46 Vocal Learning and Spoken Language: Insights from Animal Models with an Emphasis on Genetic Contributions

CONSTANCE SCHARFF, MIRJAM KNÖRNSCHILD, AND ERICH D. JARVIS

1 From Language and Speech to Biolinguistics

Language is a specialized human trait, tightly linked to our intellectual capacity (Hauser, Chomsky, & Fitch, 2002). Scores of philosophers, linguists, and biologists have commented on the astonishing complexity of language and proposed different scenarios of how it might have arisen (Christiansen & Kirby, 2003; Di Sciullo & Boeckx, 2011; Nowak, 2000). We learn language as infants and use it throughout our lives to express our thoughts. One feature that makes this possible is that spoken language combines a limited repertoire of sounds into an essentially unlimited combination of words and sentences. This fact already fascinated medieval Arab and Jewish religious scholars (Eco, 1995) and eventually, Wilhelm von Humboldt (1836, p. 106) put it succinctly as "the infinite use of finite means."

From a biolinguistic perspective (Boeckx & Grohmann, 2013; Hauser et al., 2002), one can ask whether language evolution required unique components, including human-unique genes, or whether it was the result of a unique combination or of a continuum of traits that exist to varying degrees in other animals (Fitch, 2010; Larson, Déprez, & Yamakido, 2010; Petkov & Jarvis, 2012; Scharff & Petri, 2011). Which of these traits are relevant to language evolution is a matter of debate (Arbib, Liebal, & Pika, 2008; Bolhuis, Tattersall, Chomsky, & Berwick, 2014; Tomasello, 2005). Speech is one feature suitable for comparative studies because many animals also communicate vocally and in a subset of those the production of those vocalizations are learned. How does the neural control of jaw, orofacial, and laryngeal muscles necessary for speech relate to sound production mechanisms in vocally communicating animals? These and other key components of spoken language, such as vocal learning, syntax, and semantics, which concern the acquisition, structure, and meaning of a signal, respectively, are increasingly the focus of comparative biolinguistic studies (Christiansen & Kirby, 2003; Fitch 2010; Larson et al., 2010; Petkov & Jarvis, 2012).

1.1. Definitions of Social Learning in Vocal Communication To understand the relevance of animal models to study genes relevant for language learning, we briefly review some relevant terms (Janik & Slater, 1997, 2000; Jarvis, 2004; Petkov & Jarvis, 2012). *Vocal learning*, or more precisely, *vocal production learning* (VPL), is the ability to learn to produce new vocalizations based on auditory feedback (vocal imitation) or reshape existing vocalizations based on social influences (vocal modification). *Vocal usage learning* (VUL) is the ability of a sender to learn which sound signal (whether it be an innate or learned sound) to produce in a specific context. *Auditory comprehension learning* is the ability of a receiver to learn which message a sound signal carries in a particular context (table 46.1).

Of these types of learning, VPL is the rarest. It is documented in only three orders of birds (songbirds, parrots, and hummingbirds) and in five orders of mammals (humans, cetaceans, bats, elephants, and pinnipeds) (Janik & Slater, 1997, 2000; Petkov & Jarvis, 2012). Songbirds and parrots are particularly accomplished in a special form of VPL, *vocal mimicry*, where even heterospecific sounds are copied, including human words and sentences (Garamszegi, Eens, Pavlova, Avilés, & Møller, 2007; Kelley, Coe, Madden, & Healy, 2008; Pepperberg, 2010). In these instances, mimicry of other species often occurs by socialization with them and absence of their own species. Except for parrots, imitation of human speech (Pepperberg, 1981) has been reported in an Asian elephant (Stoeger et al.,

Table 46.1

Definitions of terms used in the chapter with comparison to terms used in Janik and Slater (2000) and Petkov and Jarvis (2012)

Present chapter	Janik and Slater (2000)	Petkov and Jarvis (2012)
Auditory comprehension learning: "the ability of a receiver to learn which message a sound signal carries in a particular context."	*Comprehension learning*: "a receiver comes to extract a novel meaning from a signal as a result of experience with the usage of signals by other individuals" (table 1).	*Auditory learning*: "an animal learns to perceive something novel or behaviorally react to sounds differently as a result of experience" (p. 4).
Vocal usage learning: "the ability of a sender to learn which sound signal to produce in a specific context."	*Usage learning:* "an existing signal is produced in a new context as a result of experience with the usage of signals by other individuals" (table 1).	*Usage vocal learning*: "an animal learns to use acoustically innate or already learned vocalizations in a new context" (p. 4).
Vocal production learning (VPL): "the ability to learn to produce new vocalizations based on auditory feedback (*vocal imitation*) or reshape existing vocalizations based on social influences (*vocal modification*)."	*Production learning*: "signals are modified in form as a result of experience with those of other individuals. This can lead to signals that are either similar or dissimilar to the model" (table 1).	*Production vocal learning*: the ability of an animal to produce novel vocalizations.
Limited VPL: the ability to learn to produce new vocalizations is limited. The adult vocal repertoire of species with limited VPL may not need to be imitatively learned from an external source, but can be (see obligatory/facultative VPL).		Limited vocal learning
Extensive VPL: the ability to learn to produce new vocalizations is extensive and adult vocalizations of species with extensive VPL are usually imitatively learned from an external source. Note that different vocalizations (song, calls) can differ in the degree of VPL.		High vocal learning
Oobligatory VPL: VPL is obligatory in species that cannot acquire their species-typical sound repertoire in the absence of an external source.		

2012), bottlenose dolphins (Lilly, 1965), a male harbor seal (Ralls, Fiorelli, & Gish, 1985), and a male beluga whale (Ridgway, Carder, Jeffries, & Todd, 2012). Heterospecific imitation was described in a young Risso's dolphin cross-fostered with bottlenose dolphins (Favaro et al., 2016), a juvenile free-ranging orca separated from its natal group that imitated the barks of sea lions (Foote et al., 2006), and a single African elephant housed with Asian elephants (Poole, Tyack, & Stoeger-Horwath, 2005). Cross-species VPL has also been noted in children deprived of human contact but fortunately, examples are few (Bettelheim, 1959; Ogburn, 1959).

In contrast, auditory comprehension learning is widespread among animals (Seyfarth & Cheney, 2010; Petkov & Jarvis, 2012). For example, dogs can learn to respond to the sound *sit* by sitting down (Jarvis, 2004; Kaminski, Call, & Fisher, 2004), but they cannot learn to produce the sound *sit*. VUL is also not uncommon. Some animals learn which vocalizations to use when faced with particular types of predators (Blumstein, 1999; Seyfarth & Cheney, 2003; Scarantino & Clay, 2015). For example, vervet monkeys utter different types of alarm calls for different predators, and conspecifics respond with the appropriate avoidance behavior. Juveniles need not learn how to produce these alarm calls but they must learn when to use them (Seyfarth, Cheney, & Marler, 1980a, 1980b). Rhesus macaques can be trained to produce specific calls in response to specific visual stimuli (Hage, Gavrilov, & Nieder, 2013; Hage & Nieder, 2013) and walruses in response to hand gestures (Schusterman & Reichmuth, 2008). The fact that through VUL such species can be taught to utter particular vocalizations to request particular items (Pepperberg, 1981; Richards, 1986) suggests that at least some animals make the connection between arbitrary sounds and objects (and in the case of Alex the parrot, among sounds and the shape, color, or material of objects; Pepperberg, 2010). Given that mapping of arbitrary sound to meaning is a prerequisite for the semanticity of language, claims about the unsuitability of animal communication to inform domains of language besides speech should be reconsidered (Berwick, Okanoya, Beckers, & Bolhuis, 2011; Mozzi et al., 2016).

Traditionally, VPL has been discussed in a binary way; some species are capable of VPL (e.g., humans and songbirds) and others are not (e.g., monkeys and mice). However, work showing modifications of vocalizations in species previously considered incapable of VPL has led to the continuum hypothesis of VPL (Arriaga & Jarvis, 2013; Petkov & Jarvis, 2012). In this framework, VPL is not a dichotomous trait but a continuum (figure 46.1), ranging from subtle modifications of vocalizations in mice (Arriaga, Zhou, & Jarvis, 2012), goats (Briefer & McElligott, 2012), and chimpanzees (Watson et al., 2015); to social influences on repertoire maturation in bats (Prat, Taub, & Yovel, 2015) and marmosets (Gultekin & Hage, 2017; Takahashi et al., 2015; Takahashi, Liao, & Ghazanfar, 2017); to full imitative acquisition of communication sounds in many songbirds and speech in humans (Doupe & Kuhl, 1999). Species capable of VPL do not always use it, it can be facultative (Hammerschmidt et al., 2012). For instance, laboratory mice can acquire normal song without external input (Hammerschmidt et al., 2012; Mahrt, Perkel, Tong, Rubel, & Portfors, 2013), but when available, external input can lead to modified song (Arriaga et al., 2012).

In the context of the continuum theory, it is noteworthy that among the thousands of species of songbirds there is a range of how much of the song is learned from an external auditory source (often an adult "tutor") and how much of the song is internally represented ("innate"). Humans and many songbirds develop abnormal vocal communication when deprived of external input (Fromkin, Krashen, Curtiss, Rigler, & Rigler, 1974; Price, 1979; Williams, Kilander, & Sotanski, 1993), making VPL obligatory for developing species-typical communication sounds. However, some songbirds can develop normal sounding song without exposure to song models. As long as they can hear themselves they use this to guide their own vocalizations toward an internal model. Examples are European sedge warblers (Leitner, Nicholson, Leisler, DeVoogd, & Catchpole, 2002), gray catbirds (Kroodsma, Houlihan, Fallon, & Wells, 1997), and chipping sparrows (Liu & Nottebohm, 2007). The different strengths of internal predispositions toward species-typical vocalizations are often only uncovered in the absence of the normally occurring external, that is, auditory, input (Hammerschmidt et al., 2012; Mets & Brainard, 2018). For instance, Fehér, Wang, Saar, Mitra, and Tchernichovski (2009) found that zebra finches, which typically copy their tutor's song with high fidelity, do not faithfully copy experimentally induced "bad" zebra finch songs. Instead, they seem to be internally biased toward species-typical song features and improve the "bad" tutor model toward more species-typical song features.

In summary, the ability to modify communicative vocalizations in response to auditory feedback occurs on a continuum; on one end are species that communicate with sounds that mostly have been imitatively learned from external models (hereafter *extensive VPL*). At the other end are species that can develop their sound repertoire without reference to an external model, but those sounds can still be modified in limited ways (hereafter *limited VPL*). Importantly, both extensive and limited VPL require auditory feedback, whereas changes to vocalizations through arousal do not (Arriaga & Jarvis, 2013; Janik & Slater, 1997; Konishi, 1985; Petkov & Jarvis, 2012).

1.2. Occurrence of Vocal Production Learning in Birds and Mammals Extensive VPL is widespread in 3 of the 42 avian orders (Kroodsma, 1982; figure 46.2A, table 46.2) and 5 of the 28 mammalian orders mentioned earlier (figure 46.2B, table 46.3). Among the three avian orders, oscine songbirds are the most numerous, with more than 4,000 species, followed by ~300 species each of parrots and hummingbirds, of the over 10,500 species of birds. VPL in songbirds has been studied in more detail in songbirds (Catchpole & Slater, 1995; Kroodsma & Miller, 1996) than in parrots (Farabaugh, Linzenbold, & Dooling, 1994; Pepperberg, 2010) or hummingbirds (Baptista & Schuchmann, 1990; Ferreira, Smulders, Sameshima, Mello, & Jarvis, 2006; Gahr, 2000; Gaunt, Baptista, Sánchez, & Hernandez, 1994). Songbirds' close relatives, the suboscines, are generally regarded as not capable of VPL (Kroodsma, 1989; Kroodsma & Konishi, 1991), but bellbirds are (Kroodsma et al., 2013; Saranathan, Hamilton, Powell, Kroodsma, & Prum, 2007). Experimentally demonstrated absence of extensive VPL has also been reported for several nonpasserine birds such as domestic chickens (Konishi, 1963), ring doves (Nottebohm & Nottebohm, 1971), barn owls, and gulls (Gahr, 2000); these species can develop normal vocalizations without auditory feedback (Gahr, 2000; Jarvis; 2004; Petkov & Jarvis, 2012), but it is possible that more detailed analysis would reveal limited VPL. Although most avian orders have not been analyzed for the absence or presence of extensive VPL in detail so far, there are no

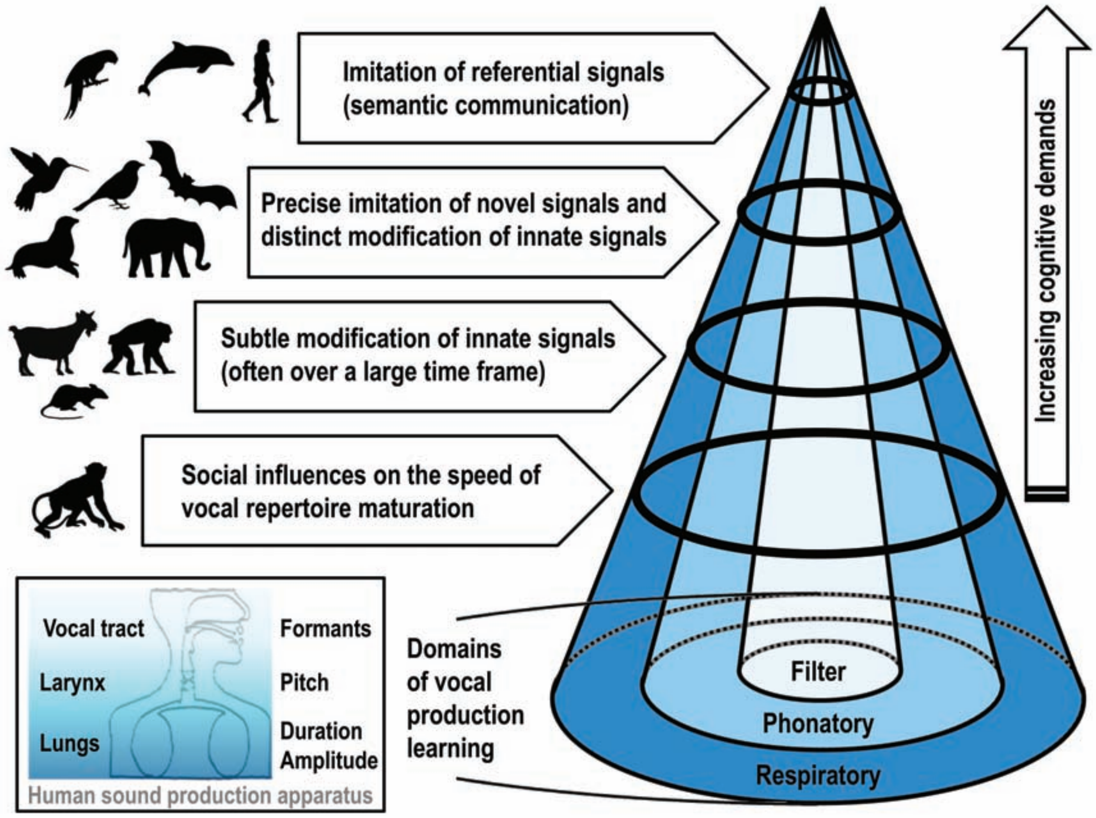

FIGURE 46.1 Different forms of VPL are arranged on a continuum of ability. Basic forms of VPL, such as social influences on repertoire maturation or subtle modifications of innate vocalizations over a long time period, constitute one end of the continuum, while complex forms of VPL, such as the imitation of referential signals, constitute the other end. Three domains of vocal production learning (respiratory, phonatory, and filter learning), their association with the sound producing apparatus, and the resulting signal characteristics are depicted as well. Sketches provide graphic references to avian and mammalian vocal production learners (on the order level).

known obvious vocal mimics such as songbirds and parrots.

Among mammals, VPL of varying degrees has been described in multiple species of cetaceans (Janik, 2014), pinnipeds (e.g., harbor seal: Ralls, Fiorelli, & Gish, 1985; walrus: Schusterman & Reichmuth, 2008), and bats (Knörnschild, 2014; Prat, Azoulay, Dor, & Yovel, 2017; Prat et al., 2015). Limited VPL exists in three mammalian orders: even-toed ungulates (Briefer & McElligott, 2012); nonhuman primates (chimpanzees: Watson et al., 2015; common marmosets: Gultekin & Hage, 2017; Takahashi et al., 2015; Takahashi et al., 2017; pygmy marmosets: Snowdon & Elowson, 1999, but also see Fischer, Wheeler, & Higham, 2015; Mahrt et al., 2013), and murids (Arriaga et al., 2012, but see Hammerschmidt et al., 2012). Concerning the capacity for VPL in mice, it is conceivable that VPL plays a bigger role in wild mice than in inbred laboratory mice for lack of sexual selection pressure in the latter (Chalfin et al., 2014; von Merten, Hoier, Pfeifle, & Tautz, 2014).

1.3. DOMAINS OF VOCAL PRODUCTION THAT CAN BE SHAPED BY LEARNING

Vocalizations are defined by temporal and spectral parameters, among them duration, pitch, timbre, and the order in which sound elements are produced. Those parameters are shaped by different components of the body's sound production apparatus (Fitch, 2006; Kriengwatana & Beckers, chapter 48 of this volume) and their neural control by the brain (Arriaga & Jarvis, 2013; Jürgens, 2009; Mackevicius & Fee, 2017; Petkov & Jarvis 2012). Because the development and function of this system involve the concerted activity of many genes, we will point out the parts of the sound production system associated with different domains of VPL (Janik & Slater, 1997, 2000; figure 46.1).

The respiratory domain of vocal production is associated with the lungs and shapes the duration, interval, and amplitude of vocalizations. The phonatory domain of vocal production is associated with the larynx in mammals and the syrinx in birds, the sound source that shapes the pitch and some other acoustic features

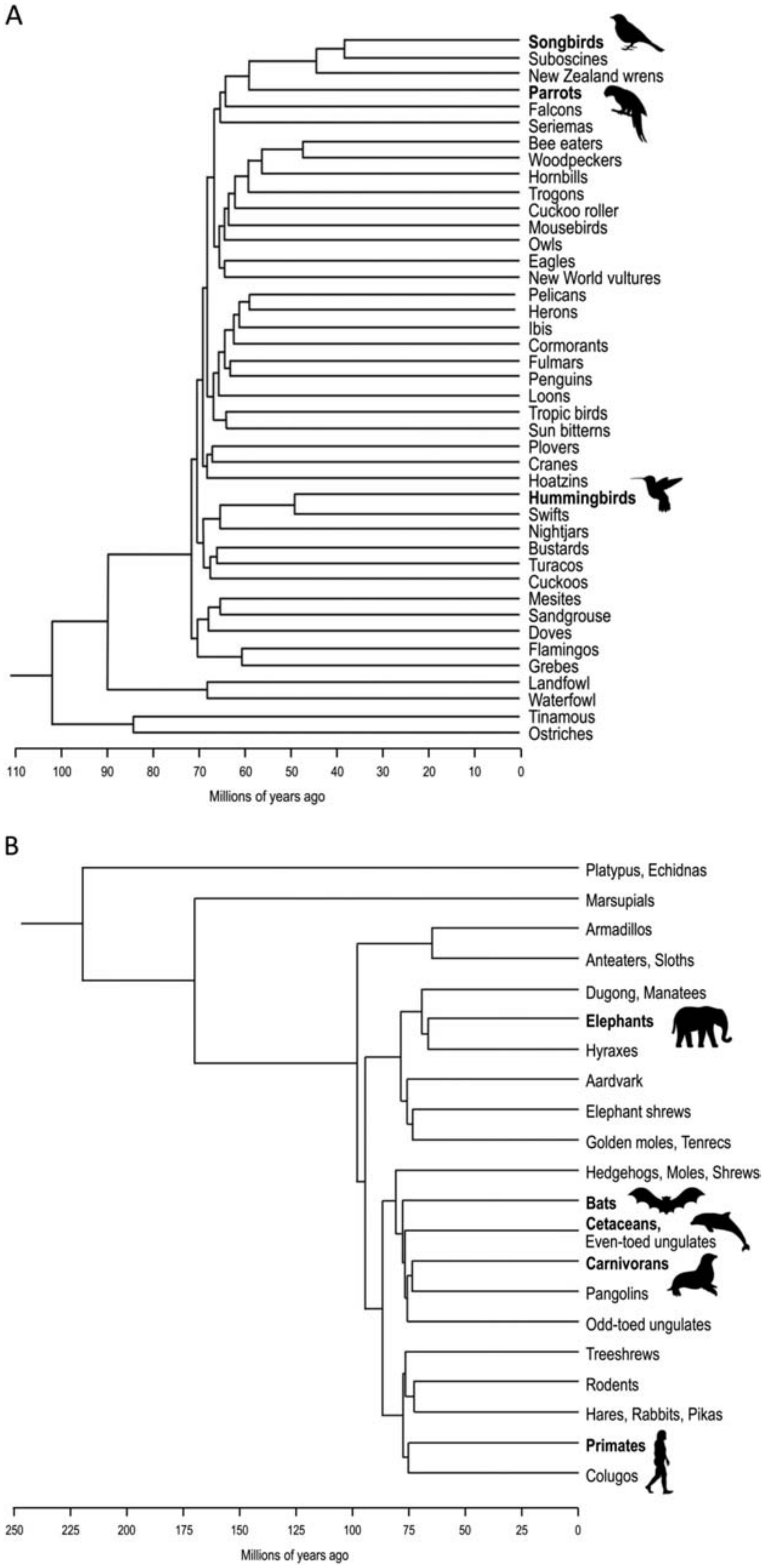

Figure 46.2 Phylogenetic tree of extant avian and mammalian orders highlighting the occurrence of VPL. (A) The avian tree has three orders capable of extensive VPL (designated by bold), namely songbirds, parrots, and hummingbirds. The evidence for extensive VPL in suboscines is currently limited to bellbirds. (B) The mammalian tree has five orders, which have at least one species capable of extensive VPL, namely humans among primates, bats, pinnipeds, cetaceans, and elephants. Four additional taxa (murids, apes, marmosets, goats) have at least one species with limited VPL capacities. Trees modified from Jarvis et al. (2014) and Foley et al. (2011).

Table 46.2

Selected cases of avian VPL

English name	Latin name	Family	Order	VPL type	Vocalization type	Description	References
Zebra finch	*Taeniopygia guttata*	Estrildidae	Passeriformes (oscines)	Imitation	Male song	Slightly imprecise copying of adult tutor song (normally father's or, rarer, sibling's song) results in an individual song type per male; song learning possible in adulthood when no suitable tutor was available before; species-specific song type can be established de novo over several generations	Immelmann (1969); Böhner (1983, 1990), Slater et al. (1988), Tchernichovski et al. (2001); Derégnaucourt and Gahr (2013); Eales (1985), Morrison and Nottebohm (1993), Jones et al. (1996); Fehér et al. (2009)
Canary	*Serinus canaria*	Fringillidae	Passeriformes (oscines)	Imitation	Male song	Male song is copied from adult tutors; open-ended learning; improvisation in the absence of tutors is also possible	Poulsen (1959), Waser and Marler (1977); Nottebohm & Nottebohm (1978); Metfessel (1935)
White-crowned sparrow	*Zonotrichia leucophrys*	Emberizidae	Passeriformes (oscines)	Imitation	Male song	Improvisation and precise imitation of adult tutor songs, regional dialects; song overproduction during development is followed by selective attrition	Marler and Tamura (1962), Marler (1970), Baptista (1977); Nelson et al. (2004); Nelson (2000)
Swamp sparrow	*Melospiza georgiana*	Emberizidae	Passeriformes (oscines)	Imitation	Male song	Male song is copied from adult conspecifics, but not heterospecifics; song overproduction during development is followed by selective attrition	Marler and Peters (1977, 1988); Marler and Peters (1982)
Song sparrow	*Melospiza melodia*	Emberizidae	Passeriformes (oscines)	Imitation	Male song	Male song is copied from adult conspecific tutors (and in captivity also from heterospecific tutors); young males copy song types from their neighbors	Marler and Peters (1987); Beecher et al. (1994, 1996), Nordby et al. (2001)

Chipping sparrow	*Spizella passerina*	Emberizidae	Passeriformes (oscines)	Imitation	Male song	Songs are learned by imitating a single adult tutor (neighbor after dispersal)	Liu and Kroodsma (1999, 2006)
Savannah sparrow	*Passerculus sandwichensis*	Emberizidae	Passeriformes (oscines)	Imitation	Male song	Specific song elements are copied from different adult tutors and combined into one song type per male; regional dialects; long-term cultural changes affect each song element differently	Wheelwright et al. (2008); Chew (1981); Williams et al. (2013)
Sedge warbler	*Acrocephalus schoenobaenus*	Acrocephalidae	Passeriformes (oscines)	Imitation	Male song	Improvisation and precise imitation of adult tutor songs, open-ended learning, neighboring males share song types; normal song development when raised in acoustic isolation	Nicholson et al. (2007); Leitner et al. 2002
Nightingale	*Luscinia megarhynchos*	Muscicapidae	Passeriformes (oscines)	Imitation	Male song	Improvisation and precise imitation of adult tutor songs; open-ended learning; neighboring males share song types but also invent new ones	Todt et al. (1979); Todt and Geberzahn (2003); Hutsch and Todt (1981), Sprau and Mundry (2010)
Brown-headed cowbird	*Molothrus ater*	Icteridae	Passeriformes (oscines)	Imitation	Male song	Male song is copied from adult tutors; improvisation instead of precise imitation in the presence of female conspecifics	King and West (1989)
European starling	*Sturnus vulgaris*	Sturnidae	Passeriformes (oscines)	Mimicry	Songs/calls from sympatric birds, artificial sounds	Mimicry of sympatric bird species and artificial sounds	Hindmarsh 1986)
				Imitation	Male song	Male song is copied from adult tutors; open-ended learning	Eens et al. (1992); Chaiken et al. (1993)

(*continued*)

Table 46.2 (*continued*)
Selected cases of avian VPL

English name	Latin name	Family	Order	VPL type	Vocalization type	Description	References
Common hill mynah	*Gracula religiosa*	Sturnidae	Passeriformes (oscines)	Mmimicry	Human speech (English)	Mimicry of human speech, close association with humans	Grosslight et al. (1962), Ginsburg (1963)
				Imitation	Whole call repertoire	Calls are learned by imitating adult neighbors (both sexes); no heterospecific mimicry in the wild	Bertram (1970)
Superb lyrebird	*Menura novaehollandiae*	Menuridae	Passeriformes (oscines)	Mimicry	Songs/calls from sympatric birds, artificial sounds	Mimicry of sympatric bird species and artificial sounds	Zann and Dunstan (2008)
Bare-throated bellbird	*Procnias nudicollis*	Cotingidae	Passeriformes (suboscines)	Mimicry	Calls of chopi blackbird	Mimicry of heterospecific calls, close association with a chopi blackbird (*Gnorimopsar chopi*)	Kroodsma et al. (2013)
Three-wattled bellbird	*Procnias tricarunculata*	Cotingidae	Passeriformes (suboscines)	Imitation	Male song	Regional dialect	Saranathan et al. (2007), Kroodsma et al. (2013)
African gray parrot	*Psittacus erithacus*	Psittacidae	Psittaciformes	Mimicry (referential)	Human speech (English)	Referential use of English words as a result of intensive training	Pepperberg (1981, 1994, 2006, 2010)
				Mimicry	Human speech (German); songs/calls from nine sympatric birds and one bat	Mimicry of human speech, close association with humans; mimicry of sympatric species	Todt (1975), Cruickshank et al. (1993)
Budgerigar	*Melopsittacus undulatus*	Psittacidae	Psittaciformes	Imitation	Contact call	Flock mates imitate each others' calls; males imitate female contact calls after pair bonding	Farabaugh et al. (1994), Hile & Striedter (2000); Hile et al. (2000, 2005)

Yellow-naped amazon	*Amazona auropalliata*	Psittacidae	Psittaciformes	Imitation	Contact call	Regional dialects; imitation of local call types by immigrant birds after dispersal	Wright (1996), Wright et al. (2005, 2008), Salinas-Melgoza and Wright (2012)
Orange-fronted conure/parakeet	*Aratinga canicularis*	Psittacidae	Psittaciformes	Imitation	Contact call	Regional dialect; vocal convergence or divergence prior to flock fusions	Bradbury et al. (2001); Balsby and Scarl (2008), Balsby and Bradbury (2009)
Green-rumped parrotlet	*Forpus passerinus*	Psittacidae	Psittaciformes	Imitation	Contact call	Individual signature character of nestlings' contact calls is learned from both parents	Berg et al. (2012)
Anna's hummingbird	*Calypte anna*	Trochilidae	Apodiformes	Imitation	Male song	Regional dialect; males raised in isolation develop abnormal song	Baptista and Schuchmann (1990)
Sparkling violetear	*Colibri coruscans*	Trochilidae	Apodiformes	Imitation	Male song	Regional dialect	Gaunt et al. (1994)
Green violetear	*Colibri thalassinus*	Trochilidae	Apodiformes	Imitation	Male song	Regional dialect	Gaunt et al. (1994)
Little hermit	*Phaethornis longuemareus*	Trochilidae	Apodiformes	Social modification and imitation	Male song	Regional dialect; open-ended song type learning	Snow (1968), Wiley (1971); Araya-Salas and Wright (2013)

Table 46.3

Occurrence of mammalian VPL

English name	Latin name	Family	Order	VPL type	Vocalization type	Description	References
Common bottlenose dolphin	*Tursiops truncatus*	Delphinidae	Cetacea	Mimicry	Artificial sounds	Referential use of learned acoustic signals (artificial)	Richards et al. (1984), Reiss and McCowan (1993)
				Imitation	Cohesion calls (signature whistles)	Copying of conspecific whistles	Janik (2000), King et al. (2013), King and Janik (2013)
				Social modification	Cohesion calls (signature whistles)	Whistles of young dolphins converge toward whistles of group members	Fripp et al. (2005)
Killer whale	*Orcinus orca*	Delphinidae	Cetacea	Mimicry	Barks of California sea lions	Mimicry of sea lion barks, close association with California sea lions	Foote et al. (2006)
				Mimicry	Human speech (English)	Mimicry of human speech, close association with humans	Abramson et al. (2018)
				Social modification	Different social calls	Regional dialects	Deecke et al. (2000), Weiß et al. (2011), Kremers et al. (2012), Filatova et al. (2012)
Risso's dolphin	*Grampus griseus*	Delphinidae	Cetacea	Mimicry	Whistles	Mimicry of bottlenose dolphin whistles, close association with bottlenose dolphin	Favaro et al. (2016)
Beluga	*Delphinapterus leucas*	Monodontidae	Cetacea	Mimicry	Human speech (English)	Mimicry of human speech, close association with humans	Eaton (1979), Ridgway et al. (2012), Murayama et al. (2014)
Humpback whale	*Megaptera novaeangliae*	Balaenopteridae	Cetacea	Imitation	Male song	Regional dialects, culturally induced change in song motifs	Noad et al. (2000), Garland et al. (2011)
Harbor seal	*Phoca vitulina*	Phocidae	Carnivora	Mimicry	Human speech (English)	Mimicry of human speech, close association with humans	Ralls et al. (1985)
				Invention	Novel sounds	Novel sound production through reinforcement training	Schustermann (2008)

Northern elephant seal	*Mirounga angustirostris*	Phocidae	Carnivora	Social modification	Male agonistic vocalizations	Regional dialects	Le Boeuf and Peterson (1969), Le Boeuf and Petrinovich (1974)
Southern elephant seal	*Mirounga leonina*	Phocidae	Carnivora	Social modification or imitation	Male agonistic vocalizations	Young adults' vocalizations closely resemble the ones of dominant males present during subadult period of development	Sanvito et al. (2007)
Walrus	*Odobenus rosmarus*	Odobenidae	Carnivora	Invention	Novel sounds	Novel sound production through reinforcement training	Schustermann and Reichmuth (2008)
Egyptian fruit bat	*Rousettus aegyptiacus*	Pteropodidae	Chiroptera	Social modification	Different social calls	Immature vocal repertoire when raised in acoustic isolation; frequency shift to conspecifics' calls	Prat et al. (2015), Prat et al. (2017)
Greater sac-winged bat	*Saccopteryx bilineata*	Emballonuridae	Chiroptera	Imitation	Male territorial songs	Juveniles imitate adult tutor song	Knörnschild et al. (2010)
				Social modification	Juvenile isolation calls	Juvenile isolation calls converge toward fellow juveniles' calls	Knörnschild et al. (2012)
Greater spear-nosed bat	*Phyllostomus hastatus*	Phyllostomidae	Chiroptera	Social modification	Short-distance contact calls (screeches)	Screeches of unrelated group members converge	Boughman (1998)
Pale spear-nosed bat	*Phyllostomus discolor*	Phyllostomidae	Chiroptera	Social modification	Juvenile isolation calls	Isolation calls converge toward maternal directive calls	Esser and Schmidt (1989), Esser (1994)
Greater horseshoe bat	*Rhinolophus ferrumequinum*	Rhinolophidae	Chiroptera	Social modification	Echolocation calls	Resting frequency of juvenile echolocation calls converges toward maternal echolocation calls	Jones and Ransome (1993)
Asian elephant	*Elephas maximus*	Elephantidae	Proboscidea	Mimicry	Human speech (Korean)	Mimicry of human speech, close association with humans	Stoeger et al. (2012)
African elephant	*Loxodonta africana*	Elephantidae	Proboscidea	Mimicry	Artificial sound, chirps of Asian elephants	Mimicry of Asian elephant chirps, close association with Asian elephants	Poole et al. (2005)

(*continued*)

Table 46.3 (*continued*)
Occurrence of mammalian VPL

English name	Latin name	Family	Order	VPL type	Vocalization type	Description	References
Common chimpanzee	*Pan troglodytes*	Hominidae	Primates	Social modification (subtle)	Long-distance call (pant hoot)	Adult males modify pant-hoots to be different from their neighbors; novel pant-hoot variant invented by one male spread in its social group	Crockford et al. (2004); Marshall et al. (1999)
				Social modification (subtle)	Food grunt	Call convergence of group members once affiliative social relationships are formed	Watson et al. (2015)
Pygmy marmoset	*Cebuella pygmaea*	Callitrichidae	Primates	Social modification (subtle)	Short-distance contact call (trill)	Convergence in trill structure between adult pair members	Snowdon and Elowson (1999)
Common marmoset	*Callithrix jacchus*	Callitrichidae	Primates	Social modification	Long-distance contact call (phee call)	Contingent vocal responses from parents influence the timing of the transition from infant cries to adult-like phee calls by reinforcing the production of phee calls	Takahashi et al. (2015, 2017)
				Social modification	Different social calls	Immature vocal repertoire when raised in acoustic isolation	Gultekin and Hage (2017)
Domestic goat	*Capra hircus*	Bovidae	Artiodactyla	Social modification (subtle)	Juvenile contact calls	Juvenile contact calls converge toward fellow juveniles' calls	Briefer and McElligott (2012)
Mouse	*Mus musculus (B6 strain)*	Muridae	Rodentia	Social modification (subtle)	Courtship vocalizations	B6 males decrease pitch toward the range of BxD males when housed under competitive social conditions	Arriaga et al. (2012)

of vocalizations. The filter domain of vocal production (also called supralaryngeal/-syringeal articulatory system) is associated with the vocal tract and shapes the dispersion and transition of formants in vocalizations. As vocal tract resonances, formants are concentrations of energy at certain frequencies and thus shape the spectrum of a vocalization. Formants constitute the primary phonetic cue in speech (Lieberman & Blumstein, 1988) and animals can perceive them in con- and heterospecific vocalizations (Fitch, 1997; Fitch & Kelley, 2000; Ohms, Escudero, Lammers, & ten Cate, 2012; Kriengwatana & Beckers, chapter 48 of this volume; Townsend & Manser, 2011). Changing temporal characteristics (duration, interval) is deemed to be easier than changing spectral parameters (pitch, formants), which is why the respiratory domain of VPL is considered to be less cognitively demanding than the phonatory or filter domain of VPL (Boughman & Moss, 2003; Janik & Slater, 1997, 2000). However, these three domains influence one another; rapid amplitude modulations by the respiratory domain, for instance, cause additional harmonics in a vocalization and therefore interact also with the phonatory domain (Fitch, 2006; Fitch, Neubauer, & Herzel, 2002; Wilden, Herzel, Peters, & Tembrock, 1998). The sequences in which the individual vocalization syllables/calls are produced are thought to be controlled by pattern generators in the brain (Mackevicius & Fee, 2017). Complex vocalizations require sophisticated coordination between respiratory, laryngeal/syringeal, and articulatory muscles, and thus involve all domains of production (Beckers, 2011; Fitch, 2000, 2006; Janik & Slater, 2000).

1.4. Diversity in Development of Vocal Learning Strategies The development of extensive VPL has been studied in much greater depth in songbirds and humans than in any other species capable of VPL (Brainard & Doupe, 2002; Catchpole & Slater, 1995; Konishi, 1985; Kroodsma & Miller, 1996). Like human infants, the young of many songbirds listen to conspecific vocalizations and memorize them (sensory or perception phase) before they start imitating them during a babbling phase (sensorimotor or production phase (Brainard & Doupe, 2002; Doupe & Kuhl, 1999; Wilbrecht & Nottebohm, 2003). In songbirds, the initial rehearsal periods (subsong and plastic song) can commence weeks or months after the sensory phase, depending on the species. To fully master adult song, different species need different amounts of time, from a few months to more than a year. In some species, for example, zebra finches, the sensory phase, and the sensory-motor phase overlap completely (Roper & Zann, 2006; Slater, Eales, & Clayton, 1988). Like humans, some avian species are capable of extensive VPL throughout their lives, for example, canaries (Nottebohm, Nottebohm, & Crane, 1986), starlings (Mountjoy & Lemon, 1995), and many parrots (Bradbury & Balsby, 2016; Wright & Dahlin, 2017), whereas VPL only occurs in a sensitive/critical period during development in other species, for example, the white-crowned sparrow (Marler, 1970), zebra finch (Slater et al., 1988), and song sparrow (Beecher, 2017). Diversity also exists in song repertoire size (small vs. large), the accuracy of learning (precise imitation vs. improvisation), the importance of early exposure (development of species-typical song only after early exposure vs. even when raised in isolation), and learning flexibility (imitation within vs. outside species-specific constraints). This diversity of song-learning strategies, both within the same species (e.g., zebra finches: Liu, Gardner, & Nottebohm, 2004) and between different species (Beecher & Brenowitz, 2005; Brenowitz & Beecher, 2005) is an opportunity to understand the evolution and mechanisms of VPL and spoken language. For different aspects of language learning, there is likely to be a suitable songbird species to study.

1.5. Selection for Vocal Learning Avian and mammalian vocalizations are under strong sexual and natural selection, depending on the vocalization types and the sex involved (Catchpole & Slater, 1995; Garland et al., 2011; Gaunt et al., 1994; Knörnschild, 2014; Price, 2015; Sanvito, Galimberti, & Miller, 2007; Sjare, Stirling, & Spencer, 2003). Sexually selected songs and calls facilitate mate attraction and rival deterrence, functioning as advertisement and territorial signals, respectively. VPL can serve to increase song/call complexity, song/call type sharing with neighbors, and local dialects (Janik & Slater, 1997, 2000).

The function of learned vocalizations under natural selection pressures often concerns recognition processes on the individual, group, or population level and facilitates pair/group cohesion, alliance maintenance, cooperation, and mother-offspring reunion (Bradbury & Balsby, 2016; Boughman, 1998; Farabaugh et al., 1994; Janik & Slater, 1997, 2000; King, Sayigh, Wells, Fellner, & Janik, 2013; Nowicki, 1989). Whether VPL in humans is still relevant to sexual and natural selection is a fascinating question from a behavioral ecologist perspective (Lange, Henninghausen, Bril, & Schwab, 2016; Miller, 2013). Behavioral selection for the VPL trait is expected to be associated with molecular selection of genetic changes required for the trait and is the subject of section 2.

2. Comparative Neurogenetics

Many misconceptions exist about the relationship among genes, brains, and language. One of those concerns the assumption that because human language is unique, the genes coding for the proteins required to build a "language-ready" brain should also exist solely in humans, and that such brain regions are also unique in humans. The evidence to date has not supported this view. Novel traits can evolve even when the coding sequence of genes remains the same but their regulation changes during the course of evolution (Carroll, 2005; Romero, Ruvinsky, & Gilad, 2012). As a result, the protein may be the same, but where and when it is expressed, and how much there is of it, can vary between species and result in different phenotypes. In addition, many genes that already existed in the primate lineage before humans emerged are extant in humans. Slightly changed versions of genes can lead to altered cellular functions (Castellano et al., 2014). Some of these alterations might also have contributed toward new, human-specific traits. Finally, a small number of genes are truly human unique, for instance as a result of incomplete duplication and fusion with another gene (Dougherty et al., 2017). Together, these changes probably synergized in developing speech and language in the hominin lineage.

Many of the aspects of language mentioned have to come together for it to function. Speech requires the concerted action of respiration, mouth and tongue muscles, and the language regions of the brain, all of which are affected by the activity of genes. However, many of those genes are also involved in processes and behaviors unrelated to language. The same genes that are necessary for development of the brain circuits involved in language perception or production can also contribute to the development of brain regions that process sounds other than language or that control movements of the mouth required for eating. As a result, there are likely many genes required for language, and none thus far have been found to be exclusively associated with it (see section 2.3.2). Of particular interest are genes required for the specialized neural mechanisms that subserve our elaborate form of vocal communication, allowing us to externalize our thoughts through sound. To understand the neurogenetics behind language, we must also understand brain mechanisms of language.

2.1. Brain Pathways for Vocal Learning and Spoken Language Consistent with the widespread auditory comprehension learning, all vertebrate species examined to date have an auditory pathway that reaches the forebrain and is thought to be involved in auditory comprehension learning. This pathway begins at the sensory hair cells inside the ear and connects through the brain stem to the forebrain where it forms a network with the auditory cortex and other connected regions (figure 46.3).

Species with extensive VPL possess a specialized forebrain pathway that controls the learning and production of vocalizations (Jarvis, 2004; Petkov & Jarvis, 2012). This VPL pathway has been studied in songbirds and humans. It consists of an anterior component through the frontal cortex, basal ganglia, and thalamus that is required for vocal imitation, and a posterior motor cortex component that is required to produce learned vocalizations (Jürgens, 2002, 2009; Mackevicius & Fee, 2017; Scharff & Nottebohm, 1991; figure 46.3). The VPL pathway is embedded within a motor pathway that controls nonvocal behaviors, and both VPL and motor pathways receive auditory and other sensory input (Belyk & Brown, 2017; Chakraborty & Jarvis, 2015; Feenders et al., 2008). One specialized feature of the posterior component of the VPL pathway is that cortical neurons in layer 5 and their equivalent neurons in the robust nucleus of the arcopallium (RA) in songbirds project directly onto brain stem vocal motor neurons (figure 46.3). Direct projections from the cortex to brain stem motor neurons correlate with greater fine motor control of the associated muscles that might be necessary for human speech and learned birdsong.

In most species with limited or no VPL, a VPL pathway has not been found, with the possible exception of nonhuman primates, which have been proposed to have a laryngeal motor cortex (LMC) and rudimentary Broca's area (Jürgens, 2009; Simonyan, 2014). Jürgens (2002, 2009) proposed that the main difference between humans and nonhuman primates is the direct human projection from LMC to the vocal motor neurons, as opposed to the indirect connection in nonhuman primates. In contrast, Rilling and colleagues considered the main difference to be the direct projection from the higher auditory cortex (Wernicke's area) to speech-language cortex (Broca's area), which is either indirect or absent in monkeys or weak in great apes (Rilling, 2014; Rilling et al., 2008). Since mice also have a putative LMC region that directly projects, although very sparsely, to brain stem vocal motor neurons, and it receives a direct robust projection from secondary auditory cortex (Arriaga et al., 2012), it seems that this type of connectivity is not a hallmark exclusively associated with extensive VPL. Like in humans and song-learning birds, the putative LMC in mice consists of motor and premotor parts and forms a loop through the basal ganglia and thalamus. However, unlike humans and song-learning birds, in the mouse, the direct projection from

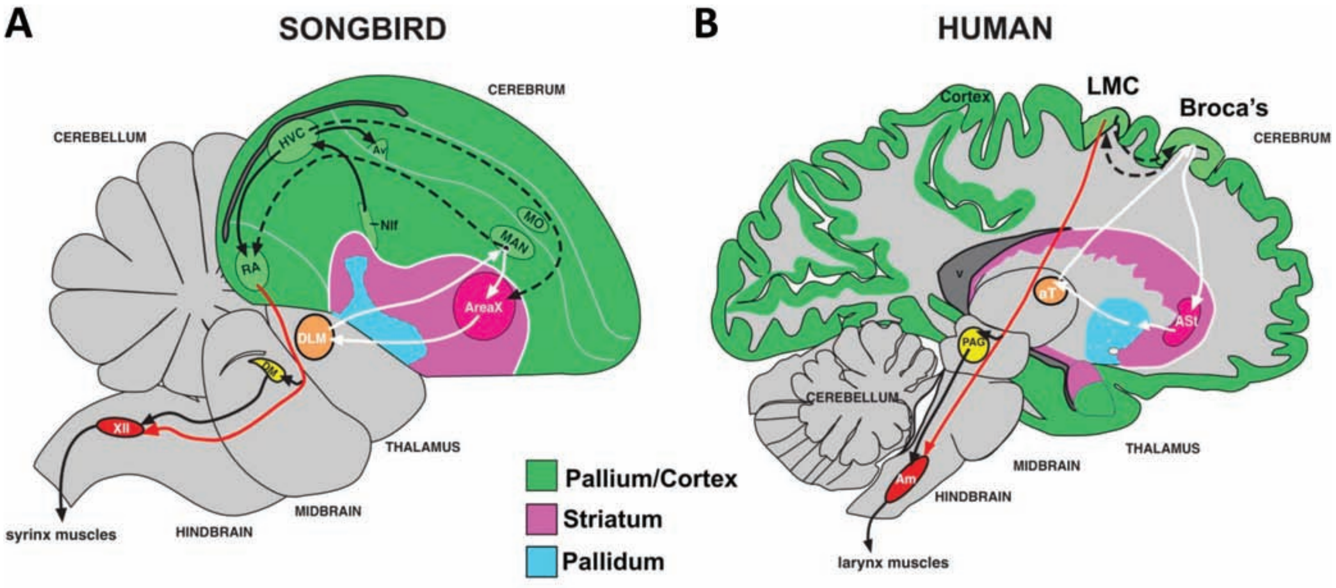

Figure 46.3 Brain pathways for vocal learning and spoken language in songbirds and humans. (A) Drawing of a songbird brain section showing connectivity of posterior (a vocal nucleus [HVC], RA, avalanche [Av], interfacial nucleus of the nidopallium [NIf]) and anterior (MAN, Area X, oval nucleus of the anterior mesopallium [MO]) song pathways. (B) Drawing of a human brain section showing proposed vocal pathway connectivity including LMC and part of anterior striatum (ASt) that shows convergence with songbird RA and Area X (Pfenning et al., 2014). Solid black arrows indicate the connections and regions of the posterior vocal motor pathway; white arrows indicate the connections and regions of the anterior vocal pathway. Dashed black arrows indicate the connections between the two pathways. Red arrows indicate the dense direct projection found only in vocal learners, from vocal motor cortex regions to brain stem vocal motor neurons. Am = nucleus ambiguous; aT = anterior thalamus speech area; DLM = dorsolateral nucleus of the thalamus; DM = dorsal medial nucleus of the midbrain; ; PAG = peri-aqueductal gray; v = ventricle space; XII = 12th vocal motor neurons in birds. Figure from Wang et al. (2015), modified from Arriaga et al. (2012); Petkov and Jarvis (2012); Pfenning et al. (2014).

LMC to vocal motor neurons is very sparse and lacks the specialized gene regulation found in language-relevant human brain regions and their counterparts in song-learning birds (discussed further in section 2.4). Moreover, the mouse LMC modulates vocalizations (Arriaga et al., 2012), but it is not required for their production (Arriaga et al., 2012; Hammerschmidt, Whelan, Eichele, & Fischer, 2015). These findings in primates, mice, and similar findings in a suboscine bird species (Liu, Wada, Jarvis, & Nottebohm, 2013) are consistent with the continuum hypothesis of VPL.

A synthesis of the anatomical and behavioral findings along a continuum of no VPL to extensive VPL is provided by the motor theory of vocal learning origin (Feenders et al., 2008). This theory proposes that in the common ancestor of all vertebrates, a forebrain pathway mediating general motor learning was present but that VPL was absent. The general motor learning pathway was duplicated and the new pathway, still embedded in the evolutionary older one, formed connections with the brain stem vocal motor system, facilitating limited VPL. In extensive VPLers, the new circuit segregated anatomically and functionally from the general motor learning pathway became more specialized and developed robust direct projections controlling the brain stem vocal motor and respiratory systems. In parrots and humans, a second duplication event of the VPL circuit generated two parallel VPL circuits allowing for greater VPL complexity (Chakraborty & Jarvis, 2015). Direct input from the auditory system is proposed to have been already present in the general motor learning pathway before extensive VPL evolved, and the VPL pathway inherited that input, allowing for auditory-vocal motor integration. The trait of VPL and its anatomical underpinnings might be evolutionarily quite plastic, leading to the development and the disappearance of VPL multiple times. For instance, in the majority of songbirds VPL exists in both sexes. However, in songbird species where only males sing, like zebra finches, females apparently lost extensive VPL (Odom, Hall, Riebel, Omland, & Langmore, 2014). In line with this notion, female zebra finches retain a vestigial VPL circuit that can be activated by hormone treatments during development, leading to the development of a VPL pathway and extensive VPL ability (Gurney, 1982). It is within this context of a continuum of neural circuitry and behavior predating speech and the anatomy underlying it, that we discuss how we look for the associated genes.

2.2. GENE HUNTING STRATEGIES There are two main ways to identify genes associated with vocal learning and language (see also chapter 41 by Vernes in this volume). First, by searching for the mutation within an individual gene associated with a particular disorder affecting language. Such monogenic disorders are rare. Usually, gene networks, not a single gene, underlie complex traits including language. However, in some cases, a mutation in a single gene can have serious consequences for the production and comprehension of speech, as is the case for the *FOXP2* gene,[1] which we will discuss in section 2.3. In the second approach, one correlates variation in the language faculty with the variation that occurs in genes. Individuals' genomes differ at millions of sites by single nucleotides (SNPs). The totality of those SNP variations is an individual's genetic fingerprint. One can search for associations between specific SNPs and a particular language phenotype, be it an impairment or a talent (see Deriziotis & Fisher, 2017, and chapter 39 by Luciano & Bates in this volume, for more details on methods and findings). This approach can be extended by probing for genes required for specific components of language, such as VPL. The diversity of VPL types among species is a great resource to identify genes involved in different aspects of VPL. Next we discuss some of the most informative discoveries thus far, a single gene approach and multigenic approaches comparing species that exhibit obligate VPL with those that do not.

2.3. THE FOXP GENE FAMILY IN HUMANS AND NONHUMAN ANIMALS

2.3.1. Core phenotype of patients with FOXP2 mutations The first gene associated specifically (but not exclusively) with speech was *FOXP2* (Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001). FOXP2 codes for a "transcription factor," that is, a protein that regulates the transcription of many other genes, called *target genes* (Spiteri et al., 2007). Mutations in FOXP2 cause Developmental Verbal Dyspraxia (also referred to as Childhood Apraxia of Speech), a severe speech disorder (Vargha-Khadem, Gadian, Copp, & Mishkin, 2005). Patients perform poorly in tests of productive aspects of language. Perceptive aspects of language and learning of other motor tasks are less affected (Morgan, Fisher, Scheffer, & Hildebrand, 2017). FOXP2 mutations lead to altered structure and function of corticostriatal and corticocerebellar circuits (Liégeois et al., 2016).

2.3.2. FOXP is evolutionary conserved The association of *FOXP2* with speech led to speculations that it might be "the language gene," unique to humans and the magic bullet to understand language evolution. However, *specifically associated with some aspects of language* does not need to translate to *unique to humans*. In fact, the coding sequence is highly conserved among most vertebrates (Li, Wang, Rossiter, Jones, & Zhang, 2007; Scharff & Haesler, 2005; Zhang, Webb, & Podlaha, 2002). In bats (Li et al., 2007) and teleost fish (Song, Wang, & Tang, 2013), the FoxP2 sequences are also conserved but diverged substantially more during the course of evolution. Importantly, FoxP2 is expressed in many brain regions that are relevant for speech in humans and for VPL in songbirds. Yet, FoxP2 expression is not limited to those neuron populations but it is also expressed in neurons relevant for other behaviors (Geerling et al., 2016; Haesler et al., 2004; Teramitsu, Kudo, London, Geshwind, & White, 2004; Verstegen, Vanderhorst, Gray, Zeidel, & Geerling, 2017). Its expression has been mapped in human embryos (Lai, Gerrelli, Monaco, Fisher, & Copp, 2003; Teramitsu et al., 2004), mice (Ferland, Cherry, Preware, Morrisey, & Walsh, 2003), songbirds (Haesler et al., 2004; Vicario, Mendoza, Abellan, Scharff, & Medina, 2017), fish (Bonkowsky & Chien, 2005), bats (Rodenas-Cuadrado et al., 2018), and even in drosophila and honeybees (DasGupta, Ferreira, & Miesenböck, 2014; Lawton, Wassmer & Deitcher, 2014; Schatton & Scharff, 2017). The fact that FoxP2 is expressed in similar brain circuits in many vertebrates suggests that this transcription factor fulfills important and potentially similar functions across a large variety of animals (Wohlgemuth, Adam, & Scharff, 2014). FoxP2 expression in invertebrates and vertebrates commences during embryogenesis and marks specific neuron populations, consistent with its demonstrated role in neuronal differentiation (Chiu et al., 2014).

2.3.3. FoxP2 manipulations in animal models affect brain development and vocalizations In the developing mouse brain, experimental manipulations of Foxp2 cause abnormal cortical neurogenesis (Tsui, Vessey, Tomita, Kaplan, & Miller, 2013), cerebellar foliation (Groszer et al., 2008), and spinal cord motor neuron delamination (Rousso et al., 2012). FoxP2 continues to function in the adult nervous system, shown by gene function studies of *FoxP2* in songbirds and mice (Adam, Mendoza, Kobalz, Wohlgemuth, & Scharff, 2016; Castellucci, McGinley, & McCormick, 2016; Chabout et al., 2016; Fisher & Scharff, 2009; Gaub, Fisher, & Ehret, 2016; Haesler et al., 2007; Heston & White, 2015; Murugan, Harward, Scharff, & Mooney, 2013; Schreiweis et al., 2014). In zebra finches, during the song-learning phase, both too little and too much FoxP2 in the striatal song nucleus Area X, an anterior forebrain component of the VPL pathway required for song learning (figure 46.3), prevents birds from developing normal song (Haesler et al., 2007; Murugan et al., 2013; Heston & White, 2015). After virus-mediated downregulation in Area X, juveniles developed abnormal song bearing phenotypic parallels to speech features of patients with FOXP2 mutations (Haesler et al., 2007), namely a smaller number of elements, shorter phrases, and inaccurate, variable delivery (Watkins, Dronkers, & Vargha-Khadem, 2002). In mice, a heterozygous Foxp2 missense mutation (Foxp2-R552H $^{+/-}$) that is equivalent to a mutation that causes speech impairments in humans affects their ability to produce complex sequences of sound units when courting a female, but does not affect simple sequences or their acoustic structure (Castellucci et al., 2016; Chabout et al., 2016; Gaub et al., 2016). Furthermore, the layer 5 LMC neurons in these mice are not properly coalesced in the LMC location (Chabout et al., 2016).

Comparing the findings in zebra finches and lab mice with those of human patients are intriguing. In humans carrying a heterozygous FOXP2-R553H point mutation, the speech deficit affects both the acoustic features of individual speech sounds and the sequencing of sound units that make up multisyllable words (Watkins et al., 2002). In zebra finches with reduced levels of FoxP2 in Area X, the acoustic structure of individual song elements was more impaired than their sequential delivery (Haesler et al., 2007). In Foxp2-R552H $^{+/-}$ mice, the acoustic features of the individual sound units are not affected but complex sequencing is. Thus, in species with extensive VPL, humans, and songbirds, it appears that FoxP2 influences circuit control of the acoustic structure of vocalizations, whereas in species with no or limited VPL, it does not play a role in acoustic structure.

2.3.4. FoxP2 affects striatal function in humans and animal models When comparing these results one should bear in mind that in mice and humans, the FoxP2 mutation was in the genome, impacting all cells in the body that express the gene. In contrast, in songbirds, the experimental manipulations were limited to Area X, the brain region important for imitative song learning, thus directly linking the relevance of FoxP for VPL to striatal activity. Follow-up studies showed that the FoxP2 downregulation in Area X alters the spine density of the striatal spiny neurons and glutamatergic transmission, impacting Dopamine 1 Receptor (D1R)–mediated modulation of signal propagation through Area X (Adam et al., 2016; Murugan et al., 2013; Schulz, Haesler, Scharff, & Rochefort, 2010). Concomitantly, D1R and DARPP-32 protein levels in Area X were reduced. In addition, social context-dependent song variability and the associated context-dependent neural variability in the cortical nucleus (lateral magnocellular nucleus of the anterior nidopallium or LMAN) that projects to Area X (figure 46.3) were abolished. Further support for the relevance of FoxP2 in striatal function is provided by in vivo electrophysiological recordings in the striatum of awake-behaving mice that were genetically modified to carry the same mutation occurring in human FOXP2 patients (French et al., 2012). Likewise, the striatum in FOXP2 patients is structurally and functionally affected (Liégeois et al., 2003).

2.3.5. FoxP2 levels can be dynamically regulated by behavior FoxP2 is transiently upregulated in Area X of young zebra finches during the developmental song-learning period and of adult canaries during seasonal song plasticity (Haesler et al., 2004; Teramitsu, Poopatanapong, Torrisi, & White, 2010; Teramitsu & White, 2006). In addition, singing behavior in young and adult zebra finches can acutely downregulate FoxP2 expression in Area X (Miller et al., 2008; Teramitsu et al., 2010; Thompson et al., 2013); intriguingly, in budgerigars (a parrot), FoxP2 is constitutively low in the region equivalent to Area X and does not decrease further with production of learned calls (learned song was not tested) (Hara et al., 2015), possibly related to the life-long vocal plasticity in budgerigars (Farabaugh et al., 1994; Hile, Plummer, & Striedter, 2000). In mice, changes in Foxp2 expression after vocalizing were not addressed, but FoxP2 is upregulated by sound exposure in the auditory thalamic (medial geniculate) nucleus (Horng et al., 2009).

2.3.6. FoxP2 modulates target genes relevant for synaptic function Hundreds of *FoxP2* target genes were discovered in experiments with human fetal brain, human neuron-like cells, and developing mouse brain (Vernes et al., 2007; Vernes et al., 2008; Vernes et al., 2011; Vernes, chapter 41 of this volume). The expression of two of those target genes is positively correlated with FoxP2 expression levels in zebra finch Area X: the very low-density lipoprotein receptor (*VLDLR*), encoding one of the reelin receptors, and Contactin-associated protein-like 2 (*CNTNAP2*), encoding a neurexin, both are important for synaptic function (Rodenas-Cuadrado, Ho, & Vernes, 2014). Zebra finch FoxP2 binds to the promoters of *VLDLR* and *CNTNAP2* (Adam et al., 2016, 2017) and experimental downregulation of FoxP2 in Area X results in reduced expression of *VLDLR* and *CNTNAP2*. Further findings raise the possibility that the regulatory relationship between FoxP2 and VLDLR guides structural plasticity toward the subset of FoxP2-positive medium spiny neurons of Area X in a singing-dependent manner via the reelin pathway (Adam et al., 2016, 2017). This highlights the need to think about the regulation of FoxP2 (Shi et al., 2018) and its targets in an activity-dependent and cell-specific manner associated with moment-to-moment changes in behavior (Adam et al., 2016, 2017; Becker, Devanna, Fisher, & Vernes, 2018; Panaitof, Abrahams, Dong, Geschwind, & White, 2010).

2.3.7. FoxP2 interacts with FoxP1 and FoxP4 FoxP2 and its paralogs, FoxP1 and FoxP4, are co-expressed in Area X, with most medium spiny neurons expressing all three FoxPs, but all other combinations also exist, with different frequencies (Mendoza & Scharff, 2017). In human cell culture experiments, the combinatorial protein-protein interaction among FoxP1, FoxP2, and FoxP4 regulates the transcription of target genes differentially (Sin, Li, & Crawford, 2015). Likewise, in zebra finches, the three FoxPs regulate the CNTNAP2 promoter differentially: FoxP1 activates it, FoxP2 represses it, and FoxP4 neither binds to nor regulates it (Mendoza & Scharff, 2017). Together, these results emphasize the functional importance of the protein-protein interactions among the FoxP family members in regulating their target genes and predict an involvement of FoxP1 and FoxP4 in speech and language. Indeed, FOXP1 mutations also affect speech and language, in addition to a variety of other clinical symptoms (Meerschaut et al., 2017; Siper et al., 2017). A FOXP4 mutation has only been reported in one patient that was developmentally delayed and had larynx problems (Charng et al., 2016).

2.3.8. FoxP1 and FoxP2 expression levels can be sex-specific FOXP2 protein levels in the left-hemispheric Brodmann's area 44 (part of Broca's area) of four-year-old girls (detected by Western blot from postmortem tissue) are higher than those of age-matched boys (Bowers, Perez-Pouchoulen, Edwards, & McCarthy, 2013). This leads to the question of whether the FOXP2 haploinsufficiency (e.g., loss of 50% of functioning protein) in humans with FOXP2 mutations might affect males more than females. However, a phenotypic difference between the sexes in patients with FOXP2 mutations has not been reported so far (Watkins et al., 2002). Sex differences in FoxP1 or FoxP2 expression in Area X were not noted in budgerigars or strawberry finch songbirds (*Amandava amandava*) (Haesler et al. 2004; Hara et al., 2015), whereas the absolute density of FoxP2 expression in Area X neurons was higher in male than in female Black-Capped Chickadees (*Poecile atricapillus*), which is consistent with males being the more vocal sex in this species (Phillmore, MacGillivray, Wilson, & Martin, 2015). Yet, a sex difference in expression levels also exists in multiple brain regions of rat pups, albeit in the opposite direction from the one reported in humans, with higher levels in males than females. Moreover, male pups produce more ultrasound vocalizations when separated from their mother than female pups do, and experimental downregulation of Foxp2 levels in the males leads to more female-like calling (Bowers et al., 2013). Very similar findings were found for Foxp1 in mice, linking sexual dimorphic pup separation calls to cortical and subcortical Foxp1 expression levels (Fröhlich, Rafiullah, Schmitt, Abele, & Rappold, 2017). This suggests a mechanistic link between the sex specificity of call activity and sexually dimorphic Foxp2 expression levels in rodents (Bowers et al., 2013), with androgens as a mediator (Bowers, Perez-Pouchoulen, Roby, Ryan, & McCarthy, 2014; Hamson, Csupity, Gaspar, & Watson, 2009).

2.3.9. FoxP affects behavioral function in fruit flies During the transition from invertebrates to vertebrates, a single *FoxP* gene gave rise to the four vertebrate FoxP gene paralogs, as a result of two postulated genome duplication events (Santos, Athanasiadis, Leitao, DuPasquier, & Sucena, 2011; Song, Tang, & Wang, 2016). Given the similarity of the DNA-binding domain of the single *FoxP* gene in *Drosophila* and the four vertebrate paralogs prompted studies into the behavioral consequences of FoxP loss of function in fruit flies. *Drosophila* mutants (dFoxPS-SZ-3955) have altered levels of the corresponding messenger RNA compared to wild-type flies and are incapable of learning to fly away from a noxious stimulus if only somatosensory cues are available to them (Mendoza et al., 2014). In a different study, dFoxPS-SZ-3955 mutants were slower to translate decisions based on different odor concentrations into a turning-away movement (DasGupta et al., 2014). A third study using a different mutant found problems with courtship behavior (Lawton et al., 2014). These different behavioral phenotypes in FoxP fly mutants are consistent with the hypothesis that tightly regulated FoxP protein levels in particular neurons are needed to translate stored information (e.g., where the body is in space) into motor programs (Schatton & Scharff 2017). Whether these are superficial similarities to the behavioral effects of FoxP1 and FoxP2 manipulations in vertebrates or whether they are based on true "deep homology" of FoxP and its associated molecular modules requires further scrutiny.

2.3.10. The human version of FOXP2 affects brain development and behavior It seems clear from the preceding summary that FoxP2 was relevant for motor behaviors including vocal communication well before spoken language evolved. But human spoken language only evolved once. Does the fact that the human FOXP2 sequence differs from all other FoxP2 sequences provide a clue to the particular specialization of human speech and language? The human *FOXP2* gene differs from other primates in only two amino acids (Enard et al., 2002) and one of those also differs from all mammals investigated (Zhang et al., 2002). The two human amino acid substitutions (T303N, N325S) were apparently subject to positive selection and became fixed in the human population after the human split from the chimpanzee lineage. This led to the notion that T303N and N325S contributed to the evolution of human language. To address this, the human-specific amino acids were introduced to the mouse Foxp2 gene, Foxp2^{hum} (Enard et al., 2002). Foxp2^{hum} pups produce qualitatively different ultrasonic calls when separated from their mothers. As adults, these mice were less exploratory but better at learning specific labyrinth tasks than their "nonhumanized Foxp2" siblings. In the brain, the Foxp2^{hum} had lower dopamine concentrations and longer dendrites in the cerebral cortex, thalamus, and striatum and increased synaptic plasticity (Long Term Depression; LTD) in the striatal medium spiny neurons (Enard et al., 2009; Reimers-Kipping, Hevers, Pääbo, & Enard, 2011; Schreiweis et al., 2014). These findings underscore the fact that even though the human and the mouse FoxP2 protein differ in only 3 of over 700 amino acids, small changes have the potential to alter neural development and brain function for vocalizations and could have promoted human cognitive development during evolution, for example, enhanced synaptic plasticity and behavioral flexibility.

2.4. Genome- and Transcriptome-wide Changes in VPLers If VPL and its contribution to speech is a multigenic trait, more genes than *FOXP2* should be discoverable. This is supported by recent informative findings of RNA expression of thousands of genes, the transcriptome. Using oligo-microarrays representing the brain transcriptome revealed that the song nucleus RA in songbirds and its analog in parrots and hummingbirds have convergent gene expression specializations with each other and with the human LMC (Pfenning et al., 2014; figure 46.3). These gene expression specializations are not found in birds or nonhuman primates that have little or no VPL. A gene expression specialization is where a gene's RNA or protein product is increased or decreased relative to the surrounding brain region, in this case, the adjacent nonvocal avian motor arcopallium or mammalian primary motor cortex. Of a total of 7,000 genes examined, there are ~55 such specialized genes in RA and the LMC; many of which are related to neural connectivity and neural development. One of the candidate genes is *SLIT1*, a ligand for the ROBO1 receptor, which is important for axon guidance (Pfenning et al., 2014; Wang et al., 2015). Mutations in SLIT1 and ROBO1 are associated with language deficits and autism (Wang et al., 2015). Interestingly, *SLIT1* and other genes relevant for neural connectivity were downregulated in avian RA and human LMC. One idea why this might be so is that since SLIT1 is a repulsive molecule, high expression in species without VPL may prevent the direct projection from RA/LMC to brain stem vocal motor neurons typical for VPL from forming (Wang et al, 2015). This can be tested by experimentally downregulating SLIT1 in the motor cortex of limited VPL or upregulating it the cortical RA/LMC of extensive vocal learners. There is precedent for such a possibility. In the study of another repulsive axon guidance ligand, PlexinA1, recapitulating its downregulation in human motor cortex in the mouse brain allowed direct projections to spinal cord motor neurons to remain after infant development (Gu et al., 2017).

Brain regions with convergent specialized transcriptome expression are also songbird Area X and a portion of the anterior human striatum (across the boundary of the caudate-putamen) that is activated during speech production (Pfenning et al., 2014; figure 46.3). The over 70 genes identified with specialized expression are overrepresented in neurotransmission and movement disorder functions, consistent with the known function of Area X. These genes include the D2 dopamine receptor that is important for neurotransmission and synaptic plasticity in the striatum and ROBO1. The identified human striatal region overlaps a part of the striatum that is affected in people with a FOXP2 point mutation (Belton, Salmond, Watkins, Vargha-Khadem, & Gadian, 2003). Other avian song learning brain regions showed only weak trends of convergent expression with other human speech regions (e.g., songbird LMAN and human Broca's area), but these also did not have the immediate surrounding brain regions in birds profiled (Pfenning et al., 2014). Preliminary experiments profiling all the surrounding brain regions indicate that most if not all song learning nuclei of song-learning birds have convergent gene expression specializations with human spoken-language brain regions (Gedman, Pfenning, Wirthlin, Audet, & Jarvis, 2017).

One explanation, postulated in the motor theory of vocal learning, for the convergent molecular similarities is that the brain regions surrounding the specialized song-learning and spoken-language areas are homologous and served as a pre-existing substrate out of which similar brain pathways could evolve(Feenders et al., 2008). In this regard, the songbird RA shows an apparent homologous gene expression profile with mammalian motor cortex layer 5 neurons; a vocal nucleus called HVC, which projects to RA, shows molecular similarity to layers 2 and 3 (Pfenning et al., 2014). This finding supports Karten's 1969 hypothesis, originally proposed for the auditory and visual cortices, that different pallial populations of the avian brain are homologous to different cortical layers of the mammalian brain (Jarvis et al., 2005; Wang, Brzozowska-Precht, & Karten, 2010). Interestingly, the auditory regions of both song-learning birds and humans do not show as high a level of gene expression specialization relative to the surrounding sensory cortices, nor do they show convergent expression. Taken together, these findings are consistent with known species differences in the different vocal production learning types. They indicate that the most anatomically and molecularly specialized components of the avian song and human language systems are those involved in song and spoken-language acquisition and production.

In humans, additional molecular changes have occurred that affect the brain. This is the case in the Slit-Robo GTPase 2 gene (*SRGAP2*), which interacts with the speech relevant ROBO1 receptor and modulates its function. In humans, SRGAP2 is partially duplicated, and the duplicated copies act as competitive inhibitors of the parent gene (Charrier et al., 2012; Dennis et al., 2012). Partial inhibition of SRGAP2 and thus of ROBO1 receptor function causes dendrites of cortical neurons in humans to remain in a more immature state throughout life, similar to the finding of Foxp2$^{\mathrm{hum}}$ mice. The immature state allows for continued higher levels of learning into adulthood compared to learning in other species without such duplication. Putting studies together, one can imagine how the consequences of specialized expression of SLIT1 and ROBO1, regulated by FoxP2, in some spoken-language VPL brain regions, is in turn enhanced by the SRGAP2 duplication in humans.

The specialized gene regulation in speech brain regions is presumably controlled by changes in regulatory regions of those genes or in the coding sequences of genes that regulate them. Studies have looked for genome-wide sequence changes in VPL (Zhang et al., 2014) as well as in humans with speech disorders (Chen et al., 2017). Such investigations are subject to high rates of false positive sequence changes not associated with the trait, but that are either neutral or associated with another shared trait. To mitigate these false positives, multiple pieces of evidence are necessary. In this regard, noncoding sequences (nucleotide changes) near some of the genes with specialized expression in songbirds evolved faster than in non-VPL species (Zhang et al., 2014). Searching for convergent protein coding sequences among extensive VPL birds, one study identified genes previously found to be associated with speech deficits, including ROBO1, and some overlap with those specialized in speech brain regions (Lei et al., 2017).

In a recent sequencing study of 43 families with individuals that have specific language impairment or dyslexia, rare mutations were identified in dozens of genes (Chen et al., 2017) as well as a common variant (SNP) (Devanna et al., 2018). Of these, 14 genes had previously been associated with language/dyslexia deficits when mutated, including a receptor involved in neurotransmission and plasticity (*GRIN2B*), neurotransmitter release (*ERC1*), and a urokinase receptor ligand (*SPRX2*) involved in promoting synapse formation. Novel candidates included some present in multiple affected families or multiple mutations in the same gene, including genes involved in cellular lipid activity (*STARD9*), potassium channels (*SCN9A*), and histone H3 methyltransferase (*KMT2D*) that modulates gene regulation of the chromatin. Thus far, none of the mutations have been studied functionally in nonhuman animals to address the mechanism that may cause the language deficit. When such studies are done, they will need to consider the type of VPL behavior and whether the brain pathways examined are specialized for song and spoken language.

3. Conclusion

In conclusion, studying nonhuman animals has provided insights into components of human speech and its evolution at the behavioral, neuroanatomical, and genetic levels. As expected, the molecular profiles linked to VPL in nonhuman animals and speech in humans are related to many genes. The challenges for the future will be to determine the remaining molecular players, narrowing down the most important ones, and deciphering their mechanisms of action and their gene networks for spoken language. This will enable scientists to decipher what is ubiquitous among vocally communicating species, what is specialized in different abilities of VPL, and what may be unique to humans. Eventually, this will allow us to disambiguate between the discrete and continuum hypothesis of vocal production learning and thereby spoken language.

NOTE

1. Following standard nomenclature, genes are denoted in italics, proteins in regular font. Uppercase letters denote the human version of the gene (i.e., *FOXP2*), lowercase the mouse version of the gene (i.e., *Foxp2*), for all other species *FoxP2* is used.

REFERENCES

Abramson, J. Z., Hernández-Lloreda, M. V., García L., Colmenares, F., Aboitiz, F., & Call, J. (2018). Imitation of novel conspecific and human speech sounds in the killer whale (*Orcinus orca*). *Proceedings of the Royal Society B: Biological Sciences*, *285*(1871), 20172171.

Adam, I., Mendoza, E., Kobalz, U., Wohlgemuth, S., & Scharff, C. (2016). FoxP2 directly regulates the reelin receptor VLDLR developmentally and by singing. *Molecular and Cellular Neuroscience*, *74*, 96–105.

Adam, I., Mendoza, E., Kobalz, U., Wohlgemuth, S., & Scharff, C. (2017). CNTNAP2 is a direct FoxP2 target in vitro and in vivo in zebra finches: Complex regulation by age and activity. *Genes, Brain and Behavior*, *16*(6), 635–642.

Araya-Salas, M., & Wright, T. (2013). Open-ended song learning in a hummingbird. *Biology letters*, *9*, 20130625.

Arbib, M. A., Liebal, K., & Pika, S. (2008). Primate vocalization, gesture, and the evolution of human language. *Current Anthropology*, *49*(6), 1053–1076.

Arriaga, G., & Jarvis, E. D. (2013). Mouse vocal communication system: Are ultrasounds learned or innate? *Brain and Language*, *124*(1), 96–116.

Arriaga, G., Zhou, E. P., & 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.

Balsby, T. J. S., & Bradbury, J. W. (2009). Vocal matching by orange-fronted conures (*Aratinga canicularis*). *Behavioural processes*, *82*(2), 133–139.

Balsby, T. J. & Scarl, J. C. (2008). Sex-specific responses to vocal convergence and divergence of contact calls in orange-fronted conures (*Aratinga canicularis*). *Proceedings of the Royal Society B: Biological Sciences*, *275*(1647), 2147–2154.

Baptista, L. F. (1977). Geographic variation in song and dialects of the Puget Sound white-crowned sparrow. *The Condor*, *79*, 356–370.

Baptista, L. F., & Schuchmann, K.-L. (1990). Song learning in the Anna hummingbird (*Calypte anna*). *Ethology*, *84*(1), 15–26.

Becker, M., Devanna, P., Fisher, S. E., & Vernes, S. C. (2018). Mapping of human FOXP2 enhancers reveals complex regulation. *Frontiers in Molecular Neuroscience*, *11*, 47.

Beckers, G. J. (2011). Bird speech perception and vocal production: A comparison with humans. *Human Biology*, *83*(2), 191–212.

Beecher, M. D. (2017). Birdsong learning as a social process. *Animal Behaviour*, *124*, 233–246.

Beecher, M. D., & Brenowitz, E. A. (2005). Functional aspects of song learning in songbirds. *Trends in Ecology and Evolution*, *20*(3), 143–149.

Beecher, M. D., Campbell, S. E. & Stoddard, P. K. (1994). Correlation of song learning and territory establishment strategies in the song sparrow. *Proceedings of the National Academy of Sciences*, *91*, 1450–1454.

Beecher, M. D., Stoddard, P. K., Campbell, S. E. & Horning, C. L. (1996). Repertoire matching between neighbouring song sparrows. *Animal Behaviour*, *51*, 917–923.

Belton, E., Salmond, C. H., Watkins, K. E., Vargha-Khadem, F., & Gadian, D. G. (2003). Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia. *Human Brain Mapping*, *18*(3), 194–200.

Belyk, M., & Brown, S. (2017). The origins of the vocal brain in humans. *Neuroscience and Biobehavioral Reviews*, *77*, 177–193.

Berg, K. S., Delgado, S., Cortopassi, K. A., Beissinger, S. R., & Bradbury, J. W. (2011). Vertical transmission of learned signatures in a wild parrot. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1728), 585–591.

Bertram, B. (1970). The Vocal Behaviour of the Indian Hill Mynah, Gracula religiosa. *Animal Behaviour Monographs*, *3*(2), 79–192, IN1.

Berwick, R. C., Okanoya, K., Beckers, G. J., & Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends in Cognitive Sciences*, *15*(3), 113–121.

Bettelheim, B. (1959). Feral children and autistic children. *American Journal of Sociology*, *64*(5), 455–467.

Blumstein, D. T. (1999). The evolution of functionally referential alarm communication: Multiple adaptations; multiple constraints. *Evolution of Communication*, *3*(2), 135–147.

Boeckx, C., & Grohmann, K. K. (2013). *The Cambridge handbook of biolinguistics*. Cambridge: Cambridge University Press.

Böhner, J. (1983). Song learning in the zebra finch (*Taeniopygia guttata*): selectivity in the choice of a tutor and accuracy of song copies. *Animal Behaviour*, *31*, 231–237.

Böhner, J. (1990). Early acquisition of song in the zebra finch, *Taeniopygia guttata*. *Animal Behaviour*, *39*, 369–374.

Bolhuis, J. J., Tattersall, I., Chomsky, N., & Berwick, R. C. (2014). How could language have evolved? *PLOS Biology*, *12*(8), e1001934.

Bonkowsky, J. L., & Chien, C.-B. (2005). Molecular cloning and developmental expression of foxP2 in zebrafish. *Developmental Dynamics*, *234*(3), 740–746.

Boughman, J. W. (1998). Vocal learning by greater spear-nosed bats. *Proceedings of the Royal Society B: Biological Sciences*, *265*(1392), 227–233.

Boughman, J. W., & Moss, C. F. (2003). Social sounds: Vocal learning and development of mammal and bird calls. In A. M. Simmons, R. R. Fay, & A. N. Popper (Eds.), *Acoustic communication* (pp. 138–224). New York: Springer.

Bowers, J. M., Perez-Pouchoulen, M., Edwards, N. S., & McCarthy, M. M. (2013). Foxp2 mediates sex differences in ultrasonic vocalization by rat pups and directs order of maternal retrieval. *Journal of Neuroscience*, *33*(8), 3276–3283.

Bowers, J. M., Perez-Pouchoulen, M., Roby, C. R., Ryan, T. E., & McCarthy, M. M. (2014). Androgen modulation of Foxp1 and Foxp2 in the developing rat brain: impact on sex specific vocalization. *Endocrinology*, *155*(12), 4881–4894.

Bradbury, J. W., & Balsby, T. J. (2016). The functions of vocal learning in parrots. *Behavioral Ecology and Sociobiology*, *70*(3), 293–312.

Bradbury, J. W., Cortopassi, K. A. & Clemmons, J. R.(2001). Geographical variation in the contact calls of orange-fronted parakeets. *The Auk*, *118*(4), 958–972.

Brainard, M. S., & Doupe, A. J. (2002). What songbirds teach us about learning. *Nature*, *417*, 351.

Brenowitz, E. A., & Beecher, M. D. (2005). Song learning in birds: Diversity and plasticity, opportunities and challenges. *Trends in Neurosciences*, *28*(3), 127–132.

Briefer, E. F. & McElligott, A. G. (2012). Social effects on vocal ontogeny in an ungulate, the goat, *Capra hircus*. *Animal Behaviour*, *83*(4), 991–1000.

Carroll, S. B. (2005). Evolution at two levels: On genes and form. *PLOS Biology*, *3*(7), e245.

Castellano, S., Parra, G., Sanchez-Quinto, F. A., Racimo, F., Kuhlwilm, M., Kircher, M., … Pääbo, S. (2014). Patterns of coding variation in the complete exomes of three Neandertals. *Proceedings of the National Academy of Sciences*, *111*(18), 6666–6671.

Castellucci, G. A., McGinley, M. J., & McCormick, D. A. (2016). Knockout of Foxp2 disrupts vocal development in mice. *Scientific Reports*, *6*, 23305.

Catchpole, C., & Slater, P. (1995). *Bird song: Biological themes and variations*. Cambridge: Cambridge University Press.

Chabout, J., Sarkar, A., Patel, S. R., Radden, T., Dunson, D. B., Fisher S. E., & Jarvis, E. D. (2016). A Foxp2 mutation implicated in human speech deficits alters sequencing of ultrasonic vocalizations in adult male mice. *Frontiers in Behavioral Neuroscience*, *10*, 197.

Chaiken, M., Boehner, J., & Marler, E. (1993). Song acquisition in European starlings, Sturnus vulgaris: A comparison of the songs of live-tutored, tape-tutored, untutored, and wild-caught males. *Animal Behaviour*, *46*, 1079–1090.

Chakraborty, M., & Jarvis, E. D. (2015). Brain evolution by brain pathway duplication. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1684), pii: 20150056.

Chalfin, L., Dayan, M., Levy, D. R., Austad, S. N., Miller, R. A., Iraqi, F. A., … Kimchi, T. (2014). Mapping ecologically relevant social behaviours by gene knockout in wild mice. *Nature Communications*, *5*, 4569.

Charng, W. L., Karaca, E., Coban Akdemir, Z., Gambin, T., Atik, M. M., Gu, S., … Lupski, J. R. (2016). Exome sequencing in mostly consanguineous Arab families with neurologic disease provides a high potential molecular diagnosis rate. *BMC Medical Genomics*, *9*(1), 42.

Charrier, C., Joshi, K., Coutinho-Budd, J., Kim, J. E., Lambert, N., de Marchena, J., … Polleux, F. (2012). Inhibition of SRGAP2 function by its human-specific paralogs induces neoteny during spine maturation. *Cell*, *149*(4), 923–935.

Chen, X. S., Reader, R. H., Hoischen, A., Veltman, J. A., Simpson, N. H., Francks, C., … Fisher, S. E. (2017). Next-generation DNA sequencing identifies novel gene variants and pathways involved in specific language impairment. *Scientific Reports*, *7*, 46105.

Chew, L. (1981). Geographic and individual variation in the morphology and sequential organization of the song of the Savannah sparrow (*Passerculus sandwichensis*). *Canadian Journal of Zoology*, *59*, 702–713.

Chiu, Y.-C., Li, M.-Y., Liu, Y.-H., Ding, J.-Y., Yu, J.-Y., & Wang, T.-W. (2014). Foxp2 regulates neuronal differentiation and neuronal subtype specification. *Developmental Neurobiology*, *74*(7), 723–738.

Christiansen, M. H., & Kirby, S. (2003). Language evolution: Consensus and controversies. *Trends in Cognitive Sciences*, *7*(7), 300–307.

Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004) Wild chimpanzees produce group-specific calls: a case for vocal learning? *Ethology*, *110*, 221–243.

Cruickshank, A. J., Gautier, J. & Chappuis, C. (1993). Vocal mimicry in wild African grey parrot *Psittacus erithacus. Ibis*, *135*, 293–299.

DasGupta, S., Ferreira, C. H., & Miesenböck, G. (2014). FoxP influences the speed and accuracy of a perceptual decision in *Drosophila*. *Science*, *344*(6186), 901–904.

Deecke, V. B., Ford, J. K., Spong, P. (2000). Dialect change in resident killer whales: implications for vocal learning and cultural transmission. *Animal Behaviour*, *60*, 629–638.

Dennis, M. Y., Nuttle, X., Sudmant, P. H., Antonacci, F., Graves, T. A., Nefedov, M., … Eichler, E. E. (2012). Evolution of human-specific neural SRGAP2 genes by incomplete segmental duplication. *Cell*, *149*(4), 912–922.

Derégnaucourt, S., & Gahr, M. (2013). Horizontal transmission of the father's song in the zebra finch (*Taeniopygia guttata*). *Biology Letters*, *9*, 20130247.

Deriziotis, P., & Fisher, S. E. (2017). Speech and language: Translating the genome. *Trends in Genetics*, *33*(9), 642–656.

Devanna, P., Chen, X. S., Ho, J., Gajewski, D., Smith, S. D., Gialluisi, A., … Vernes, S. C. (2018). Next-gen sequencing identifies non-coding variation disrupting miRNA-binding sites in neurological disorders. *Molecular Psychiatry*, *23*(5), 1375–1384.

Di Sciullo, A.-M., & Boeckx, C. (2011). *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*. Oxford: Oxford University Press.

Dougherty, M. L., Nuttle, X., Penn, O., Nelson, B. J., Huddleston, J., Baker, C., … Eichler, E. E. (2017). The birth of a human-specific neural gene by incomplete duplication and gene fusion. *Genome Biology*, *18*(1), 49.

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience*, *22*(1), 567–631.

Eales, L. A. (1985). Song learning in zebra finches: Some effects of song model availability on what is learnt and when. *Animal Behaviour*, *33*(4), 1293–1300.

Eaton, R. L. (1979). A beluga whale imitates human speech. *Carnivore*, *2*, 22–23.

Eco, U. (1995). *The search for the perfect language*. Hoboken, NJ: Wiley-Blackwell.

Eens, M., Pinxten, R., & Verheyen, R. F. (1992). Song learning in captive European starlings, *Sturnus vulgaris*. *Animal Behaviour*, *44*(6), 1131–1143.

Enard, W., Gehre, S., Hammerschmidt, K., Hölter, S. M., Blass, T., Somel, M., … Pääbo, S. (2009). A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell*, *137*(5), 961–971.

Enard, W., Khaitovich, P., Klose, J., Zöllner, S., Heissig, F., Giavalisco, P., … Ravid, R. (2002). Intra- and interspecific variation in primate gene expression patterns. *Science*, *296*(5566), 340–343.

Esser, K. H. (1994). Audio-vocal learning in a non-human mammal: the lesser spear-nosed bat *Phyllostomus discolor*. *Neuroreport*, *5*(14), 1718–1720.

Esser, K. H., & Schmidt, U. (1989). Mother-Infant Communication in the Lesser Spear-nosed Bat *Phyllostomus discolor* (*Chiroptera, Phyllostomidae*) - Evidence for Acoustic Learning. *Ethology*, *82*(2), 156–168.

Farabaugh, S. M., Linzenbold, A., & Dooling, R. J. (1994). Vocal plasticity in budgerigars (*Melopsittacus undulatus*): Evidence for social factors in the learning of contact calls. *Journal of Comparative Psychology*, *108*(1), 81–92.

Favaro, L., Neves, S., Furlati, S., Pessani, D., Martin, V., & Janik, V. M. (2016). Evidence suggests vocal production learning in a cross-fostered Risso's dolphin (*Grampus griseus*). *Animal Cognition*, *19*(4), 847–853.

Feenders, G., Liedvogel, M., Rivas, M., Zapka, M., Horita, H., Hara, E., … Jarvis, E. D. (2008). Molecular mapping of movement-associated areas in the avian brain: A motor theory for vocal learning origin. *PLOS ONE*, *3*(3), e1768.

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.

Ferland, R. J., Cherry, T. J., Preware, P. O., Morrisey, E. E., & Walsh, C. A. (2003). Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain. *Journal of Comparative Neurology*, *460*(2), 266–279.

Ferreira, A. R., Smulders, T. V., Sameshima, K., Mello, C. V., & Jarvis, E. D. (2006). Vocalizations and associated behaviors of the sombre hummingbird (*Aphantochroa cirrhochloris*) and the rufous-breasted hermit (*Glaucis hirsutus*). *Auk*, *123*(4), 1129–1148.

Filatova, O. A., Burdin, A. M., & Hoyt, E. (2010). Horizontal transmission of vocal traditions in killer whale (*Orcinus orca*) dialects. *Biology Bulletin*, *37*(9), 965–971.

Fisher, S. E., & Scharff, C. (2009). FOXP2 as a molecular window into speech and language. *Trends in Genetics*, *25*(4), 166–177.

Fischer, J., Wheeler, B. C., & J. P. Higham (2015). Is there any evidence for vocal learning in chimpanzee food calls? *Current Biology*, *25*(21), R1028–R1029.

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, *102*(2), 1213–1222.

Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, *4*(7), 258–267.

Fitch, W. T. (2006). Production of vocalizations in mammals. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (2nd ed., pp. 115–121). Oxford: Elsevier.

Fitch, W. T. (2010). *The evolution of language*. Cambridge: Cambridge University Press.

Fitch, W. T., & Kelley, J. P. (2000). Perception of vocal tract resonances by whooping cranes *Grus americana*. *Ethology*, *106*(6), 559–574.

Fitch, W. T., Neubauer, J., & Herzel, H (2002). Calls out of chaos: The adaptive significance of nonlinear phenomena in mammalian vocal production. *Animal Behaviour*, *63*(3), 407–418.

Foote, A. D., Griffin, R. M., Howitt, D., Larsson, L., Miller, P. J., & Hoelzel, A. R. (2006). Killer whales are capable of vocal learning. *Biology Letters*, *2*(4), 509–512.

French, C. A., Jin, X., Campbell, T. G., Gerfen, E., Groszer, M., Fisher, S. E., & Costa, R. M. (2012). An aetiological Foxp2 mutation causes aberrant striatal activity and alters plasticity during skill learning. *Molecular Psychiatry*, *17*(11), 1077–1085.

Fripp, D., Own, C., Quintana-Rizzo, E., Shapiro, A., Buckstaff, K., & Jankowski, K. (2005). Bottlenose dolphin (*Tursiops truncatus*) calves appear to model their signature whistles on the signature whistles of community members. *Animal Cognition*, *8*, 17–26.

Fröhlich, H., Rafiullah, R., Schmitt, N., Abele, S., & Rappold, G. A. (2017). Foxp1 expression is essential for sex-specific murine neonatal ultrasonic vocalization. *Human Molecular Genetics*, *26*(8), 1511–1521.

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. *Journal of Comparative Neurology*, *426*(2), 182–196.

Garamszegi, L. Z., Eens, M., Pavlova, D. Z., Avilés, J. M., & Møller, A. P. (2007). A comparative study of the function of heterospecific vocal mimicry in European passerines. *Behavioral Ecology*, *18*(6), 1001–1009.

Garland, E. C., Goldizen, A. W., Rekdahl, M. L., Constantine, R., Garrigue, C., Hauser, N. D., … Noad, M. J. (2011). Dynamic horizontal cultural transmission of humpback whale song at the ocean basin scale. *Current Biology*, *21*(8), 687–691.

Gaub, S., Fisher, S. E., & Ehret, G. (2016). Ultrasonic vocalizations of adult male Foxp2-mutant mice: Behavioral contexts of arousal and emotion. *Genes, Brain and Behavior*, *15*(2), 243–259.

Gaunt, S. L., Baptista, L. F., Sánchez, J. E., & Hernandez, D. (1994). Song learning as evidenced from song sharing in two hummingbird species (*Colibri coruscans* and *C. thalassinus*). *Auk*, *111*(1), 87–103.

Gedman, G., Pfenning, A. R., Wirthlin, M., Audet, J.-N., & Jarvis, E. D. (2017). *Further resolution of hypotheses on convergent brain regions for learned song in songbirds and speech in humans*. Washington, DC: Society for Neuroscience.

Geerling, J. C., Kim, M., Mahoney, C. E., Abbott, S. B. G., Agostinelli, L. J., Garfield, A. S., … Scammell, T. E. (2016). Genetic identity of thermosensory relay neurons in the lateral parabrachial nucleus. *American Journal of Physiology—Regulatory, Integrative and Comparative Physiology*, *310*(1), R41.

Ginsburg, N. (1963). Conditioned talking in the Mynah bird. *Journal of comparative and physiological psychology*, *56*, 1061–1063.

Grosslight, J. H., Harrison, P. C., & Weiser, C. M. (1962). Reinforcement control of vocal responses in the Mynah bird (*Gracula religiosa*). *The Psychological Record*, *12*, 193–201.

Groszer, M., Keays, D. A., Deacon, R. M., De Bono, J. P., Prasad-Mulcare, S., Gaub, S., … Enard, W. (2008). Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. *Current Biology*, *18*(5), 354–362.

Gu, Z., Kalambogias, J., Yoshioka, S., Han, W., Li, Z., Kawasawa, Y. I., … Yoshida, Y. (2017). Control of species-dependent cortico-motoneuronal connections underlying manual dexterity. *Science*, *357*(6349), 400–404.

Gultekin, Y. B., & Hage, S. R. (2017). Limiting parental feedback disrupts vocal development in marmoset monkeys. *Nature Communications*, *8*, 14046.

Gurney, M. E. (1982). Behavioral correlates of sexual differentiation in the zebra finch song system. *Brain Research*, *231*(1), 153–172.

Haesler, S., Rochefort, C., Georgi, B., Licznerski, P., Osten, P., & Scharff, C. (2007). Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus Area X. *PLOS Biology*, *5*(12), e321.

Haesler, S., Wada, K., Nshdejan, A., Morrisey, E. E., Lints,, T., Jarvis, E. D., & Scharff, C. (2004). FoxP2 expression in avian vocal learners and non-learners. *Journal of Neuroscience*, *24*(13), 3164–3175.

Hage, S. R., Gavrilov, N., & Nieder, A. (2013). Cognitive control of distinct vocalizations in rhesus monkeys. *Journal of Cognitive Neuroscience*, *25*(10), 1692–1701.

Hage, S. R., & Nieder, A. (2013). Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nature Communications*, *4*, 2409.

Hammerschmidt, K., Reisinger, E., Westekemper, K., Ehrenreich, L., Strenzke, N., & Fischer, J. (2012). Mice do not require auditory input for the normal development of their ultrasonic vocalizations. *BMC Neuroscience*, *13*, 40.

Hammerschmidt, K., Whelan, G., Eichele, G., & Fischer, J. (2015). Mice lacking the cerebral cortex develop normal song: Insights into the foundations of vocal learning. *Scientific Reports*, *5*, 8808.

Hamson, D. K., Csupity, A. S., Gaspar, J. M., & Watson, N. V. (2009). Analysis of Foxp2 expression in the cerebellum reveals a possible sex difference. *Neuroreport*, *20*(6), 611–616.

Hara, E., Perez, J. M., Whitney, O., Chen, Q., White, S. A., & Wright, T. F. (2015). Neural FoxP2 and FoxP1 expression in the budgerigar, an avian species with adult vocal learning. *Behavioural Brain Research*, *283*, 22–29.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*(5598), 1569–1579.

Heston, J. B., & White, S. A. (2015). Behavior-linked FoxP2 regulation enables zebra finch vocal learning. *Journal of Neuroscience*, *35*(7), 2885–2894.

Hile, A. G., Burley, N. T., Coopersmith, C. B., Foster, V. S. & Striedter, G. F. (2005), Effects of Male Vocal Learning on Female Behavior in the Budgerigar, *Melopsittacus undulatus. Ethology*, *111*: 901–923.

Hile, A. G., Plummer, T. K., & Striedter, G. F. (2000). Male vocal imitation produces call convergence during pair bonding in budgerigars, *Melopsittacus undulatus. Animal Behavior*, *59*(6), 1209–1218.

Hile, A. G., & Striedter, G. F. (2000). Call convergence within groups of female budgerigars (*Melopsittacus undulatus*). *Ethology*, *106*(12), 1105–1114.

Hindmarsh, A. (1986). The functional significance of vocal mimicry in song. *Behaviour*, *99*, 87–100.

Horng, S., Kreiman, G., Ellsworth, C., Page, D., Blank, M., Millen, K., & Sur, M. (2009). Differential gene expression in the developing lateral geniculate nucleus and medial geniculate nucleus reveals novel roles for Zic4 and Foxp2 in visual and auditory pathway development. *Journal of Neuroscience*, *29*(43), 13672–13683.

Hultsch, H., & Todt, D. (1981). Repertoire sharing and songpost distance in nightingales (*Luscinia megarhynchos*). *Behavioral Ecology and Sociobiology*, *8*, 183–188.

Immelmann, K. (1969). Song development in the zebra finch and other estrildid finches. In R. A. Hinde (Ed.), *Bird Vocalizations* (pp. 61–74). Cambridge: Cambridge University Press.

Janik, V. M. (2000). Whistle matching in wild bottlenose dolphins (*Tursiops truncatus*). *Science*, *289*, 1355–1356.

Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology*, *28*, 60–65.

Janik, V. M., & Slater, P. J. B. (1997). Vocal learning in mammals. *Advances in the Study of Behaviour*, *26*, 59–100.

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*(1), 749–777.

Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., … Butler, A. B. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience*, *6*(2), 151–159.

Jarvis, E. D., Mirarab, S., Aberer, A. J., Li, B., Houde, P., Li, C., … Zhang, G. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*, *346*(6215), 1320–1331.

Jones, A. E., ten Cate, C., & Slater, P. J. B. (1996). Early experience and plasticity of song in adult male zebra finches (*Taeniopygia guttata*). *Journal of Comparative Psychology*, *110*(4), 354–369.

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. Series B: Biological Sciences*, *252*(1334), https://doi.org/10.1098/rspb.1993.0055

Jürgens, U. (2002). 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.

Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: Evidence for fast mapping. *Science*, *304*(5677), 1682–1683.

Karten, H. J. (1969). The organization of the avian telencephalon and some speculations on the phylogeny of the amniote telencephalon. *Annals of the New York Academy of Sciences*, *167*(1), 164–179.

Kelley, L. A., Coe, R. L., Madden, J. R., & Healy, S. D. (2008). Vocal mimicry in songbirds. *Animal Behaviour*, *76*(3), 521–528.

King, A. P., & West, M. J. (1989). Presence of female cowbirds (*Molothrus ater ater*) affects vocal imitation and improvisation in males. *Journal of Comparative Psychology*, *103*(1), 39–44.

King, S. L., & Janik, V. M. (2013). Bottlenose dolphins use learned vocal labels to address each other. *Proceedings of the National Academy of Sciences*, *110*, 13216–13221

King, S. L., Sayigh, L. S., Wells, R. S., Fellner, W., & Janik, V. M. (2013). Vocal copying of individually distinctive signature whistles in bottlenose dolphins. *Proceedings of the Royal Society of London B: Biological Sciences*, *280*(1757), 20130053.

Knörnschild, M. (2014). Vocal production learning in bats. *Current Opinion in Neurobiology*, *28*, 80–85.

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.

Knörnschild, M., Nagy, M., Metz, M., Mayer, F., & von Helversen, O. (2012). Learned vocal group signatures in the polygynous bat *Saccopteryx bilineata. Animal Behaviour*, *84*(4), 671–679.

Konishi, M. (1963). The role of auditory feedback in the vocal behavior of the domestic fowl. *Ethology*, *20*(3), 349–367.

Konishi, M. (1985). Birdsong: From behavior to neuron. *Annual Review of Neuroscience*, *8*(1), 125–170.

Kremers, D., Lemasson, A., Almunia, J., & Wanker, R. (2012). Vocal Sharing and Individual Acoustic Distinctiveness within a Group of Captive Orcas (*Orcinus orca*). *Journal of comparative psychology*, *126*, 433–445.

Kroodsma, D. E. (1982). Appendix: A world survey of evidence for vocal learning in birds. In D. E. Kroodsma & E. H. Miller, *Acoustic communication in birds* (pp. 311–337). New York: Academic Press.

Kroodsma, D. E. (1989). Male eastern phoebes (*Sayornis phoebe*; *Tyrannidae, Passeriformes*) fail to imitate songs. *Journal of Comparative Psychology*, *103*(3), 227.

Kroodsma, D., Hamilton, D., Sánchez, J. E., Byers, B. E., Fandiño-Mariño, H., Stemple, D. W., … Powell, G. V. (2013). Behavioral evidence for song learning in the suboscine bellbirds (*Procnias* spp.; Cotingidae). *Wilson Journal of Ornithology*, *125*(1), 1–14.

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*(3), 477–487.

Kroodsma, D. E., & Miller, E. H. (1996). *Ecology and evolution of acoustic communication in birds.* Ithaca, NY: Comstock Publishing Associates.

Lai, C. S. L., 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*, 519.

Lai, C. S. L., Gerrelli, D., Monaco, A. P., Fisher, S. E., & Copp, A. J. (2003). FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain*, *126*(11), 2455–2462.

Lange, B. P. H., Henninghausen, C., Brill, M., & Schwab, F. (2016). Only cheap talk after all? New experimental psychological findings on the role of verbal proficiency in mate choice. *Psychology of Language and Communication*, *20*(1), 1–22.

Larson, R. K., Déprez, V., & Yamakido, H. (2010). *The evolution of human language.* Cambridge: Cambridge University Press.

Lawton, K. J., Wassmer, T. L., & Deitcher, D. L. (2014). Conserved role of *Drosophila melanogaster* FoxP in motor coordination and courtship song. *Behavioural Brain Research*, *268*, 213–221.

Le Boeuf, B. J., & Peterson, R. S. (1969). Dialects in elephant seals. *Science*, *166*, 1654–1656.

Le Boeuf, B. J., & Petrinovich, L. F. (1974). Dialects of northern elephant seals, *Mirounga angustirostris*: Origin and reliability. *Animal Behaviour, 22*(3), 656–663.

Lei, H., Yan, Z., Sun, X., Zhang, Y., Wang, J., Ma, C., … Sun, Z. (2017). Axon guidance pathways served as common targets for human speech/language evolution and related disorders. *Brain and Language*, *174*, 1–8.

Leitner, S., Nicholson, J., Leisler, B., DeVoogd, T. J., & Catchpole, C. K. (2002). Song and the song control pathway in the brain can develop independently of exposure to song in the sedge warbler. *Proceedings of the Royal Society of London B: Biological Sciences*, *269*(1509), 2519–2524.

Li, G., Wang, J., Rossiter, S. J., Jones, G., & Zhang, S. (2007). Accelerated FoxP2 evolution in echolocating bats. *PLOS ONE*, *2*(9), e900.

Lieberman, P., & Blumstein, S. E. (1988). *Speech physiology, speech perception, and acoustic phonetics*. Cambridge: Cambridge University Press.

Liégeois, F., Baldeweg, T., Connelly, A., Gadian, D. G., Mishkin M., & Vargha-Khadem, F. (2003). Language fMRI abnormalities associated with FOXP2 gene mutation. *Nature Neuroscience*, *6*(11), 1230–1237.

Liégeois, F. J., Hildebrand, M. S., Bonthrone, A., Turner, S. J, Scheffer, I. E., Bahlo, M., … Morgan, A. T. (2016). Early neuroimaging markers of FOXP2 intragenic deletion. *Scientific Reports*, *6*, 35192.

Lilly, J. C. (1965). Vocal mimicry in *Tursiops*: Ability to match numbers and durations of human vocal bursts. *Science*, *147*(3655), 300–301.

Liu, W.-C., Gardner, T. J., & Nottebohm, F. (2004). Juvenile zebra finches can use multiple strategies to learn the same song. *Proceedings of the National Academy of Sciences*, *101*(52), 18177–18182.

Liu,W.-C. & Kroodsma, D. E. (1999). Song development by chipping sparrows and field sparrows. *Animal Behaviour, 57*(6), 1275–1286.

Liu,W.-C. & Kroodsma, D. E. (2006). Song Learning by Chipping Sparrows: When, Where, and from Whom. *The Condor*, *108*(3), 509–517.

Liu, W.-C., & Nottebohm, F. (2007). A learning program that ensures prompt and versatile vocal imitation. *Proceedings of the National Academy of Sciences*, *104*(51), 20398–20403.

Liu, W.-C., Wada, K., Jarvis, E. D., & Nottebohm, F. (2013). Rudimentary substrates for vocal learning in a suboscine. *Nature Communications*, *4*, 2082.

Mackevicius, E. L., & Fee, M. S. (2017). Building a state space for song learning. *Current Opinion in Neurobiology*, *49*, 59–68.

Mahrt, E. J., Perkel, D. J., Tong, L., Rubel, E. W., & Portfors, C. V. (2013). Engineered deafness reveals that mouse courtship vocalizations do not require auditory experience. *Journal of Neuroscience*, *33*(13), 5573–5583.

Marler, P. (1970). A comparative approach to vocal learning: Song development in white-crowned sparrows. *Journal of Comparative and Physiological Psychology*, *71*(2), 1.

Marler P., & Peters S. (1977). Selective vocal learning in a sparrow. *Science*, *198*, 519–521.

Marler P., & Peters S. (1982). Structural changes in song ontogeny in the swamp sparrow *Melospiza georgiana. The Auk*, *99*, 446–458.

Marler, P., & Peters, S. (1987). A sensitive period for song acquisition in the song sparrow, *Melospiza melodia*: A case of age-limited learning. *Ethology*, *76*, 89–100.

Marler P., & Peters S. (1988). Sensitive periods for song acquisition from tape recordings and live tutors in the swamp sparrow, *Melospiza georgiana*. *Ethology*, *77*, 76–84.

Marler, P., & Tamura, M. (1962). Song dialects in three populations of white-crowned sparrows. *The Condor*, *64*, 368–377.

Marshall, A. J., Wrangham, R. W., & Arcadi, A. C. (1999). Does learning affect the structure of vocalizations in chimpanzees? *Animal Behaviour*, *58*, 825–830.

Meerschaut, I., Rochefort, D., Revencu, N., Petre, J., Corsello, C., Rouleau, G. A., … Callewaert, B. (2017). FOXP1-related intellectual disability syndrome: A recognisable entity. *Journal of Medical Genetics*, *54*(9), 613–623.

Mendoza, E., Colomb, J., Rybak, J., Pfluger, H. J., Zars, T., Scharff, C., & Brembs, B. (2014). *Drosophila* FoxP mutants are deficient in operant self-learning. *PLOS ONE*, *9*(6), e100648.

Mendoza, E., & Scharff, C. (2017). Protein-protein interaction among the FoxP family members and their regulation of two target genes, VLDLR and CNTNAP2 in the zebra finch song system. *Frontiers in Molecular Neuroscience, 10*, 112.

Meredith, R. W., Janecka, J. E., Gatesy, J., Ryder, O. A., Fisher, C. A., Teeling, E. C., … Murphy, W. J. (2011). Impacts of the cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science*, *334*(6055), 521–524.

Metfessel, M. (1935). Roller canary song produced without learning from external sources. *Science*, *81*, 470.

Mets, D. G., & Brainard, M. S. (2018). Genetic variation interacts with experience to determine interindividual differences in learned song. *Proceedings of the National Academy of Sciences*, *115*, 421–426.

Miller, G. F. (2013). Mutual mate choice models as the red pill in evolutionary psychology: Long delayed, much needed, ideologically challenging, and hard to swallow. *Psychological Inquiry*, *24*(3), 207–210.

Miller, J. E., Spiteri, E., Condro, M. C., Dosumu-Johnson, R. T., Geschwind, D. H., & White, S. A. (2008). Birdsong decreases protein levels of FoxP2, a molecule required for human speech. *Journal of Neurophysiology*, *100*(4), 2015–2025.

Morgan, A., Fisher, S. E., Scheffer, I., & Hildebrand, M. (2017). FOXP2-related speech and language disorders. In M. P. Adam, H. H. Ardinger, R. A. Pagon, et al. (Eds). *GeneReviews®* [Internet]. Seattle: University of Washington, Seattle; 1993–2019. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK368474/.

Morrison, R. G. & Nottebohm, F. (1993). Role of a telencephalic nucleus in the delayed song learning of socially isolated zebra finches. *Journal of Neurobiology*, *24*(8), 1045- 1064.

Mountjoy, J. D., & Lemon, R. E. (1995). Extended song learning in wild European starlings. *Animal Behaviour*, *49*(2), 357–366.

Mozzi, A., Forni, D., Clerici, M., Pozzoli, U., Mascheretti, S., Guerini, F. R., … Sironi, M. (2016). The evolutionary history of genes involved in spoken and written language: Beyond FOXP2. *Scientific Reports*, *6*, 22157.

Murayama, T., Iijima, S., Katsumata, H., Arai, K. (2014). Vocal imitation of human speech, synthetic sounds and Beluga sounds, by a Beluga (*Delphinapterus leucas*). *International Journal of Comparative Psychology*, *27*(3), 369–384.

Murugan, M., Harward, S., Scharff, C., & Mooney, R. (2013). Diminished FoxP2 levels affect dopaminergic modulation of corticostriatal signaling important to song variability. *Neuron*, *80*(6), 1464–1476.

Nelson, D. A. (2000). Song overproduction, selective attrition, and vocal dialects in the white-crowned sparrow. *Animal Behaviour*, *60*, 887–898.

Nelson, D. A., Hallberg, K. I. & Soha, J. A. (2004). Cultural evolution of Puget Sound white-crowned sparrow song dialects. *Ethology*, *110*, 879–908.

Nicholson, J. S., Buchanan, K. L., Marshall, R. C. & Catchpole, C. K. (2007). Song sharing and repertoire size in the sedge warbler, *Acrocephalus schoenobaenus*: changes within and between years. *Animal Behaviour*, *74*, 1585–1592.

Noad, M. J., Cato, D. H., Bryden, M. M., Jenner, M. N., & Jenner, K. C. S. (2000). Cultural revolution in whale songs. *Nature*, *408*(6812), 537–537.

Nordby, J. C., Campbell, & Beecher, M. D. (2001). Late song learning in song sparrows. *Animal Behaviour*, *61*, 835–846.

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., & Nottebohm, M. E. (1978). Relationship between song repertoire and age in the canary, *Serinus canarius*. *Zeitschrift für Tierpsychologie*, *46*, 298–305.

Nottebohm, F., Nottebohm, M. E., & Crane, L. (1986). Developmental and seasonal changes in canary song and their relation to changes in the anatomy of song-control nuclei. *Behavioral and Neural Biology*, *46*(3), 445–471.

Nowak, M. A. (2000). Evolutionary biology of language. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *355*(1403), 1615–1622.

Nowicki, S. (1989). Vocal plasticity in captive black-capped chickadees: The acoustic basis and rate of call convergence. *Animal Behaviour*, *37*, 64–73.

Odom, K. J., Hall, M. L., Riebel, K., Omland, K. E., & Langmore, N. E. (2014). Female song is widespread and ancestral in songbirds. *Nature Communications*, *5*, 3379.

Ogburn, W. F. (1959). The wolf boy of Agra. *American Journal of Sociology*, *64*(5), 449–454.

Ohms, V. R., Escudero, P., Lammers, K., & ten Cate, C. (2012). Zebra finches and Dutch adults exhibit the same cue weighting bias in vowel perception. *Animal Cognition*, *15*(2), 155–161.

Panaitof, S. C., Abrahams, B. S., Dong, H., Geschwind, D. H., & White, S. A. (2010). Language-related Cntnap2 gene is differentially expressed in sexually dimorphic song nuclei essential for vocal learning in songbirds. *Journal of Comparative Neurology*, *518*(11), 1995–2018.

Pepperberg, I. M. (1981). Functional vocalizations by an African Grey parrot (*Psittacus erithacus*). *Ethology*, *55*(2), 139–160.

Pepperberg, I. M. (1994). Vocal learning in grey parrots (*Psittacus erithacus*): effects of social interaction, reference, and context. *The Auk*, *111*(2), 300–313.

Pepperberg, I. M. (2006). Cognitive and communicative abilities of grey parrots. *Applied Animal Behaviour Science*, *100*(1–2), 77–86.

Pepperberg, I. M. (2010). Vocal learning in grey parrots: A brief review of perception, production, and cross-species comparisons. *Brain and Language*, *115*(1), 81–91.

Petkov, C. I., & Jarvis, E. D. (2012). Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, *4*, 12.

Pfenning, A. R., Hara, E., Whitney, O., Rivas, M. V., Wang, R., Roulhac, P. L, … Jarvis, E. D. (2014). Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science*, *346*(6215), 1256846.

Phillmore, L. S., MacGillivray, H. L., Wilson, K. R., & Martin, S. (2015). Effects of sex and seasonality on the song control system and FoxP2 protein expression in black-capped chickadees (*Poecile atricapillus*). *Developmental Neurobiology*, *75*(2), 203–216.

Poole, J. H., Tyack, P. L., Stoeger-Horwath, A. S., & Watwood, S. (2005). Animal behaviour: Elephants are capable of vocal learning. *Nature*, *434*(7032), 455–456.

Poulsen, H. (1959). Song learning in the domestic canary. *Zeitschrift für Tierpsychologie*, *16*, 173–178.

Prat, Y., Azoulay, L., Dor, R., & Yovel, Y. (2017). Crowd vocal learning induces vocal dialects in bats: Playback of conspecifics shapes fundamental frequency usage by pups. *PLOS Biology*, *15*(10), e2002556.

Prat, Y., Taub, M., & Yovel, Y. (2015). Vocal learning in a social mammal: Demonstrated by isolation and playback experiments in bats. *Science Advances*, *1*(2), e1500019.

Price, J. J. (2015). Rethinking our assumptions about the evolution of bird song and other sexually dimorphic signals. *Frontiers in Ecology and Evolution*, *3*, 40.

Price, P. H. (1979). Developmental determinants of structure in zebra finch song. *Journal of Comparative and Physiological Psychology*, *93*(2), 260.

Ralls, K., Fiorelli, P., & Gish, S. (1985). Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Canadian Journal of Zoology*, *63*(5), 1050–1056.

Reimers-Kipping, S., Hevers, W., Pääbo, S., & Enard, W. (2011). Humanized Foxp2 specifically affects cortico-basal ganglia circuits. *Neuroscience*, *175*, 75–84.

Reiss, D., & McCowan, B. (1993). Spontaneous vocal mimicry and production by bottlenose dolphins (*Tursiops truncatus*): Evidence for vocal learning. *Journal of Comparative Psychology*, *107*(3), 301–312.

Richards, D. G. (1986). Dolphin vocal mimicry and vocal object labeling. In R. J. Schusterman, J. A. Thomas, F. G. Wood (Eds.), *Dolphin Cognition and Behavior: A Comparative Approach* (pp. 273–288). Hillsdale, NJ: L. Erlbaum Associates.

Richards , D. G., Wolz, J. P., & Herman, L. M. (1984). Vocal mimicry of computer-generated sounds and vocal labeling of objects by a bottlenose dolphin, *Tursiops truncatus*. *Journal of Comparative Psychology*, *98*, 10–28.

Ridgway, S., Carder, D., Jeffries, M. & Todd, M. (2012). Spontaneous human speech mimicry by a cetacean. *Current Biology*, *22*(20), R860–R861.

Rilling, J. K. (2014). Comparative primate neurobiology and the evolution of brain and language usage systems. *Current Opinion in Neurobiology*, *28*, 10–14.

Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., Hu, X., & Behrens, T. E. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience*, *11*(4), 426–428.

Rodenas-Cuadrado, P., Ho, J., & Vernes, S. C. (2014). Shining a light on CNTNAP2: Complex functions to complex disorders. *European Journal of Human Genetics*, *22*(2), 171–178.

Rodenas-Cuadrado, P. M., Mengede, J., Baas, L., Devanna, P., Schmid, T. A., Yartsev, M., … Vernes, S. C. (2018). Mapping the distribution of language related genes FoxP1, FoxP2 and CntnaP2 in the brains of vocal learning bat species. *Journal of Comparative Neurology*, *526*(8), 1235–1266.

Romero, I. G., Ruvinsky, I., & Gilad, Y. (2012). Comparative studies of gene expression and the evolution of gene regulation. *Nature Reviews Genetics*, *13*(7), 505–516.

Roper, A., & Zann, R. (2006). The onset of song learning and song tutor selection in fledgling zebra finches. *Ethology*, *112*(5), 458–470.

Rousso, D. L., Pearson, C. A., Gaber, Z. B., Miquelajauregui, A., Li, S., Portera-Cailliau, C., … Novitch, B. G. (2012). Foxp-mediated suppression of N-cadherin regulates neuroepithelial character and progenitor maintenance in the CNS. *Neuron*, *74*(2), 314–330.

Salinas-Melgoza, A., & Wright, T. (2012). Evidence for vocal learning and limited dispersal as dual mechanisms for dialect maintenance in a parrot. *PLoS ONE*, *7*(11), e48667.

Santos, M. E., Athanasiadis, A., Leitao, A. B., DuPasquier, L., & Sucena, E. (2011). Alternative splicing and gene duplication in the evolution of the FoxP gene subfamily. *Molecular Biology and Evolution*, *28*(1), 237–247.

Sanvito, S., Galimberti, F., & Miller, E. H. (2007). Observational evidences of vocal learning in southern elephant seals: A longitudinal study. *Ethology*, *113*(2), 137–146.

Saranathan, V., Hamilton, D., Powell, G. V., Kroodsma, D. E., & Prum, R. O. (2007). Genetic evidence supports song learning in the three-wattled bellbird *Procnias tricarunculata* (Cotingidae). *Molecular Ecology*, *16*(17), 3689–3702.

Scarantino, A., & Clay, Z. (2015). Contextually variable signals can be functionally referential. *Animal Behaviour*, *100*, e1–e8.

Scharff, C., & Haesler, S. (2005). An evolutionary perspective on FoxP2: Strictly for the birds? *Current Opinion in Neurobiology*, *15*(6), 694–703.

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. *Journal of Neuroscience*, *11*(9), 2896–2913.

Scharff, C., & Petri, J. (2011). Evo-devo, deep homology and FoxP2: Implications for the evolution of speech and language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1574), 2124.

Schatton, A., & Scharff, C. (2017). FoxP expression identifies a Kenyon cell subtype in the honeybee mushroom bodies linking them to fruit fly alphabeta$_c$ neurons. *European Journal of Neuroscience*, *46*(9), 2534–2541.

Schreiweis, C., Bornschein, U., Burguière, E., Kerimoglu, C., Schreiter, S., Dannemann, M., … Graybiel, A. M. (2014). Humanized Foxp2 accelerates learning by enhancing transitions from declarative to procedural performance. *Proceedings of the National Academy of Sciences*, *111*(39), 14253–14258.

Schulz, S. B., Haesler, S., Scharff, C., & Rochefort, C. (2010). Knockdown of FoxP2 alters spine density in Area X of the zebra finch. *Genes, Brain and Behavior*, *9*(7), 732–740.

Schusterman, R. J. (2008). Vocal learning in mammals with special emphasis on pinnipeds. In D. Kimbrough Oller & U. Griebel (Eds.), *The Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication*. MIT Press Scholarship Online, 10.7551/mitpress/9780262151214.001.0001

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.

Seyfarth, R. M., & Cheney, D. L. (2003). Meaning and emotion in animal vocalizations. *Annals of the New York Academy of Sciences*, *1000*(1), 32–55.

Seyfarth, R. M., & Cheney, D. L. (2010). Production, usage, and comprehension in animal vocalizations. *Brain and Language*, *115*(1), 92–100.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980a). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, *210*(4471), 801–803.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980b). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour*, *28*(4), 1070–1094.

Shi, Z., Piccus, Z., Zhang, X., Yang, H., Jarrell, H., Ding, Y., … & Li, X. (2018). miR-9 regulates basal ganglia-dependent developmental vocal learning and adult vocal performance in songbirds. *eLife*, *7*, e29087.

Simonyan, K. (2014). The laryngeal motor cortex: Its organization and connectivity. *Current Opinion in Neurobiology*, *28*, 15–21.

Sin, C., Li, H., & Crawford, D. A. (2015). Transcriptional regulation by FOXP1, FOXP2, and FOXP4 dimerization. *Journal of Molecular Neuroscience*, *55*(2), 437–448.

Siper, P. M., de Rubeis, S., Trelles, M. D. P., Durkin, A., Di Marino, D., Muratet, F., … Buxbaum, J. D. (2017). Prospective investigation of FOXP1 syndrome. *Molecular Autism*, *8*, 57.

Sjare, B., Stirling, I., & Spencer, C. (2003). Structural variation in the songs of Atlantic walruses breeding in the Canadian High Arctic. *Aquatic Mammals*, *29*(2), 297–318.

Slater, P. J., Eales, L. A., & Clayton, N. (1988). Song learning in zebra finches (*Taeniopygia guttata*), progress and prospects. *Advances in the Study of Behavior*, *18*, 1–34.

Snow, D.W. (1968). The singing assemblies of little hermits. *Living Bird*, *7*, 47–55.

Snowdon, C. T., & Elowson, A. M. (1999). Pygmy marmosets modify call structure when paired. *Ethology*, *105*(10), 893–908.

Song, X., Tang, Y., & Wang, Y. (2016). Genesis of the vertebrate FoxP subfamily member genes occurred during two ancestral whole genome duplication events. *Gene*, *588*(2), 156–162.

Song, X., Wang, Y., & Tang, Y. (2013). Rapid diversification of FoxP2 in teleosts through gene duplication in the teleost-specific whole genome duplication event. *PLOS ONE*, *8*(12), e83858.

Spiteri, E., Konopka, G., Coppola, G., Bomar, J., Oldham, M., Ou, J., … Geschwind, D. H. (2007). Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. *American Journal of Human Genetics*, *81*(6), 1144–1157.

Sprau, P., & Mundry, R. (2010). Song type sharing in common nightingales, Luscinia megarhynchos, and its implications for cultural evolution. *Animal Behaviour*, *80*, 427–434.

Stoeger, A. S., Mietchen, D., Oh, S., de Silva, S., Herbst, C. T., Kwon, S., & Fitch, W. T. (2012). An Asian elephant imitates human speech. *Current Biology*, *22*(22), 2144–2148.

Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D., Borjon, J. I., Holmes, P., & Ghazanfar, A. A. (2015). The developmental dynamics of marmoset monkey vocal production. *Science*, *349*(6249), 734–738.

Takahashi, D. Y., Liao, D. A., & Ghazanfar, A. A. (2017). Vocal learning via social reinforcement by infant marmoset monkeys. *Current Biology*, *27*(12), 1844–1852.

Tchernichovski, O., Mitra, P. P., & Nottebohm, T. L. F. (2001). Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science*, *291*(5513), 2564–2569.

Teramitsu, I., Kudo, L. C., London, S. E., Geshwind, D. H., & White, S. A. (2004). Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction. *Journal of Neuroscience*, *24*(13), 3152–3163.

Teramitsu, I., Poopatanapong, A., Torrisi, S., & White, S. A. (2010). Striatal FoxP2 is actively regulated during songbird sensorimotor learning. *PLOS ONE*, *5*(1), e8548.

Teramitsu, I., & White, S. A. (2006). FoxP2 regulation during undirected singing in adult songbirds. *Journal of Neuroscience*, *26*(28), 7390–7394.

Thompson, C. K., Schwabe, F., Schoof, A., Mendoza, E., Gampe, J., Rochefort, C., & Scharff, C. (2013). Young and intense: FoxP2 immunoreactivity in Area X varies with age, song stereotypy, and singing in male zebra finches. *Front Neural Circuits*, *7*, 24.

Todt, D. (1975) Social learning of vocal patterns and modes of their application in grey parrots. *Zeitschrift für Tierpsychologie*, *39*, 178–188.

Todt, D., & Geberzahn, N. (2003). Age-dependent effects of song exposure: Song crystallization sets a boundary between fast and delayed vocal imitation. *Animal Behaviour*, *65*(5), 971–979.

Todt, D., Hultsch, H. & Heike, D. (1979). Conditions affecting song acquisition in nightingales. *Zeitschrift für Tierpsychologie*, *51*, 23–35.

Tomasello, M. (2005). Beyond formalities: The case of language acquisition. *Linguistic Review*, *22*(2–4), 183–197.

Townsend, S. W., & Manser, M. B. (2011). The function of nonlinear phenomena in meerkat alarm calls. *Biology Letters*, *7*(1), 47–49.

Tsui, D., Vessey, J. P., Tomita, H., Kaplan, D. R., & Miller, F. D. (2013). FoxP2 regulates neurogenesis during embryonic cortical development. *Journal of Neuroscience*, *33*(1), 244–258.

Vargha-Khadem, F., Gadian, D. G., Copp, A., & Mishkin, M. (2005). FOXP2 and the neuroanatomy of speech and language. *Nature Reviews Neuroscience*, *6*, 131.

Vernes, S. C., Newbury, D. F., Abrahams, B. S., Winchester, L., Nicod, J., Groszer, M., … Fisher, S. E. (2008). A functional genetic link between distinct developmental language disorders. *New England Journal of Medicine*, *359*(22), 2337–2345.

Vernes, S. C., Oliver, P. L., Spiteri, E., Lockstone, H. E. Puliyadi, R., Taylor, J. M., … Fisher, S. E. (2011). Foxp2 regulates gene networks implicated in neurite outgrowth in the developing brain. *PLOS Genetics*, *7*(7), e1002145.

Vernes, S. C., Spiteri, E., Nicod, J., Groszer, M., Taylor, J. M., Davies, K. E., … Fisher, S. E. (2007). High-throughput analysis of promoter occupancy reveals direct neural targets of FOXP2, a gene mutated in speech and language disorders. *American Journal of Human Genetics*, *81*(6), 1232–1250.

Verstegen, A. M. J., Vanderhorst, V., Gray, P. A., Zeidel, M. L., & Geerling, J. C. (2017). Barrington's nucleus: Neuroanatomic landscape of the mouse pontine micturition center. *Journal of Comparative Neurology*, *525*(10), 2287–2309.

Vicario, A., Mendoza, E., Abellan, A., Scharff, C., & Medina, L. (2017). Genoarchitecture of the extended amygdala in zebra finch, and expression of FoxP2 in cell corridors of different genetic profile. *Brain Structure and Function*, *222*(1), 481–514.

von Humboldt, W. (1836). *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluß aus die geistige Entwickelung des Menschengeschlechts*. Berlin: Königliche Akademie der Wissenschaften.

von Merten, S., Hoier, S., Pfeifle, C., & Tautz, D. (2014). A role for ultrasonic vocalisation in social communication and divergence of natural populations of the house mouse (*Mus musculus domesticus*). *PLOS ONE*, *9*(5), e97244.

Wang, R., Chen, C. C., Hara, E., Rivas, M. V., Roulhac, P. L., Howard, J. T., … Jarvis, E. D. (2015). Convergent differential regulation of SLIT-ROBO axon guidance genes in the brains of vocal learners. *Journal of Comparative Neurology*, *523*(6), 892–906.

Wang, Y., Brzozowska-Prechtl, A., & Karten, H. J. (2010). Laminar and columnar auditory cortex in avian brain. *Proceedings of the National Academy of Sciences*, *107*(28), 12676–12681.

Waser, M. S., & Marler, P. (1977). Song learning in canaries. *Journal of Comparative and Physiological Psychology*, *91*(1), 1–7.

Watkins, K. E., Dronkers, N. F., & Vargha-Khadem, F. (2002). Behavioural analysis of an inherited speech and language disorder: Comparison with acquired aphasia. *Brain*, *125*(Pt 3), 452–464.

Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Wallace, E. K., Cheng, L., … Slocombe, K. E. (2015). Vocal learning in the functionally referential food grunts of chimpanzees. *Current Biology*, *25*(4), 495–499.

Weiß, B. M., Symonds, H., Spong, P., Ladich, F. (2011). Call sharing across vocal clans of killer whales: evidence for vocal imitation? *Marine Mammal Science*, *27*(2), E1–E13.

Wheelwright, N. T., Swett, M. B., Levin, I. I., Kroodsma, D. E., Freeman-Gallant, C. R., & Williams, H. (2008). The influence of different tutor types on song learning in a natural bird population. *Animal Behaviour*, *75*(4), 1479–1493.

Wilbrecht, L., & Nottebohm, F. (2003). Vocal learning in birds and humans. *Developmental Disabilities Research Reviews*, *9*(3), 135–148.

Wilden, I., Herzel, H., Peters, G., & Tembrock, G. (1998). Subharmonics, biphonation, and deterministic chaos in mammal vocalization. *Bioacoustics*, *9*(3), 171–196.

Wiley, R. H. (1971). Song groups in a singing assembly of little hermits. *The Condor*, *73*(1), 28–35.

Williams, H., Kilander, K., & Sotanski, M. L. (1993). Untutored song, reproductive success and song learning. *Animal Behaviour*, *45*(4), 695–705.

Williams, H., Levin, I. I., Norris, D. R., Newman, A. E. M., Wheelwright, N. T. (2013). Three decades of cultural evolution in Savannah sparrow songs. *Animal Behaviour*, *85*, 213–223.Wohlgemuth, S., Adam, I., & Scharff, C. (2014). FoxP2 in songbirds. *Current Opinion in Neurobiology*, *28*(Suppl C), 86–93.

Wright, T. F. (1996). Regional dialects in the contact call of a parrot. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *263*(1372), 867–872.

Wright, T. F. D., & Dahlin, C. R. (2017). Vocal dialects in parrots: Patterns and processes of cultural evolution. *Emu-Austral Ornithology*, *118*(1), 50–66.

Wright, T. F., Dahlin, C. R., & Salinas-Melgoza, A. (2008). Stability and change in vocal dialects of the yellow-naped amazon. *Animal behaviour, 76*(3), 1017–1027.

Wright, T. F., Rodriguez, A. M., & Fleischer, R. C. (2005). Vocal dialects, sex-biased dispersal, and microsatellite population structure in the parrot *Amazona auropalliata*. *Molecular Ecology*, *14*, 1197–1205.

Zann, R., & Dunstan, E. (2008). Mimetic song in superb lyrebirds: species mimicked and mimetic accuracy in different populations and age classes. *Animal Behaviour, 76*, 1043–1054.

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., … Wang, J. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, *346*(6215), 1311–1320.

Zhang, J., Webb, D. M., & Podlaha, O. (2002). Accelerated protein evolution and origins of human-specific features: FOXP2 as an example. *Genetics*, *162*(4), 1825.