a projection from secondary auditory cortex. These features are similar to those of known vocal production circuits in humans and song learning birds (Fig. 5). These findings suggest that a cortico-bulbar projection to vocal motoneurons is not unique to vocal learning birds and humans amongst mammals, as previously thought (Deacon, 2007; Fischer & Hammerschmidt, 2010; Fitch et al., 2010; Jarvis, 2004; Jürgens, 1982; Jürgens et al., 1982; Kirzinger & Jürgens, 1982; Okanoya, 2004; Simonyan & Horwitz, 2011; Simonyan & Jürgens, 2003).

4. Innate and learned features of mouse vocalizations

Like input from motor cortex, auditory experience seems to be more important for the production of learned vocalizations than innate calls. In humans and songbirds auditory experience plays a critical role at multiple stages in the ontogeny of vocal behavior: (1) a sensory phase during which an auditory memory or 'template' is formed following exposure to an appropriate model; (2) a sensorimotor phase during which vocal output is monitored and compared to the model in a guided learning process; (3) an adult maintenance phase during which auditory feedback is used to maintain vocal output over the long-term (Doupe & Kuhl, 1999; Marler, 1970a). We posit that the main difference between learning by imitation and improvisation is the dependence on the first stage. In imitation, the model or template is acquired externally. In improvisation there is no external model against which to measure progress, so another instructive signal must guide the learning process; however, this strategy likely involves a similar mechanism of auditory self-monitoring followed by selection and retention of preferred learned features. Auditory experience is critical under either learning paradigm. Accordingly, experiments testing for vocal learning have typically focused on modifying, disrupting, or removing auditory information at the various developmental phases. We briefly review the results from known vocal learning and non-learning species, then discuss results from recent studies performed on mice.

4.1. Effects of deafening on innate and learned vocalizations

It has been demonstrated in various mammalian (Hammerschmidt et al., 2001; Romand & Ehret, 1984; Talmage-Riggs, Winter, Ploog, & Mayer, 1972) and avian (Konishi, 1964; Kroodsma & Konishi, 1991; Nottebohm & Nottebohm, 1971) species that the acoustic structure of innate vocalizations does not depend on auditory experience at any developmental stage. Eastern phoebes (*Sayornis phoebe*), a sub-oscine vocal non-learning songbird species, develop normal species-specific songs after being mechanically deafened by cochlear removal before the onset of singing behavior (Kroodsma & Konishi, 1991), despite being very closely related to vocal learning songbirds. Similar results have been reported in the more distantly related ringdove (Nottebohm & Nottebohm, 1971) and chicken (Konishi, 1963). In non-human primates, neither hereditary deafness (Hammerschmidt et al., 2001) nor deafening by cochlear coagulation (Talmage-Riggs et al., 1972) affect normal vocal behavior. Unsurprisingly, the less severe auditory deprivation caused by social isolation also has no reported effect on monkey call spectral structure (Hammerschmidt et al., 2001; Winter, Handley, Ploog, & Schott, 1973). Even innate calls in male zebra finches, a vocal learner, are not affected by deafening (Simpson & Vicario, 1990).

In contrast, learned vocalizations are susceptible to elimination or disturbance of auditory feedback at various stages in development. In songbirds early deafening in the sensory acquisition (Marler & Waser, 1977) or sensorimotor phase of song learning

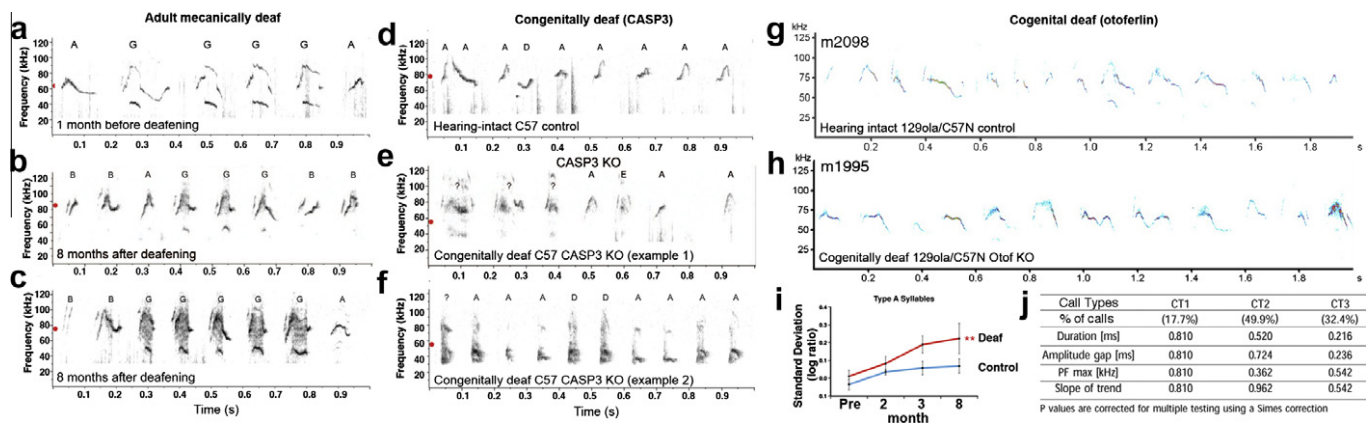


Fig. 7. Example results of deafening experiments in mice. (a) Sonograms representing 1 s of ultrasonic song from an adult mouse 1 month before deafening. (b and c) Same mouse 8 months after deafening (bilateral cochlear removal) showing the smaller (b) and larger (c) effects seen. (d) Sonogram of wild-type B6 mouse; (e and f) Same mouse strain but congenitally deaf due to knock out of the CASP3 gene, showing the smaller (e) and larger (f) effects seen on song. Red dots represent the average pitch over the entire recording session for that individual animal. (g) Sonogram of wild type *ola1/B6* male USVs. (h) Same mouse strain but congenitally deaf due to knock out of the otoferlin gene. (i) Standard deviation of the pitch of Type A syllables (expressed as a log-ratio) over 8 post-operative months (** = $p < 0.01$; repeated measures ANOVA with the Bonferroni–Dunn post hoc test comparing within-group means across recording months). Data are plotted as means \pm s.e.m. (j) p -Values for comparisons of syllable features of three major category types (CT1–CT3) between hearing-intact and otoferlin knockout mice. Panels a–f and i used with permission from Arriaga et al. (2012), and g, h, and j from Hammerschmidt et al. (2012).

(Konishi, 1965a, 1965b) has a dramatic effect, resulting in severely degraded songs characterized by a small repertoire with highly variable and unstable notes. Songbirds raised in social isolation develop highly abnormal ‘isolate song’ (Marler, 1970a, 1970b; Marler & Waser, 1977). Taken together these findings reveal that songbirds need to hear others to learn what to mimic and to hear themselves to practice their own copy. But songbirds continue to depend on auditory information even after learning and stabilizing normal songs. For example, adult Bengalese and zebra finches suffer rapid deterioration of syntax and phonology when deafened (Horita, Wada, & Jarvis, 2008; Lombardino & Nottebohm, 2000; Okanoya & Yamaguchi, 1997; Woolley & Rubel, 1997). Even the milder treatment of disrupting auditory feedback signals in real-time without deafening is sufficient to cause a destabilization of learned song features (Leonardo & Konishi, 1999; Sakata & Brainard, 2006). Thus, songbirds clearly rely heavily on auditory experience throughout the entire song development process, including for maintenance and stabilization of songs learned early in life.

Human speech shares with birdsong a dependence on auditory information throughout life (Doupe & Kuhl, 1999). Like in song learning birds, in humans early language deprivation by social isolation severely disrupts speech acquisition (Fromkin, Krashen, Curtiss, Rigler, & Rigler, 1974). In this regard, humans and some songbirds (Marler, 1970a; Thorpe, 1958) are subject to sensitive periods for vocal development. Later in life, as in song learning birds, post-lingually deaf patients suffer a degradation of speech sounds that results in decreased control of phonation, disrupted prosody, and abnormal suprasegmental properties of sentences, with younger patients being more strongly afflicted (Waldstein, 1990). Thus, vocal learners seem to make use of auditory feedback to calibrate the fine phonetic control required to produce high-quality vocalizations even after the waning of a robust vocal learning ability.

4.2. Evidence for and against a requirement of auditory feedback to maintain specific features of mouse songs

Our laboratory and several others have been conducting behavioral studies in mice to test for the presence of features found in vocal learning mammals and birds (Arriaga et al., 2012; Grimsley et al., 2011; Hammerschmidt et al., 2012; Kikusui et al., 2011).

We first focused on the role of auditory input. Based on the data from vocal learning and non-learning species discussed previously, we reasoned that if male mice learn any aspect of their courtship vocalizations, then they should require auditory information in order to maintain the spectral quality of songs. However, if songs are innate, then they should not be affected by deafening. We tested this hypothesis by mechanically deafening adult mice (Arriaga et al., 2012). Over the course of 8 months after deafening the songs of the deaf mice became spectrally distorted with some noisy looking syllables and less spectral purity than songs of sham-operated controls (Fig. 7a and b). We wondered if the noisier syllables were due to deaf mice possibly singing louder and causing microphone recording distortion, but found that the vocalizations were not on average louder than pre-deafened song. The pitch of deaf mice songs had also increased such that 6–8 months after surgery they were reliably singing at a significantly higher frequency relative to both their own pre-deafening levels and those of hearing-intact controls.

The average increase in mean pitch of post-deafening mouse songs was comparable to the 4–6 kHz increase in USVs reported for deafened horseshoe bats, an accepted vocal learning species (Rübsamen & Schäfer, 1990). The combined effects on pitch and spectral purity were similar in character and timing to changes in vocalizations observed in post-lingually deaf humans and mechanically deafened song-learning birds (Brainard & Doupe, 2000; Heaton, Dooling, & Farabaugh, 1999; Waldstein, 1990; Watanabe, Eda-Fujiwara, & Kimura, 2006; Woolley & Rubel, 1997).

We also analyzed the songs of normal hearing-intact B6 males to those of males congenitally deaf due to loss of inner ear hair cells within several days after birth resulting from knockout (KO) of the caspase 3 gene (CASP3) (Takahashi et al., 2001). We found that these mice showed larger differences in their song syllables compared to the wild type (Fig. 7c and d). Some syllables were highly degraded and barely recognizable, but with some resemblance to normal syllable categories. The changes included producing a higher proportion of the more simple Type A syllable, lower mean frequency of the pitch, greater standard deviation of the pitch, and lower spectral purity. The changes in the CASP3 KO animals songs are the largest that we are aware of for any genetically manipulated animal. However, we could still recognize features of the songs and syllables, indicative an innate component to mouse songs.

A similar study using a mouse strain congenitally deaf due to knockout of the *otoferlin* gene generated on a mixed background (129 ola and B6) found no differences in the amount of syllables/calls produced between deaf and hearing-intact mice (Fig. 7e and f) (Hammerschmidt et al., 2012). The study also found no differences in duration and amplitude, which were not affected by hearing status in our studies, but also did not find differences in pitch, although only a subset of pitch measures were assessed. From this negative result, the authors conclude that it is questionable if mice could be used as models for vocal learning.

We offer two explanations for the differing results of the deafening studies: (1) the mechanical deafening of adults and the CASP3 KO caused changes in mouse songs due to some variable other than loss of hearing; or (2) the methods used to analyze the *otoferlin* knockout mouse songs did not capture changes in the songs seen in our study. In our study, sham operated mice did not show changes in song like those in the mechanically deafened group, suggesting that disruption of the facial musculature does not explain the differences. The CASP3 gene serves many functions in different neurons, and knocking it out could have affected other brain pathways or phonatory musculature. However, CASP3 knockout did not produce overt motor deficits. For the second explanation, in the *otoferlin* knockout study, the syllables were split into only 2–3 super categories. We believe that this method groups syllables with great morphological and spectral differences, thereby potentially increasing the variability within each category. As a result, this approach risks masking effects that might be better detected by analyzing syllable types individually. For analyses on amplitude, they did split the syllables into more categories and did not find differences in the amplitude before and after deafening, similar to our own study. Another methodological difference is that the *otoferlin* study introduced an awake behaving female into the recording chamber to elicit male songs. Because females also produce some ultrasounds, it is possible that the *otoferlin* knockout mouse song recordings were contaminated with vocalizations from hearing-intact females. Moreover, the study did not report data for the three acoustic features that showed the greatest differences in our deafening experiments (mean pitch, standard deviation of the pitch distribution, and spectral purity). We believe reconciling these differences will require standardizing the experimental designs, syllable classification schemes, and spectral analysis techniques across laboratories. Until then, the methodological issues make it difficult to draw strong conclusions from the current set of different deafening results, and thus we believe the possibility of auditory dependence for normal mouse song development remains open.

Deafening-induced song deterioration alone does not demonstrate presence or absence of the vocal learning ability, but it is a strong indication that this ability may be present; to date, destabilization of vocal production after deafening has only been observed in vocal learners. However, these observations remain correlative and not diagnostic. Diagnostic test require demonstrating some form of vocal production learning, the subject of the next section.

4.3. Evidence that mouse songs are innate

Imitation of another species' vocalizations when cross-fostered, such parrots raised by humans who then imitate human speech, is the gold standard for demonstrating vocal learning. However, not even all known vocal learning species have the ability to imitate other species, and successful cultural transfer of song elements under cross-fostering can require optimal social and developmental conditions. For example, juvenile zebra finches will imitate Bengalese finch songs when raised exclusively with Bengalese finches. Yet, young zebra finches show an innate predisposition to learn

their own species song when given a choice between a Bengalese finch foster-father and a zebra finch (Clayton, 1987).

A recent study conducted a cross-fostering experiment with two strains of mice (B6 and BALB/c) that sing at different pitches, and have different distributions of syllable types in their repertoires (Kikusui et al., 2011). They cross-fostered young mice from post-natal day 0 to 21 and then scored the acoustic and syntactic structure of their songs as adults. They did not find any changes in the pitch and syllable distribution of the songs of the cross-fostered mice (Fig. 8a and b). Therefore, the authors concluded that the strains were not able to imitate each other's songs and interpreted this negative result as evidence that mouse songs are innate.

4.4. Evidence that mouse songs have some learned features

Three recent studies, including one by our own lab, have found some evidence of adaptive vocal modification of mouse USVs by examining acoustic changes that occur over the course of development (Grimsley et al., 2011), after temporary social isolation (Chabout et al., 2012), or after being housed with another male mouse with a different song in a competitive social condition (Arriaga et al., 2012). The former two showed developmental or social experience changes that could not be easily explained by innate developmental vocal trajectories, and the latter demonstrated song pitch convergence that possibly resulted from imitation.

4.4.1. Ontogeny of mouse USVs

The first study analyzed the development of CBA/CaJ mouse pup isolation calls from post-natal day 5 to post-natal day 13 and compared them to adult USVs (Grimsley et al., 2011). Using a syllable classification scheme similar to that described earlier in this review (Scattoni et al., 2008) they report changes in repertoire composition over early development (Fig. 8c and d). Notes that were flat, or contained 1 frequency jump dominated the repertoire on post-natal day 5 and post-natal day 7. From post-natal day 9 to post-natal day 13, notes with 2 frequency steps were most common. This was very different from the adult repertoire, which was dominated by one-note syllables with an upward, flat, or chevron-like trajectory. Although the relative proportions of syllables varied, all types were produced from post-natal day 7 through adulthood. The authors used a Zipf's statistic to compare the complexity of the repertoire over different developmental ages. They found that complexity steadily increased from post-natal day 5 to post-natal day 13, resulting in a more diverse and less repetitive sequence of syllables with greater higher-order structure.

Developmental changes in Syllable morphology were also reported. Generally, the duration of pup syllables tended to decrease with age. For example, the distributions of flat syllable and chevron-shaped syllable durations were tighter and had a lower mean for adult vocalizations compared to pup vocalizations. Peak frequencies of both pup syllable types were distributed bi-modally over a broad frequency range, but adult syllable peak frequencies were normally distributed over a more restricted range with a lower mean. The narrowing of the peak frequency range resulted from exclusion of the higher and lower margins of the pup peak frequency distribution for both syllable types, and syllables with dominant frequencies above 100 kHz were common in pups but rare in adults. Although the developmental trajectory of each specific syllable type varied, overall, adult syllables were shorter in duration and lower in pitch than pup syllables.

The authors concluded that the complex spectro-temporal, repertoire composition, and sequencing changes observed in mouse syllables over development could indicate a learning process, whereby pups learn to produce syllables and sequences that permit identification and more reliable retrieval, and adults

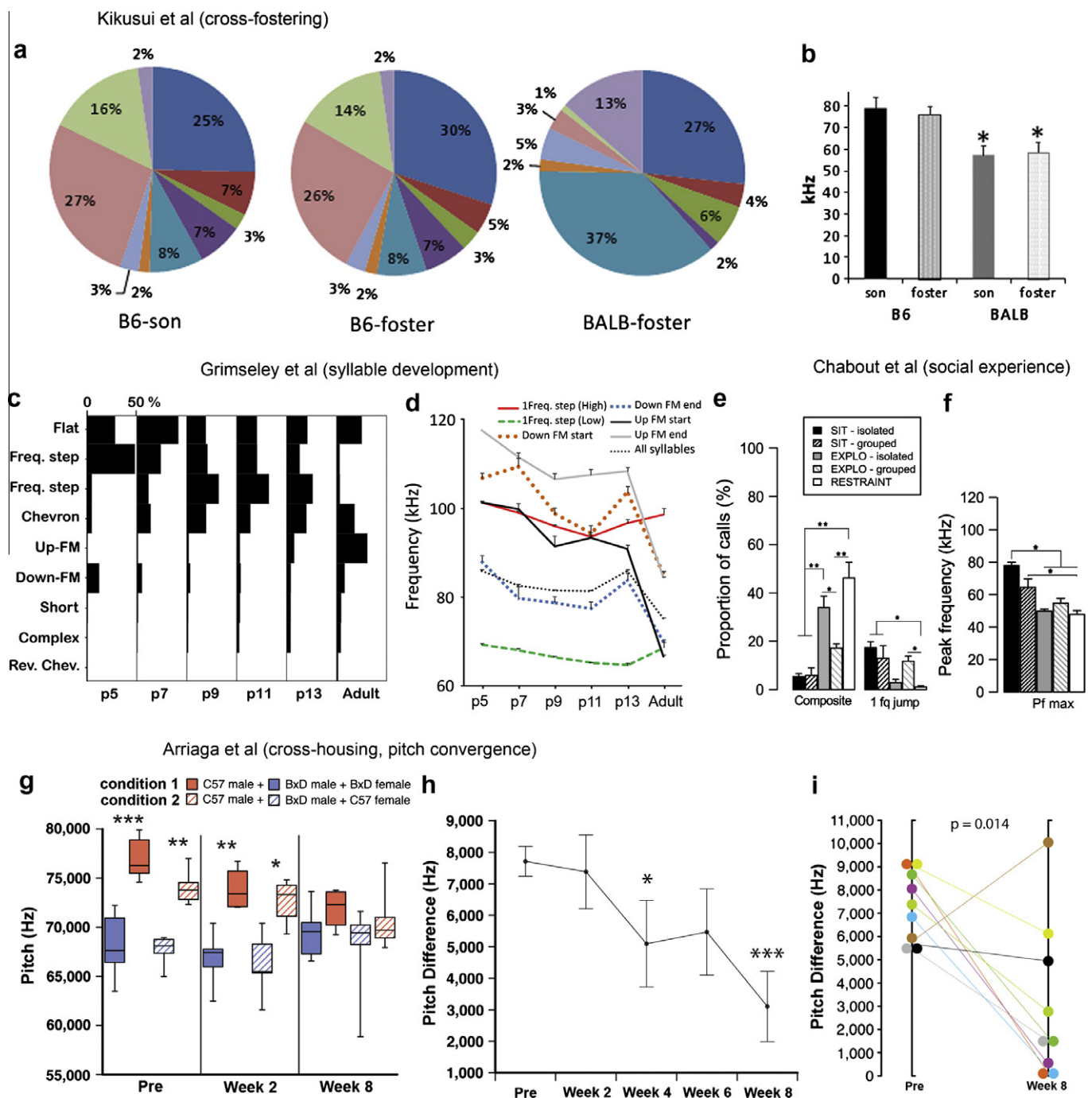


Fig. 8. Example results of vocal development and social experience on vocal behavior in mice. (a) No change in repertoire composition of syllable types (colors) of the cross fostered animals from Kikusui et al. (2011). (b) No change in mean peak frequency of biological sons and cross fostered sons of B6 and BALB mice from Kikusui et al. (2011). (c) Changes in repertoire composition (y-axis is syllable types) over development (x-axis is proportion and age) from Grimseley et al. (2011). (d) Changes in frequency (pitch) of different syllable types in different directions over development from Grimseley et al. (2011). (e) Changes repertoire composition as a result of social experience in adult mice from Chabout et al. (2012). (f) Changes in mean peak frequency as a result of social experience in adult mice from Chabout et al. (2012). (g) Convergence of pitch of Type A syllables from the songs of B6 (C57BL/6J) and BxD (B6D2F1/J) males before and over 8 weeks of cross-strain paired housing from Arriaga et al. (2012). Box plots show the median, 1st and 3rd quartile, and full range. (h) The average change in difference in pitch of Type A syllables between the two males in each B6–BxD pair from before to after 8 weeks of cross-strain paired housing (paired Student's *t*-test) from Arriaga et al. (2012). (i) The change in difference in pitch of each individual specific pair from Arriaga et al. (2012); 0 is no difference. Figures used with permission.

differentiate themselves from pups (Grimseley et al., 2011). Alternatively, there could be some complex innate maturation processes that cause the developmental patterns observed, an explanation proposed by other researchers (Hammerschmidt et al., 2012). Indeed, the authors do recognize that these data are descriptive and do not test for vocal learning capabilities, and they suggest examining vocal ontogeny in the absence of auditory feedback.

A later study found that the adult repertoire composition and some acoustic features (duration and peak frequency) of individual syllables is context-dependent (Chabout et al., 2012). Adult male mice isolated for 3 weeks produced significantly different songs than group-housed mice (Fig. 8e and f). Although not explicitly mentioned by the authors, the repertoire composition changes could represent a case of vocal usage learning through social

experience. The peak frequency changes could represent vocal production learning through social experience. One issue the authors raise is that they were unable to sort out the vocalizations between the two different mice in the dyadic social recording situation. Nevertheless, these findings suggest that social isolation of young animals could strongly affect the development of a normal song repertoire.

4.4.2. Song pitch convergence in mice

The closest evidence for some form of vocal mimicry in mice comes from our study showing syllable pitch convergence (Arriaga et al., 2012). Although overt mimicry of novel sounds is considered the gold standard for vocal learning, some researchers argue that a more limited form of vocal imitation should also be considered whereby the spectral content of innately specified conspecific calls converges (Egnor & Hauser, 2004; Snowdon, 2009; Tyack, 2008). We considered that mouse songs are produced in a mating context, and tried cross-housing sexually mature males from different strains (B6 and BxD) in a sexually competitive environment (Arriaga et al., 2012). Before crossing, the average pitch of songs from B6 and BxD males segregated into two non-overlapping distributions. After cross strain housing pairs of males along with a BxD or B6 female, over the course of 8 weeks males showed a significant convergence in pitch independent of the strain of the female present (Fig. 8g). In particular, the pitch of all B6 animals shifted downward and some BxD's shifted upward, such that after 8 weeks of cross-housing the pitches of BxD and B6 songs were no longer statistically distinguishable. Before crossing, the mean pitch difference between pairs was 8.6 ± 0.51 kHz. By 3 weeks after crossing the mean pitch difference had decreased significantly, and continued to decline to a global minimum difference of 2.1 ± 1.4 kHz at 8 weeks (Fig. 8h). Importantly, after 8 weeks of cross-housing most of the pairs had reduced their difference in pitch by more than 80 percent of their specific cage mate, and many of the pairs had converged to within 1 kHz of each other's pitch (Fig. 8i).

The results of cross-housing pairs of BxD and B6 males support the hypothesis that mice are capable of copying some features of another male's songs. The changes observed were made to an existing note type shared between both strains. Therefore, the reported change is akin to vocal convergence reported in bats. The pitch of echolocation calls of young greater horseshoe bats (*Rhinolophus ferrumequinum*) correlate strongly with the calls of their mother (Jones & Ransome, 1993). Because the pitch of a mother's calls varies with her age, the correlation with her offspring's pitch is likely to result from their learning her pitch. When female greater spear-nosed bats (*Phyllostomus hastatus*) were transferred to a new social group both the residents of the group and the new members changed the spectro-temporal features of their existing screech calls to converge on a similar call (Boughman, 1998). A recent study on greater sac-winged bats (*Saccopteryx bilineata*) showed similar convergence of young male calls onto a tutor father's call (Knörnschild, Nagy, Metz, Mayer, & von Helversen, 2010). It is unknown if the pre-convergence bat calls are innately specified or learned, but the changes are more striking than those reported for call convergence in non-human primates. Call convergence in non-human primates is based mostly on observations of within-group similarity and geographical variation in call features (Janik & Slater, 1997; Snowdon, 2009; Tyack, 2008). Some experimental evidence has been reported for pygmy marmosets (*Cebuella pygmaea*) that minimized spectral differences between each other's calls when new male/female pairs were housed in a cage together (Snowdon & Elowson, 1999).

Although the syllables we tested in mice were not novel, convergence does require the transfer of vocal elements between individuals and may reflect a rudimentary ability that could have

been expanded to include production of novel elements. The finding that B6 males changed as a group but the BxD males were relatively unaffected by cross-housing, supports our hypothesis of sexual competition. We noted that the BxD males tend to be larger and sing more than the B6 males. Therefore, the greater shift in pitch by the B6 males could reflect a tendency to try to match the pitch of a more dominant singer in the presence of a female. Another possibility is that the females co-housed with the pairs provided a selection force in the direction of their preferred range for both BxD and B6 males. While females could certainly provide a reinforcing stimulus for convergence, as in the case of cowbirds (King, 1983; West & King, 1988), the close approximation of the BxD male's pitch in most B6/BxD pairs analyzed at 8 weeks post-crossing suggests that they were likely guided by auditory information.

The pitch matching results (Arriaga et al., 2012) contradict the findings of the previously mentioned cross-fostering study (Kikusui et al., 2011). We believe the differences between studies could be explained by experimental design. First, the learning paradigm used for cross-fostering (Kikusui et al., 2011) did not ensure or test for vocal production by the foster father. Absence of tutor song production would prevent the young males from acquiring a template to mimic. Second, the cross-fostered mice were tutored at a very early age and for a very short period (21 days). For more than half of that period the pups' ear canals are closed, effectively leaving only 9 days of full auditory experience. In the pitch-matching study (Arriaga et al., 2012), the mice required at least 4–6 weeks of co-housing to begin showing pitch convergence. Lastly, prior to testing the cross-fostered mice (Kikusui et al., 2011) were returned to group housing in an acoustically unshielded colony for a much longer period (50–120 days) than the cross-fostering phase. Thus, the juveniles had more potential auditory experience with the songs of their own strain than with those of the foster father. Given the demonstrated predisposition of vocal learning species for learning their own species-typical songs, if mice are vocal learners, it is possible that the cross-fostered mice actively selected songs of their own strain for imitation during mixed housing. The mice in the pitch-matching study (Arriaga et al., 2012) were never returned to group housing during the experiment and were acoustically shielded from the songs of mice other than their cage-mate. Given the differences in design between the initial cross-fostering and pitch-matching studies (Arriaga et al., 2012; Kikusui et al., 2011), we believe that the available evidence supports the possibility of mouse song pitch learning by imitation or by improvisation.

5. Conclusions and future directions

This perspective report has examined the underlying neural circuits that support production of ultrasonic courtship songs of male laboratory mice, and described some basic capabilities of adult mice to modify and maintain the spectral content of their songs. Some of the currently available data indicate that a combination of neural and behavioral features is present in laboratory mice that had previously only been reported in humans and song learning birds. Some of these findings are being reported for the first time in non-human mammals. Further investigations will be necessary to reconcile the conflicting conclusions on auditory feedback and mouse song imitation. The discovery of brain regions and pathways involved in mouse song production should aid interpretation of past studies and inform the design of future studies investigating the effects of social, genetic, and pharmacological manipulation on vocal behavior. Additionally, the discovery of a sparse direct cortical projection to the vocal motor nucleus ambiguus, input to motor cortex from secondary auditory cortex, a controversial requirement for auditory feedback, and a capacity for adaptive

vocal modification based on social experience should inform studies investigating the distribution, development and evolution of the rare vocal learning trait. Below we propose such neurobiological and behavioral experiments that we believe may help advance the field.

5.1. Functional connections of the mouse song system

The singing-associated forebrain pathways described in this review included brain regions and connectivity similar to cortico-striatal-thalamic loops for song learning in birds and proposed loops for speech learning in humans (Fig. 5) (Jarvis, 2004; Jürgens, 2009; Lieberman, 2001). To test this idea, future experiments should investigate the proposed connections between dorsolateral striatum and the thalamus, which are likely to go through the globus pallidus. Further investigation should also test whether the cortico-striatal-thalamic circuit is dedicated to vocalization as in songbirds, a hypothesis that is difficult to test in human subjects. It is also possible that these circuits serve a non-motor function as suggested by neural activity recorded in monkey premotor cortex before and during conditioned but not spontaneous vocalizations (Coudé et al., 2011).

The direct forebrain projection to Amb in mice appears much less robust than in vocal learning birds (Wild, 1993). The analogous projection in humans also appears sparse relative to songbirds (Iwatsubo et al., 1990; Kuypers, 1958b) but stronger than in mice. We propose that density of direct motorneuron innervation could be a contributing factor to the degree of vocal learning complexity, as this aspect is known to correlate with the level of manual dexterity across mammalian species (Lemon, 2008). A recent study in rats using the same PRV-Bartha back-tracing technique employed in this study in the laryngeal muscles also found some motor cortical cells (van Daele & Cassell, 2009) as reviewed here for mice. However, they found fewer, isolated, labeled cells in primary motor cortex at a later survival time (more than 120 h after injection into laryngeal muscles). They suggest a weak and indirect connection between M1 and Amb, and propose instead that laryngeal motor cortex is located laterally in the insular cortex; however, they did not demonstrate whether it was indirect or discuss the possible implications of these findings. We did so for mice, and suggest that rats might have a rudimentary projection. The presence of a direct cortico-bulbar connection from motor cortex suggests that mice, if not rodents generally, share a neuroanatomical feature with humans not found thus far in our closest primate relatives. Finding this projection in mice makes us wonder if a similar projection may have been missed in past studies on non-human primates. Although Kuypers stated later that non-human primates lack such a direct projection (Kuypers, 1982), his first study using the neural degeneration technique in chimpanzee and macaque did state (but not show) that after M1 lesions: “Only very few, if any, degenerating elements were found among the cells of the ambiguous nuclei.” (Kuypers, 1958a)

Our experiments suggest a need for re-evaluation of a possible direct motor cortical projection to Amb in non-human primates. We performed our tracing experiments by working our way up from the laryngeal muscles, whereas the studies performed in non-human primates worked their way down from the cortex (Jürgens & Ehrenreich, 2007; Simonyan & Jürgens, 2003). We believe future investigations should try using a similar approach by injecting transynaptic tracers in the laryngeal muscles of non-human primates.

Future studies in mice should test whether the motor cortical axons detected on Amb laryngeal motorneurons make functional synaptic connections. This can be accomplished with electron microscopy, a technique that was employed previously to identify the only other known direct cortico-bulbar connection to brainstem

motoneurons in rodents from vibrissa motor cortex to VII (Grinewich, Brecht, & Osten, 2005).

5.2. Vocal mimicry

The pitch convergence after cross-strain pairings in adult mice is more pronounced than what has been reported previously for non-human primates (Snowdon & Elowson, 1999). Although the data from primates and mice are very different in nature and scale, we believe that together they could indicate a general property of limited vocal learning among mammals that was missed in prior investigations (beyond the changes to amplitude and duration that have been observed in many animal vocalizations). Furthermore, the nature and timing of the pitch convergence was similar to what has been reported for calls in bats and dolphins (Boughman, 1998; Knörnschild et al., 2010; Smolker & Pepper, 1999; Watwood, Tyack, & Wells, 2004). These results of our experiments suggest that mice are capable of at least limited vocal learning in the form of vocal convergence of existing call types. A major difference in our experiments relative Kikusui et al. (2011), was that we cross-housed animals for up to 8 weeks whereas Kikusui et al. cross-housed them for no more than 3 weeks. At 3 weeks, we did not yet see a significant group effect. To reconcile these findings, future work should be conducted on cross-fostering or tutoring for 8 weeks or longer. Future work should also investigate whether the learning abilities of mice extend beyond modification of innate templates to the generation of novel sounds or learning syllable sequences. The most convincing evidence of vocal learning would come through successful tutoring of spectral features from heterospecific, artificial, or anthropogenic sounds.

5.3. Clearly define vocal learning and categories

As a supplement to non-human primate studies and complement to songbird studies of vocal communication, mouse models can clearly serve to cover some gaps in understanding the molecular basis of vocal production, social communication dysfunctions, and the evolution of brain systems that form the basic substrates of speech. However, more work is necessary to establish how useful mouse models will be in studying the process of vocal learning. This conclusion will be chiefly determined by whether the vocal learning capabilities of mice extend beyond the limits of pitch convergence, but also requires clear definitions of what defines vocal learning.

The current framework for classifying vocal learning and non-learning species presents a dichotomous scheme whereby a species is either: (1) a vocal mimic with the associated neuroanatomical traits found shared among all vocal learning species studied to date; or (2) a vocal non-learner producing innate vocalizations without the associated neuroanatomical and developmental characteristics of learners. This schema overlooks some problematic examples, such as species that develop novel vocalizations without mimicry and the mouse, which does not appear to fully fit either category. Therefore, we propose a new scheme that we believe more accurately reflects the biophysical, ontogenetic, molecular and neuroanatomical evidence—the Continuum Hypothesis.

1. Vocalizations based on a template.
 - (a) No modification possible, strictly determined by innate central pattern generator.
 - (b) Modification of amplitude and temporal structure only.
 - (c) Modification of amplitude, temporal structure, and/or spectral structure that does not require an externally acquired target (improvisation).
 - (d) Modification of spectra-temporal structure guided by an externally acquired target (imitation-based modification of a template).

2. Vocalizations generated *de novo*.

- (a) Modification of amplitude, temporal structure, and/or spectral structure that does not require an externally acquired target (improvisation).
- (b) Modification of spectra-temporal structure guided by an externally acquired target (full mimicry).

Examples for most of the proposed categories have already been presented in this review. Based on the available data, we believe that mice should be classified in Group 1d, along with bats. Both mice and bats appear able to adaptively modify existing syllables based on experience. This represents a limited form of vocal learning. Humans and song learning birds belong to Group 2b, which can be further divided into closed-ended and open-ended learners. The latter group continues to learn as adults. Each behavioral phenotype above will likely be associated with a particular type of neural architecture, as proposed below.

1. Vocalizations controlled by midbrain.

- (a) Strictly programmed by innate central pattern generator (CPG).
- (b) Modification of CPG possible without cortical input.
- (c) Modification of CPG possible with cortical input.
- (d) Modification of CPG by cortical input guided by integrated auditory pathways.

2. Vocalizations controlled by forebrain.

- (a) Premotor control by cortical circuits without a requirement for auditory-motor integration.
- (b) Premotor control by cortical circuits guided by integrated auditory pathways (songbird system and human language circuits).

The combination of behavioral and neuroanatomical studies proposed will allow researchers to begin testing for a link between the degree of vocal learning capabilities exhibited by various species, and the distinct features of the underlying neural systems for vocalization. Properly classifying a species under this scheme will require both behavioral and neuroanatomical investigations of a given species. We predict that species able to modify the spectral content of songs will feature a direct motor cortical projection to Amb or XlIIts.

5.4. Genetically manipulating vocal learning pathways

Several recent studies have studied the effects of manipulating genes associated with speech disorders in non-human animal models. The most widely known studies investigated the FoxP2 transcription factor, a gene required for normal speech acquisition in humans and song acquisition in songbirds (Fisher & Scharff, 2009; Haesler et al., 2004, 2007; Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001). Mutating the FoxP2 gene, and introducing the human variant in mice, produced small changes in amplitude and pitch (Enard et al., 2009; Gaub, Groszer, Fisher, & Ehret, 2010). However, these studies did not employ the vocal behavior and neurobiological framework we present in this review, and the authors did not have information about the vocal neural circuits described in the present review when interpreting the effects of FoxP2. With this information, investigators can now ask if FoxP2 expression in the vocalization-activated striatal region in mice is required for pitch convergence, and whether changing the FoxP2 variant expressed in M1 (Hisaoka, Nakamura, Senba, & Morikawa, 2010) alters the strength of the projection to Amb. The identification of a direct M1 to Amb connection opens the possibility of studying the molecular basis for specifying this projection that is considered one of the most critical steps in the evolution of vocal learning. Identification of the genetic factors involved in develop-

ing this connection might even allow for inducing a connection *de novo* in non-learning species, enhancing the projection in species with limited learning abilities, and perhaps recovery of vocal learning abilities after brain injury in species that already learn vocalizations.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bandl.2012.10.002>.

References

- Aitken, P. G. (1981). Cortical control of conditioned and spontaneous vocal behavior in rhesus monkeys. *Brain and Language*, 13(1), 171–184.
- Arriaga, G., Zhou, E., & Jarvis, E. D. (2012). Of mice, birds, and men: The mouse ultrasonic song system has some features similar to humans and song-learning birds. *PLoS ONE*, 7(10), e46610.
- Bass, A. H., & McKibben, J. R. (2003). Neural mechanisms and behaviors for acoustic communication in teleost fish. *Progress in Neurobiology*, 69(1), 1–26.
- Berquist, S. W., Ho, J. P., & Metzner, W. (2010, August 23). Sound production in the isolated mouse larynx. *Society for neuroscience annual meeting*, San Diego.
- Bottjer, S. W., Halsema, K. A., Brown, S. A., & Miesner, E. A. (1989). Axonal connections of a forebrain nucleus involved with vocal learning in zebra finches. *The Journal of Comparative Neurology*, 279(2), 312–326.
- Bottjer, S. W., Miesner, E. A., & Arnold, A. P. (1984). Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science*, 224(4651), 901–903.
- Boughman, J. W. (1998). Vocal learning by greater spear-nosed bats. *Proceedings of the Royal Society of London B*, 265(1392), 227–233. <http://dx.doi.org/10.1098/rspb.1998.0286>.
- Brainard, M. S., & Doupe, A. J. (2000). Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature*, 404(6779), 762–766. <http://dx.doi.org/10.1038/35008083>.
- Branchi, I., Santucci, D., & Alleva, E. (2001). Ultrasonic vocalisation emitted by infant rodents: A tool for assessment of neurobehavioural development. *Behavioural Brain Research*, 125(1–2), 49–56.
- Broughton, W. P. (1963). *Acoustic behavior of animals*. Boston: Elsevier.
- Brudzynski, S. M. (2007). Ultrasonic calls of rats as indicator variables of negative or positive states: Acetylcholine-dopamine interaction and acoustic coding. *Behavioural Brain Research*, 182(2), 261–273. <http://dx.doi.org/10.1016/j.bbr.2007.03.004>.
- Brudzynski, S. M. (2009). Communication of adult rats by ultrasonic vocalization: Biological, sociobiological, and neuroscience approaches. *ILAR Journal/National Research Council, Institute of Laboratory Animal Resources*, 50(1), 43–50.
- Brudzynski, S. M., Kehoe, P., & Callahan, M. (1999). Sonographic structure of isolation-induced ultrasonic calls of rat pups. *Developmental Psychobiology*, 34(3), 195–204.
- Burgdorf, J., Wood, P. L., Kroes, R. A., Moskal, J. R., & Panksepp, J. (2007). Neurobiology of 50-kHz ultrasonic vocalizations in rats: Electrode mapping, lesion, and pharmacology studies. *Behavioural Brain Research*, 182(2), 274–283.
- Chabot, J., Serreau, P., Ey, E., Bellier, L., Aubin, T., Bourgeron, T., et al. (2012). Adult male mice emit context-specific ultrasonic vocalizations that are modulated by prior isolation or group rearing environment. *PLoS ONE*, 7(1), e29401. <http://dx.doi.org/10.1371/journal.pone.0029401>.
- Clayton, N. (1987). Song learning in cross-fostered zebra finches: A re-examination of the sensitive phase. *Behaviour*, 102(1/2), 67–81.
- Constantini, F., & D'Amato, F. R. (2006). Ultrasonic vocalizations in mice and rats: Social contexts and functions. *Acta Zoologica Sinica*, 52(4), 619–633.
- Coudé, G., Ferrari, P. F., Rodà, F., Maranesi, M., Borelli, E., Veroni, V., et al. (2011). Neurons controlling voluntary vocalization in the macaque ventral premotor cortex. *PLoS ONE*, 6(11), e26822. <http://dx.doi.org/10.1371/journal.pone.0026822>.
- D'Amato, F. R., Scalera, E., Sarli, C., & Moles, A. (2005). Pups call, mothers rush: Does maternal responsiveness affect the amount of ultrasonic vocalizations in mouse pups? *Behavior Genetics*, 35(1), 103–112.
- Deacon, T. W. (2007). The evolution of language systems in the human brain. In J. Kaas (Ed.), *Evolution of nervous systems* (Vol. 4, pp. 529–547). Amsterdam: Elsevier. <<http://www.teleodynamics.com/wp-content/PDF/Evolutionlanguage%20systems.pdf>>.
- Dirks, A., Fish, E. W., Kikusui, T., van der Gugten, J., Groenink, L., Olivier, B., et al. (2002). Effects of corticotropin-releasing hormone on distress vocalizations and locomotion in maternally separated mouse pups. *Pharmacology Biochemistry and Behavior*, 72(4), 993–999.
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567–631. <http://dx.doi.org/10.1146/annurev.neuro.22.1.567>.
- Dujardin, E., & Jürgens, U. (2005). Afferents of vocalization-controlling periaqueductal regions in the squirrel monkey. *Brain Research*, 1034(1–2), 114–131. <http://dx.doi.org/10.1016/j.brainres.2004.11.048>.

- Durand, S. E., Heaton, J. T., Amateau, S. K., & Brauth, S. E. (1997). Vocal control pathways through the anterior forebrain of a parrot (*Melopsittacus undulatus*). *The Journal of Comparative Neurology*, 377(2), 179–206.
- Düsterhöft, F., Häusler, U., & Jürgens, U. (2003). Neuronal activity in the periaqueductal gray and bordering structures during vocal communication in the squirrel monkey. *Neuroscience*, 123(1), 53–60.
- Egnor, S. E. R., & Hauser, M. D. (2004). A paradox in the evolution of primate vocal learning. *Trends in Neurosciences*, 27(11), 649–654. <http://dx.doi.org/10.1016/j.tins.2004.08.009>.
- Elwood, R. W., & Keeling, F. (1982). Temporal organization of ultrasonic vocalizations in infant mice. *Developmental Psychobiology*, 15(3), 221–227.
- Enard, W., Gehre, S., Hammerschmidt, K., Hölter, S. M., Blass, T., Somel, M., et al. (2009). A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell*, 137(5), 961–971. <http://dx.doi.org/10.1016/j.cell.2009.03.041>.
- Ennis, M., Xu, S.-J., & Rizvi, T. A. (1997). Discrete subregions of the rat midbrain periaqueductal gray project to nucleus ambiguus and the periambigual region. *Neuroscience*, 80(3), 829–845.
- Fee, M. S., Kozhevnikov, A. A., & Hahnloser, R. H. R. (2004). Neural mechanisms of vocal sequence generation in the songbird. *Annals of the New York Academy of Sciences*, 1016, 153–170. <http://dx.doi.org/10.1196/annals.1298.022>.
- Fehér, O., Wang, H., Saar, S., Mitra, P. P., & Tchernichovski, O. (2009). De novo establishment of wild-type song culture in the zebra finch. *Nature*, 459(7246), 564–568. <http://dx.doi.org/10.1038/nature07994>.
- Fischer, J., & Hammerschmidt, K. (2010). Ultrasonic vocalizations in mouse models for speech and socio-cognitive disorders: Insights into the evolution of vocal communication. *Genes, Brain, and Behavior*, 10(1), 17–27. <http://dx.doi.org/10.1111/j.1601-183X.2010.00610.x>.
- Fish, E. W., Faccidomo, S., Gupta, S., & Miczek, K. A. (2004). Anxiolytic-like effects of escitalopram, citalopram, and R-citalopram in maternally separated mouse pups. *The Journal of Pharmacology and Experimental Therapeutics*, 308(2), 474–480.
- Fish, E. W., Sekinda, M., Ferrari, P. F., Dirks, A., & Miczek, K. A. (2000). Distress vocalizations in maternally separated mouse pups: Modulation via 5-HT1A, 5-HT1B and GABA receptors. *Psychopharmacology*, 149(3), 277–285.
- Fisher, S. E., & Scharff, C. (2009). FOXP2 as a molecular window into speech and language. *Trends in Genetics*, 25(4), 166–177. <http://dx.doi.org/10.1016/j.tig.2009.03.002>.
- Fitch, W. T., Huber, L., & Bugnyar, T. (2010). Social cognition and the evolution of language: Constructing cognitive phylogenies. *Neuron*, 65(6), 795–814. <http://dx.doi.org/10.1016/j.neuron.2010.03.011>.
- Floody, O. R., & DeBold, J. F. (2004). Effects of midbrain lesions on lordosis and ultrasound production. *Physiology & Behavior*, 82(5), 791–804. <http://dx.doi.org/10.1016/j.physbeh.2004.06.022>.
- Foster, E. F., & Bottjer, S. W. (1998). Axonal connections of the high vocal center and surrounding cortical regions in juvenile and adult male zebra finches. *The Journal of Comparative Neurology*, 397(1), 118–138.
- Foster, E. F., & Bottjer, S. W. (2001). Lesions of a telencephalic nucleus in male zebra finches: Influences on vocal behavior in juveniles and adults. *Journal of Neurobiology*, 46(2), 142–165.
- Fromkin, V., Krashen, S., Curtiss, S., Rigler, D., & Rigler, M. (1974). The development of language in genie: A case of language acquisition beyond the “critical period”. *Brain and Language*, 1(1), 81–107.
- Gahr, M. (2000). Neural song control system of hummingbirds: Comparison to swifts, vocal learning (songbirds) and nonlearning (suboscines) passerines, and vocal learning (budgerigars) and nonlearning (dove, owl, gull, quail, chicken) nonpasserines. *The Journal of Comparative Neurology*, 426(2), 182–196.
- Gaub, S., Groszer, M., Fisher, S. E., & Ehret, G. (2010). The structure of innate vocalizations in Foxp2-deficient mouse pups. *Genes, Brain, and Behavior*, 9(4), 390–401. <http://dx.doi.org/10.1111/j.1601-183X.2010.00570.x>.
- Gourbal, B. E. F., Barthelemy, M., Petit, G., & Gabrion, C. (2004). Spectrographic analysis of the ultrasonic vocalisations of adult male and female BALB/c mice. *Naturwissenschaften*, 91(8), 381–385. <http://dx.doi.org/10.1007/s00114-004-0543-7>.
- Grimsley, J. M. S., Monaghan, J. J. M., & Wenstrup, J. J. (2011). Development of social vocalizations in mice. *PLoS ONE*, 6(3), e17460. <http://dx.doi.org/10.1371/journal.pone.0017460>.
- Grinevich, V., Brecht, M., & Osten, P. (2005). Monosynaptic pathway from rat vibrissa motor cortex to facial motor neurons revealed by lentivirus-based axonal tracing. *The Journal of Neuroscience*, 25(36), 8250–8258. <http://dx.doi.org/10.1523/JNEUROSCI.2235-05.2005>.
- Guo, Z., & Holy, T. E. (2007). Sex selectivity of mouse ultrasonic songs. *Chemical Senses*, 32(5), 463–473. <http://dx.doi.org/10.1093/chemse/bjm015>.
- Haesler, S., Rochefort, C., Georgi, B., Licznarski, P., Osten, P., & Scharff, C. (2007). Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus Area X. *PLoS Biology*, 5(12), e321. <http://dx.doi.org/10.1371/journal.pbio.0050321>.
- Haesler, S., Wada, K., Nshdejan, A., Morrissey, E. E., Lints, T., Jarvis, E. D., et al. (2004). FoxP2 expression in avian vocal learners and non-learners. *The Journal of Neuroscience*, 24(13), 3164–3175. <http://dx.doi.org/10.1523/JNEUROSCI.4369-03.2004>.
- Hage, S. R., & Jürgens, U. (2006a). Localization of a vocal pattern generator in the pontine brainstem of the squirrel monkey. *European Journal of Neuroscience*, 23(3), 840–844. <http://dx.doi.org/10.1111/j.1460-9568.2006.04595.x>.
- Hage, S. R., & Jürgens, U. (2006b). On the role of the pontine brainstem in vocal pattern generation: A telemetric single-unit recording study in the squirrel monkey. *The Journal of Neuroscience*, 26(26), 7105–7115. <http://dx.doi.org/10.1523/JNEUROSCI.1024-06.2006>.
- Hahn, M. E., Hewitt, J. K., Adams, M., & Trully, T. (1987). Genetic influences on ultrasonic vocalizations in young mice. *Behavior Genetics*, 17(2), 155–166.
- Hahnloser, R. H. R., Kozhevnikov, A. A., & Fee, M. S. (2002). An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature*, 419(6902), 65–70. <http://dx.doi.org/10.1038/nature00974>.
- Hammerschmidt, K., Freudenstein, T., & Jürgens, U. (2001). Vocal development in squirrel monkeys. *Behaviour*, 138(9), 1179–1204.
- Hammerschmidt, K., Radyushkin, K., Ehrenreich, H., & Fischer, J. (2009). Female mice respond to male ultrasonic “songs” with approach behaviour. *Biology Letters*, 5(5), 589–592. <http://dx.doi.org/10.1098/rsbl.2009.0317>.
- Hammerschmidt, K., Reisinger, E., Westkemper, K., Ehrenreich, L., Strenze, N., & Fischer, J. (2012). Mice do not require auditory input for the normal development of their ultrasonic vocalizations. *BMC Neuroscience*, 13, 40. <http://dx.doi.org/10.1186/1471-2202-13-40>.
- Hannig, S., & Jürgens, U. (2005). Projections of the ventrolateral pontine vocalization area in the squirrel monkey. *Experimental Brain Research*, 169(1), 92–105. <http://dx.doi.org/10.1007/s00221-005-0128-5>.
- Harrison, D. F. N. (1995). *The anatomy and physiology of the mammalian larynx*. Cambridge, UK: Cambridge University Press.
- Hast, M. H., Fischer, J., Wetzell, A. B., & Thompson, V. E. (1974). Cortical motor representation of the laryngeal muscles in *Macaca mulatta*. *Brain Research*, 73(2), 229–240.
- Hauser, M. D., & Konishi, M. (Eds.). (1999). *The design of animal communication*. Cambridge, MA: MIT Press.
- Heaton, J. T., Dooling, R. J., & Farabaugh, S. M. (1999). Effects of deafening on the calls and warble song of adult budgerigars (*Melopsittacus undulatus*). *Journal of the Acoustical Society of America*, 105(3), 2010–2019.
- Hisaoka, T., Nakamura, Y., Senba, E., & Morikawa, Y. (2010). The forkhead transcription factors, Foxp1 and Foxp2, identify different subpopulations of projection neurons in the mouse cerebral cortex. *Neuroscience*, 166, 551–563.
- Hofer, M. A., & Shair, H. N. (1992). Ultrasonic vocalization by rat pups during recovery from deep hypothermia. *Developmental Psychobiology*, 25(7), 511–528. <http://dx.doi.org/10.1002/dev.420250705>.
- Holy, T. E., & Guo, Z. (2005). Ultrasonic songs of male mice. *PLoS Biology*, 3(12), e386. <http://dx.doi.org/10.1371/journal.pbio.0030386>.
- Horita, H., Wada, K., & Jarvis, E. D. (2008). Early onset of deafening-induced song deterioration and differential requirements of the pallial-basal ganglia vocal pathway. *European Journal of Neuroscience*, 28(12), 2519–2532. <http://dx.doi.org/10.1111/j.1460-9568.2008.06535.x>.
- Ise, S., & Ohta, H. (2009). Power spectrum analysis of ultrasonic vocalization elicited by maternal separation in rat pups. *Brain Research*, 1283, 58–64. <http://dx.doi.org/10.1016/j.brainres.2009.06.003>.
- 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(2), 309–312.
- Janik, V. M., & Slater, P. J. B. (1997). Vocal learning in mammals. *Advances in the Study of Behavior*, 26, 59–99.
- Janik, V. M., & Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, 60(1), 1–11.
- Jarvis, E. D. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences*, 1016, 749–777.
- Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., et al. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience*, 6(2), 151–159. <http://dx.doi.org/10.1038/nrn1606>.
- Jarvis, E. D., & Mello, C. V. (2000). Molecular mapping of brain areas involved in parrot vocal communication. *The Journal of Comparative Neurology*, 419(1), 1–31.
- Jarvis, E. D., & Nottebohm, F. (1997). Motor-driven gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 94(8), 4097–4102.
- Jarvis, E. D., Ribeiro, S., Da Silva, M. L., Ventura, D., Vieliard, J., & Mello, C. V. (2000). Behaviourally driven gene expression reveals song nuclei in hummingbird brain. *Nature*, 406(6796), 628–632. <http://dx.doi.org/10.1038/35020570>.
- Jones, G., & Ransome, R. D. (1993). Echolocation calls of bats are influenced by maternal effects and change over a lifetime. *Proceedings of the Royal Society of London B*, 252(1334), 125–128.
- Jürgens, U. (1982). Afferents to the cortical larynx area in the monkey. *Brain Research*, 239(2), 377–389.
- Jürgens, U. (1983). Afferent fibers to the singular vocalization region in the squirrel monkey. *Experimental Neurology*, 80(2), 395–409.
- Jürgens, U. (1984). The efferent and afferent connections of the supplementary motor area. *Brain Research*, 300(1), 63–81.
- Jürgens, U. (1998). Neuronal control of mammalian vocalization, with special reference to the squirrel monkey. *Naturwissenschaften*, 85(8), 376–388.
- Jürgens, U. (2002a). A study of the central control of vocalization using the squirrel monkey. *Medical Engineering & Physics*, 24(7–8), 473–477.
- Jürgens, U. (2002b). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews*, 26(2), 235–258.
- Jürgens, U. (2009). The neural control of vocalization in mammals: A review. *Journal of Voice*, 23(1), 1–10. <http://dx.doi.org/10.1016/j.jvoice.2007.07.005>.
- Jürgens, U., & Alipour, M. (2002). A comparative study on the cortico-hypoglossal connections in primates, using biotin dextranamine. *Neuroscience Letters*, 328(3), 245–248.
- Jürgens, U., & Ehrenreich, L. (2007). The descending motorcortical pathway to the laryngeal motoneurons in the squirrel monkey. *Brain Research*, 1148, 90–95. <http://dx.doi.org/10.1016/j.brainres.2007.02.020>.

- Jürgens, U., Ehrenreich, L., & de Lanerolle, N. C. (2002). 2-Deoxyglucose uptake during vocalization in the squirrel monkey brain. *Behavioural Brain Research*, 136(2), 605–610.
- 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(1), 125–139.
- Jürgens, U., & Ploog, D. (1970). Cerebral representation of vocalization in the squirrel monkey. *Experimental Brain Research*, 10(5), 532–554.
- Jürgens, U., & Pratt, R. (1979). Role of the periaqueductal grey in vocal expression of emotion. *Brain Research*, 167(2), 367–378.
- Kao, M. H., & Brainard, M. S. (2006). Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *Journal of Neurophysiology*, 96(3), 1441–1455. <http://dx.doi.org/10.1152/jn.01138.2005>.
- Kao, M. H., Doupe, A. J., & Brainard, M. S. (2005). Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song. *Nature*, 433(7026), 638–643. <http://dx.doi.org/10.1038/nature03127>.
- 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(3), e17721. <http://dx.doi.org/10.1371/journal.pone.0017721>.
- King, A. P. (1983). Epigenesis of cowbird song – A joint endeavour of males and females. *Nature*, 305, 704–706.
- Kirzinger, A. (1985). Cerebellar lesion effects on vocalization of the squirrel monkey. *Behavioural Brain Research*, 16(2–3), 177–181.
- Kirzinger, A., & Jürgens, U. (1982). Cortical lesion effects and vocalization in the squirrel monkey. *Brain Research*, 233(2), 299–315.
- Kirzinger, A., & Jürgens, U. (1985). The effects of brainstem lesions on vocalization in the squirrel monkey. *Brain Research*, 358(1–2), 150–162.
- Kittelberger, J. M., Land, B. R., & Bass, A. H. (2006). Midbrain periaqueductal gray and vocal patterning in a teleost fish. *Journal of Neurophysiology*, 96(1), 71–85. <http://dx.doi.org/10.1152/jn.00067.2006>.
- Knörnschild, M., Nagy, M., Metz, M., Mayer, F., & von Helversen, O. (2010). Complex vocal imitation during ontogeny in a bat. *Biology Letters*, 6(2), 156–159. <http://dx.doi.org/10.1098/rsbl.2009.0685>.
- Knutson, B., Burgdorf, J., & Panksepp, J. (2002). Ultrasonic vocalizations as indices of affective states in rats. *Psychological Bulletin*, 128(6), 961–977.
- Konishi, M. (1963). The role of auditory feedback in the vocal behavior of the domestic fowl. *Zeitschrift für Tierpsychologie*, 20(3), 349–367.
- Konishi, M. (1964). Effects of deafening on song development in two species of juncos. *The Condor*, 66(2), 85–102.
- Konishi, M. (1965a). The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Zeitschrift für Tierpsychologie*, 22(7), 770–783.
- Konishi, M. (1965b). Effects of deafening on song development in American robins and black-headed grosbeaks. *Zeitschrift für Tierpsychologie*, 22(5), 584–599.
- Konishi, M. (1985). Birdsong: From behavior to neuron. *Annual Review of Neuroscience*, 8, 125–170.
- Kroodsma, D. E., Houlihan, P. W., Fallon, P. A., & Wells, J. A. (1997). Song development by grey catbirds. *Animal Behaviour*, 54(2), 457–464.
- Kroodsma, D. E., & Konishi, M. (1991). A suboscine bird (eastern phoebe, *Sayornis phoebe*) develops normal song without auditory feedback. *Animal Behaviour*, 42, 477–487.
- Kubikova, L., Turner, E. A., & Jarvis, E. D. (2007). The pallial basal ganglia pathway modulates the behaviorally driven gene expression of the motor pathway. *European Journal of Neuroscience*, 25(7), 2145–2160. <http://dx.doi.org/10.1111/j.1460-9568.2007.05368.x>.
- Kuypers, H. (1958a). Some projections from the peri-central cortex to the pons and lower brain stem in monkey and chimpanzee. *The Journal of Comparative Neurology*, 110(2), 221–255.
- Kuypers, H. (1958b). Corticobulbar connexions to the pons and lower brain-stem in man: An anatomical study. *Brain*, 81(3), 364–388.
- Kuypers, H. (1958c). An anatomical analysis of cortico-bulbar connexions to the pons and lower brain stem in the cat. *Journal of Anatomy*, 92(2), 198–218.
- Kuypers, H. (1958d). Pericentral cortical projections to motor and sensory nuclei. *Science*, 128(3225), 662–663.
- Kuypers, H. (1982). A new look at the organization of the motor system. *Progress in Brain Research*, 57, 381–403.
- Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, 413(6855), 519–523. <http://dx.doi.org/10.1038/35097076>.
- Lemon, R. N. (2008). Descending pathways in motor control. *Annual Review of Neuroscience*, 31, 195–218. <http://dx.doi.org/10.1146/annurev.neuro.31.060407.125547>.
- Leonardo, A., & Fee, M. S. (2005). Ensemble coding of vocal control in birdsong. *The Journal of Neuroscience*, 25(3), 652–661. <http://dx.doi.org/10.1523/JNEUROSCI.3036-04.2005>.
- Leonardo, A., & Konishi, M. (1999). Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature*, 399(6735), 466–470. <http://dx.doi.org/10.1038/20933>.
- Lieberman, P. (2001). Human language and our reptilian brain: The subcortical bases of speech, syntax, and thought. *Perspectives in Biology and Medicine*, 44(1), 32–51.
- Lombardino, A. J., & Nottebohm, F. (2000). Age at deafening affects the stability of learned song in adult male zebra finches. *The Journal of Neuroscience*, 20(13), 5054–5064.
- Lu, C. L., & Jürgens, U. (1993). Effects of chemical stimulation in the periaqueductal gray on vocalization in the squirrel monkey. *Brain Research Bulletin*, 32(2), 143–151.
- Ludlow, C. L. (2005). Central nervous system control of the laryngeal muscles in humans. *Respiratory Physiology & Neurobiology*, 147(2–3), 205–222. <http://dx.doi.org/10.1016/j.resp.2005.04.015>.
- Lütke, L., Häusler, U., & Jürgens, U. (2000). Neuronal activity in the medulla oblongata during vocalization. A single-unit recording study in the squirrel monkey. *Behavioural Brain Research*, 116(2), 197–210.
- MacLean, P. D. (1978). Effects of lesions of globus pallidus on species-typical display behavior of squirrel monkeys. *Brain Research*, 149(1), 175–196.
- Madsen, P. T., Jensen, F. H., Carder, D., & Ridgway, S. (2012). Dolphin whistles: A functional misnomer revealed by heliox breathing. *Biology Letters*, 8(2), 211–213. <http://dx.doi.org/10.1098/rsbl.2011.0701>.
- Mantyh, P. W. (1983). Connections of midbrain periaqueductal gray in the monkey. I. Ascending efferent projections. *Journal of Neurophysiology*, 49(3), 567–581.
- Marler, P. (1970a). Birdsong and speech development: Could there be parallels? *American Scientist*, 58(6), 669–673.
- Marler, P. (1970b). A comparative approach to vocal learning: Song development in white-crowned sparrows. *Journal of Comparative and Physiological Psychology*, 71(2, Pt.2), 1–25. <http://dx.doi.org/10.1037/h0029144>.
- Marler, P. (1997). Three models of song learning: Evidence from behavior. *Journal of Neurobiology*, 33(5), 501–516.
- Marler, P., & Waser, M. S. (1977). Role of auditory feedback in canary song development. *Journal of Comparative and Physiological Psychology*, 91(1), 8–16.
- Miller, C. T., DiMauro, A., Pistorio, A., Hendry, S., & Wang, X. (2010). Vocalization induced cFos expression in marmoset cortex. *Frontiers in Integrative Neuroscience*, 4(128), 1–15.
- Moles, A., Costantini, F., Garbugino, L., Zanettini, C., & D'Amato, F. R. (2007). Ultrasonic vocalizations emitted during dyadic interactions in female mice. A possible index of sociability? *Behavioural Brain Research*, 182(2), 223–230. <http://dx.doi.org/10.1016/j.bbr.2007.01.020>.
- Müller-Preuss, P., & Jürgens, U. (1976). Projections from the “cingular” vocalization area in the squirrel monkey. *Brain Research*, 103(1), 29–43.
- Müller-Preuss, P., Newman, J. D., & Jürgens, U. (1980). Anatomical and physiological evidence for a relationship between the “cingular” vocalization area and the auditory cortex in the squirrel monkey. *Brain Research*, 202(2), 307–315.
- Musolf, K., Hoffmann, F., & Penn, D. J. (2010). Ultrasonic courtship vocalizations in wild house mice, *Mus musculus musculus*. *Animal Behaviour*, 79(3), 757–764. <http://dx.doi.org/10.1016/j.anbehav.2009.12.034>.
- Noirot, E., & Pye, D. (1969). Sound analysis of ultrasonic distress calls of mouse pups as a function of their age. *Animal Behaviour*, 17(2), 340–349.
- Nottebohm, F., & Nottebohm, M. E. (1971). Vocalizations and breeding behaviour of surgically deafened ring doves (*Streptopelia risoria*). *Animal Behaviour*, 19(2), 313–327.
- Nottebohm, F., Paton, J. A., & Kelley, D. B. (1982). Connections of vocal control nuclei in the canary telencephalon. *The Journal of Comparative Neurology*, 207(4), 344–357.
- Nottebohm, F., Stokes, T. M., & Leonard, C. M. (1976). Central control of song in the canary, *Serinus canarius*. *The Journal of Comparative Neurology*, 165(4), 457–486. <http://dx.doi.org/10.1002/cne.901650405>.
- Nunez, A. A., Pomerantz, S. M., Bean, N. J., & Youngstrom, T. G. (1985). Effects of laryngeal denervation on ultrasound production and male sexual behavior in rodents. *Physiology & Behavior*, 34(6), 901–905.
- Nyby, J. (1983). Ultrasonic vocalizations during sex behavior of male house mice (*Mus musculus*): A description. *Behavioral and Neural Biology*, 39(1), 128–134.
- Okanoya, K. (2004). Functional and structural pre-adaptations to language: Insight from comparative cognitive science into the study of language origin. *Japanese Psychological Research*, 46(3), 207–215.
- Okanoya, K., & Yamaguchi, A. (1997). Adult Bengalese finches (*Lonchura striata* var. domestica) require real-time auditory feedback to produce normal song syntax. *Journal of Neurobiology*, 33(4), 343–356.
- Okuhata, S., & Saito, N. (1987). Synaptic connections of thalamo-cerebral vocal nuclei of the canary. *Brain Research Bulletin*, 18(1), 35–44.
- Olveczky, B. P., Andalman, A. S., & Fee, M. S. (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biology*, 3(5), e153. <http://dx.doi.org/10.1371/journal.pbio.0030153>.
- Paton, J. A., Manogue, K. R., & Nottebohm, F. (1981). Bilateral organization of the vocal control pathway in the budgerigar, *Melopsittacus undulatus*. *The Journal of Neuroscience*, 1(11), 1279–1288.
- Person, A. L., Gale, S. D., Farries, M. A., & Perkel, D. J. (2008). Organization of the songbird basal ganglia, including area X. *The Journal of Comparative Neurology*, 508(5), 840–866. <http://dx.doi.org/10.1002/cne.21699>.
- Pomerantz, S. M., Nunez, A. A., & Bean, N. J. (1983). Female behavior is affected by male ultrasonic vocalizations in house mice. *Physiology & Behavior*, 31(1), 91–96.
- Portfors, C. V. (2007). Types and functions of ultrasonic vocalizations in laboratory rats and mice. *Journal of the American Association for Laboratory Animal Science*, 46(1), 28–34.
- Roberts, L. H. (1975). Evidence for the laryngeal source of ultrasonic and audible cries of rodents. *Journal of Zoology*, 175(2), 243–257.
- Romand, R., & Ehret, G. (1984). Development of sound production in normal, isolated, and deafened kittens during the first postnatal months. *Developmental Psychobiology*, 17(6), 629–649. <http://dx.doi.org/10.1002/dev.420170606>.
- Rübsamen, R., & Schäfer, M. (1990). Audiovocal interactions during development? Vocalisation in deafened young horseshoe bats vs. audition in vocalisation-impaired bats. *Journal of Comparative Physiology A*, 167(6), 771–784.

- Sakata, J. T., & Brainard, M. S. (2006). Real-time contributions of auditory feedback to avian vocal motor control. *The Journal of Neuroscience*, 26(38), 9619–9628. <http://dx.doi.org/10.1523/JNEUROSCI.2027-06.2006>.
- Sales, G. D., & Smith, J. C. (1978). Comparative studies of the ultrasonic calls of infant murid rodents. *Developmental Psychobiology*, 11(6), 595–619.
- Scattoni, M. L., Crawley, J. N., & Ricceri, L. (2009). Ultrasonic vocalizations: A tool for behavioural phenotyping of mouse models of neurodevelopmental disorders. *Neuroscience and Biobehavioral Reviews*, 33(4), 508–515. <http://dx.doi.org/10.1016/j.neubiorev.2008.08.003>.
- Scattoni, M. L., Gandhi, S. U., Ricceri, L., & Crawley, J. N. (2008). Unusual repertoire of vocalizations in the BTBR T+tf/J mouse model of autism. *PLoS ONE*, 3(8), e3067. <http://dx.doi.org/10.1371/journal.pone.0003067>.
- Scharff, C., & Nottebohm, F. (1991). A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: Implications for vocal learning. *The Journal of Neuroscience*, 11(9), 2896–2913.
- Schusterman, R. J. (2008). Vocal learning in mammals with special emphasis on pinnipeds. In D. K. Oller & U. Griebel (Eds.), *Evolution of communicative flexibility: Complexity, creativity, and adaptability in human and animal communication* (pp. 41–70). Cambridge, MA: The MIT Press.
- Schusterman, R. J., & Reichmuth, C. (2008). Novel sound production through contingency learning in the Pacific walrus (*Odobenus rosmarus divergens*). *Animal Cognition*, 11(2), 319–327. <http://dx.doi.org/10.1007/s10071-007-0120-5>.
- Seyfarth, R. M., & Cheney, D. L. (1986). Vocal development in vervet monkeys. *Animal Behaviour*, 34(6), 1640–1658.
- Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, 210(4471), 801–803.
- Siebert, S., & Jürgens, U. (2003). Vocalization after periaqueductal grey inactivation with the GABA agonist muscimol in the squirrel monkey. *Neuroscience Letters*, 340(2), 111–114.
- Simões, C. S., Vianney, P. V. R., de Moura, M. M., Freire, M. A. M., Mello, L. E., Sameshima, K., et al. (2010). Activation of frontal neocortical areas by vocal production in marmosets. *Frontiers in Integrative Neuroscience*, 4. <http://dx.doi.org/10.3389/fnint.2010.00123>.
- Simonyan, K., & Horwitz, B. (2011). Laryngeal motor cortex and control of speech in humans. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, 17(2), 197–208. <http://dx.doi.org/10.1177/1073858410386727>.
- Simonyan, K., & Jürgens, U. (2002). Cortico-cortical projections of the motorcortical larynx area in the rhesus monkey. *Brain Research*, 949(1–2), 23–31.
- Simonyan, K., & Jürgens, U. (2003). Efferent subcortical projections of the laryngeal motorcortex in the rhesus monkey. *Brain Research*, 974(1–2), 43–59.
- Simonyan, K., & Jürgens, U. (2004). Afferent subcortical connections into the motor cortical larynx area in the rhesus monkey. *Neuroscience*, 130(1), 119–131. <http://dx.doi.org/10.1016/j.neuroscience.2004.06.071>.
- Simonyan, K., & Jürgens, U. (2005). Afferent cortical connections of the motor cortical larynx area in the rhesus monkey. *Neuroscience*, 130(1), 133–149. <http://dx.doi.org/10.1016/j.neuroscience.2004.08.031>.
- Simpson, H. B., & Vicario, D. S. (1990). Brain pathways for learned and unlearned vocalizations differ in zebra finches. *The Journal of Neuroscience*, 10(5), 1541–1556.
- Smolker, R., & Pepper, J. W. (1999). Whistle convergence among allied male bottlenose dolphins (*Delphinidae*, *Tursiops* sp.). *Ethology*, 105(7), 595–618.
- Snowdon, C. T. (2009). *Plasticity of communication in nonhuman primates* (Vol. 40, pp. 239–276). Elsevier. [http://dx.doi.org/10.1016/S0065-3454\(09\)40007-X](http://dx.doi.org/10.1016/S0065-3454(09)40007-X).
- Snowdon, C. T., & Elowson, A. M. (1999). Pygmy marmosets modify call structure when paired. *Ethology*, 105(10), 893–908.
- Striedter, G. F. (1994). The vocal control pathways in budgerigars differ from those in songbirds. *The Journal of Comparative Neurology*, 343(1), 35–56. <http://dx.doi.org/10.1002/cne.903430104>.
- Sutton, D., Larson, C., & Lindeman, R. C. (1974). Neocortical and limbic lesion effects on primate phonation. *Brain Research*, 71, 61–75.
- Tagliabattola, J. P., Russell, J. L., Schaeffer, J. A., & Hopkins, W. D. (2011). Chimpanzee vocal signaling points to a multimodal origin of human language. *PLoS ONE*, 6(4), e18852. <http://dx.doi.org/10.1371/journal.pone.0018852>.
- Takahashi, K., Kamiya, K., Urase, K., Suga, M., Takizawa, T., Mori, H., et al. (2001). Caspase-3-deficiency induces hyperplasia of supporting cells and degeneration of sensory cells resulting in the hearing loss. *Brain Research*, 894(2), 359–367.
- Talmage-Riggs, G., Winter, P., Ploog, D., & Mayer, W. (1972). Effect of deafening on the vocal behavior of the squirrel monkey (*Saimiri sciureus*). *Folia Primatologica*, 17(5), 404–420.
- Thomas, L. B., Stemple, J. C., Andreatta, R. D., & Andrade, F. H. (2009). Establishing a new animal model for the study of laryngeal biology and disease: An anatomic study of the mouse larynx. *Journal of Speech, Language, and Hearing Research*, 52(3), 802–811.
- Thoms, G., & Jürgens, U. (1987). Common input of the cranial motor nuclei involved in phonation in squirrel monkey. *Experimental Neurology*, 95(1), 85–99.
- Thorpe, W. H. (1958). The learning of song patterns by birds, with especial reference to the song of the chaffinch *Fringilla coelebs*. *Ibis*, 100(4), 535–570.
- Travers, J. B., & Norgren, R. (1983). Afferent projections to the oral motor nuclei in the rat. *The Journal of Comparative Neurology*, 220(3), 280–298.
- Tyack, P. L. (2008). Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. *Journal of Comparative Psychology*, 122(3), 319–331. <http://dx.doi.org/10.1037/a0013087>.
- van Daele, D. J., & Cassell, M. D. (2009). Multiple forebrain systems converge on motor neurons innervating the thyroarytenoid muscle. *Neuroscience*, 162(2), 501–524. <http://dx.doi.org/10.1016/j.neuroscience.2009.05.005>.
- Wada, K., Sakaguchi, H., Jarvis, E. D., & Hagiwara, M. (2004). Differential expression of glutamate receptors in avian neural pathways for learned vocalization. *The Journal of Comparative Neurology*, 476(1), 44–64. <http://dx.doi.org/10.1002/cne.20201>.
- Waldstein, R. S. (1990). Effects of postlingual deafness on speech production: Implications for the role of auditory feedback. *Journal of the Acoustical Society of America*, 88(5), 2099–2114.
- Watanabe, A., Eda-Fujiwara, H., & Kimura, T. (2006). Auditory feedback is necessary for long-term maintenance of high-frequency sound syllables in the song of adult male budgerigars (*Melopsittacus undulatus*). *Journal of Comparative Physiology A*, 193(1), 81–97. <http://dx.doi.org/10.1007/s00359-006-0173-y>.
- Watwood, S. L., Tyack, P. L., & Wells, R. S. (2004). Whistle sharing in paired male bottlenose dolphins, *Tursiops truncatus*. *Behavioral Ecology and Sociobiology*, 55(6), 531–543.
- West, M. J., & King, A. P. (1988). Female visual displays affect the development of male song in the cowbird. *Nature*, 334(6179), 244–246. <http://dx.doi.org/10.1038/334244a0>.
- Wild, J. M. (1993). Descending projections of the songbird nucleus robustus archistriatalis. *The Journal of Comparative Neurology*, 338(2), 225–241. <http://dx.doi.org/10.1002/cne.903380207>.
- Wild, J. M. (1994). The auditory–vocal–respiratory axis in birds. *Brain, Behavior and Evolution*, 44(4–5), 192–209.
- Wild, J. M. (1997). Neural pathways for the control of birdsong production. *Journal of Neurobiology*, 33(5), 653–670.
- Williams, H., & Mehta, N. (1999). Changes in adult zebra finch song require a forebrain nucleus that is not necessary for song production. *Journal of Neurobiology*, 39(1), 14–28.
- Winter, P., Handley, P., Ploog, D., & Schott, D. (1973). Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. *Behaviour*, 47(3), 230–239.
- Wöhr, M., Dalhoff, M., Wolf, E., Holsboer, F., Schwarting, R. K. W., & Wotjak, C. T. (2008). Effects of genetic background, gender, and early environmental factors on isolation-induced ultrasonic calling in mouse pups: An embryo-transfer study. *Behavior Genetics*, 38(6), 579–595.
- Wöhr, M., Houx, B., Schwarting, R. K., & Spruijt, B. (2008). Effects of experience and context on 50-kHz vocalizations in rats. *Physiology & Behavior*, 93(4–5), 766–776. <http://dx.doi.org/10.1016/j.physbeh.2007.11.031>.
- Woolley, S. M. N., & Rubel, E. W. (1997). Bengalese finches *Lonchura striata domestica* depend upon auditory feedback for the maintenance of adult song. *The Journal of Neuroscience*, 17(16), 6380–6390.
- Yajima, Y., Hayashi, Y., & Yoshii, N. (1982). Ambiguus motoneurons discharging closely associated with ultrasonic vocalization in rats. *Brain Research*, 238(2), 445–450.
- Yajima, Y., & Larson, C. R. (1993). Multifunctional properties of ambiguous neurons identified electrophysiologically during vocalization in the awake monkey. *Journal of Neurophysiology*, 70(2), 529–540.
- Yu, A. C., & Margoliash, D. (1996). Temporal hierarchical control of singing in birds. *Science*, 273(5283), 1871–1875.
- Zann, R. (1990). Song and call learning in wild zebra finches in south-east Australia. *Animal Behaviour*, 40(5), 811–828.