



# Is the Capacity for Vocal Learning in Vertebrates Rooted in Fish Schooling Behavior?

Matz Larsson<sup>1,2,3</sup> · Benjamin W. Abbott<sup>4</sup>

Received: 26 October 2017 / Accepted: 7 June 2018 / Published online: 13 June 2018  
© The Author(s) 2018

## Abstract

The capacity to learn and reproduce vocal sounds has evolved in phylogenetically distant tetrapod lineages. Vocal learners in all these lineages express similar neural circuitry and genetic factors when perceiving, processing, and reproducing vocalization, suggesting that brain pathways for vocal learning evolved within strong constraints from a common ancestor, potentially fish. We hypothesize that the auditory-motor circuits and genes involved in entrainment have their origins in fish schooling behavior and respiratory-motor coupling. In this *acoustic advantages* hypothesis, aural costs and benefits played a key role in shaping a wide variety of traits, which could readily be exapted for entrainment and vocal learning, including social grouping, group movement, and respiratory-motor coupling. Specifically, incidental sounds of locomotion and respiration (ISLR) may have reinforced synchronization by communicating important spatial and temporal information between school-members and extending windows of silence to improve situational awareness. This process would be mutually reinforcing. Neurons in the telencephalon, which were initially involved in linking ISLR with forelimbs, could have switched functions to serve vocal machinery (e.g. mouth, beak, tongue, larynx, syrinx). While previous vocal learning hypotheses invoke transmission of neurons from visual tasks (gestures) to the auditory channel, this hypothesis involves the auditory channel from the onset. Acoustic benefits of locomotor-respiratory coordination in fish may have selected for genetic factors and brain circuitry capable of synchronizing respiratory and limb movements, predisposing tetrapod lines to synchronized movement, vocalization, and vocal learning. We discuss how the capacity to entrain is manifest in fish, amphibians, birds, and mammals, and propose predictions to test our acoustic advantages hypothesis.

**Keywords** Vocal learning · Incidental sound · Locomotion · Respiration · Locomotor-respiratory coupling · Fish schooling · Entrainment · Synchronization · FoxP2

## Introduction

Synchronized movement at organismal and group levels is ubiquitous in vertebrates, for example, schooling fish, swarming starlings, and ballet dancers. However, the ability to synchronize movement based on external auditory cues

(entrainment) and the capacity to imitate sounds vocally (vocal learning) are much less common (Patel 2014). However, there is emerging evidence that entrainment and vocal learning are more widespread than previously believed (Condro and White 2014; Scharff and Petri 2011; Schachner et al. 2009; Wilson and Cook 2016; Arriaga et al. 2012). Entrainment appears to be linked with the capacity for vocal learning, although the mechanisms behind this relationship are hotly debated (Schachner et al. 2009; Patel et al. 2009; Patel 2014; Soma and Mori 2015; Merker et al. 2015).

At its most basic level, the ability to learn new vocalizations based on auditory input requires three related abilities: (1) perceiving and processing the relevant auditory signal, (2) temporally buffering the signal (i.e. committing it to memory), and (3) reproducing the signal with a complex motor system (in this case the vocal tract). Despite more than 300 million years of evolutionary separation, fish and

✉ Matz Larsson  
larsson.matz@gmail.com

<sup>1</sup> The Heart, Lung and Physiology Clinic, Örebro University Hospital, Örebro, Sweden

<sup>2</sup> School of Health and Medical Sciences, Örebro University, Örebro, Sweden

<sup>3</sup> Clinical Health Promotion Centre, Lund University, Lund, Sweden

<sup>4</sup> Department of Plant and Wildlife Sciences, Brigham Young University, Provo, USA

tetrapods, including humans, share similar genetic factors (See: “[FoxP2 in Vocalization](#)”) and neuroanatomical structures for vocalization (Bass et al. 2008; Scharff and Petri 2011). Furthermore, all known vocal-learning vertebrates show behavioral, anatomical, and genetic convergence (Doupe and Kuhl 1999; Jarvis 2004; Nottebohm and Liu 2010; Pfenning et al. 2014; Condro and White 2014), with vocal learning closely associated with the neural and molecular mechanisms that regulate breathing, forelimb motor control, cerebral vocal control, and vocalization in the basal ganglia (Condro and White 2014; Feenders et al. 2008). Even the independently-evolved neurological hallmark of vocal learning—vocal control by telencephalic vocal nuclei (Feenders et al. 2008; Arriaga and Jarvis 2013)—shows striking similarities to control circuits in teleost fish (Forlano et al. 2014; Goodson and Bass 2002; Kittelberger and Bass 2013). Multiple gene factors and neural circuits shared by all vocal learners suggest strong genetic or epigenetic constraints on the evolution of this trait, potentially stemming from neuroanatomical structures inherited from a common ancestor (Jarvis 2004; Scharff and Petri 2011), either reptiles or fish.

The vocal learning and rhythmic synchronization hypothesis proposes that the ability to entrain to rhythmic auditory cues evolved as a by-product of vocal learning (Patel 2006; Schachner et al. 2009). It has also been proposed that entrainment evolved to attract mates, i.e. that vocal synchrony would amplify group broadcasting to distant females (Merker et al. 2009). Here we present an alternative hypothesis, that the ability to entrain movement to external sounds is prerequisite to the development of vocal learning and that some of the auditory-motor circuits and genes for entrainment and vocal learning have their origins in schooling behavior in fish. Because auditory perception plays a central role in fish synchronization (Pitcher et al. 1976; Faucher et al. 2010; Greenwood et al. 2013), molecular and neural mechanisms associated with schooling could have been co-opted for eventual entrainment and cerebral vocal control. In this paper, we develop this *acoustic advantages* hypothesis, specifically exploring how incidental sounds of locomotion and respiration (ISLR) could provide selective pressure to synchronize respiration and movement at the individual and group levels. We discuss how the capacity to entrain is manifest in fish, amphibians, birds, and mammals, and propose several experiments to test the acoustic advantages hypothesis. While no species basal to both teleost and tetrapod lineages still exist, we compare genetic, neural, and behavioral characteristics of vertebrates across these groups to assess likelihood of deep homology.

## A Spectrum of Entrainment and Vocal Learning

Entrainment and vocal learning have typically been defined as categorical abilities expressed by a narrow group of cognitively advanced animals (e.g. Merker et al. 2009, 2015; Patel et al. 2009). While qualitative differences may exist between vocal learners and non-learners (Jarvis 2004), much of the binary view of entrainment may be an experimental artefact (Scharff and Petri 2011; Arriaga et al. 2012). The few animals that have been tested for entrainment have typically been classified categorically based on whether they synchronize movement to human music or rhythmic sound (Arriaga et al. 2012; Petkov and Jarvis 2012). While entrainment to artificial sounds allows a relatively objective classification, interpreting the evolutionary relevance of this metric is not straightforward (Petkov and Jarvis 2012). Music and dance are influenced by culture, indeed they are products of culture, and major structural features of music and responsiveness to music are shaped by the cultural transmission process itself (Merker et al. 2015). For example, depending on the choice of music, the cultural background of the subject, and the degree of musical experience, it is likely that some humans would not entrain when exposed to some types of music. This cultural mismatch becomes orders of magnitude more problematic when applied across species (Bolhuis and Wynne 2009; Petkov and Jarvis 2012). Exposure to rhythmic sound that an animal would naturally experience (e.g. flapping of wings, swimming sounds, or footfalls) could offer a more appropriate test of entrainment abilities, though entrainment and vocal learning are admittedly difficult to quantify across diverse lineages.

In contrast to the categorical view, there is a growing perspective that motor and vocal synchronization and imitation exist on a continuum (Arriaga et al. 2012), ranging from phase matching (e.g. coordination of firefly flashes) to complex learned behaviors involving listening, practice, and performance (e.g. bird, whale, and human song; Wilson and Cook 2016; Scharff and Petri 2011; Petkov and Jarvis 2012). In this view, some level of entrainment and adaptability of vocalization is the norm among vertebrates and even some invertebrates. At one end of the spectrum, simple forms of temporal synchrony arise spontaneously due to sensorimotor coupling. At the other end, humans achieve predictive timing that features both phase and period correction, so-called “entrainment with perfect synchrony” or “negative asynchrony” where the response precedes the call, depending on advanced telencephalic learning (Merker et al. 2015, 2009). Notably, only a few species have indisputably been shown to be unable to exhibit vocal learning (Kroodtsma and Konishi 1991), and unrecognized intermediary phenotypes between accurate imitative ‘production’ learning and ‘usage’ learning may exist (Janik and Slater 2000). We use

this continuum hypothesis (Arriaga et al. 2012; Arriaga and Jarvis 2013) to explore temporal coordination associated with neuroanatomical and behavioral characteristics relevant to vocal learning as expressed in modern tetrapod lineages, including respiratory-motor coupling, forelimb motor processing, and sound-producing movements occurring during group locomotion.

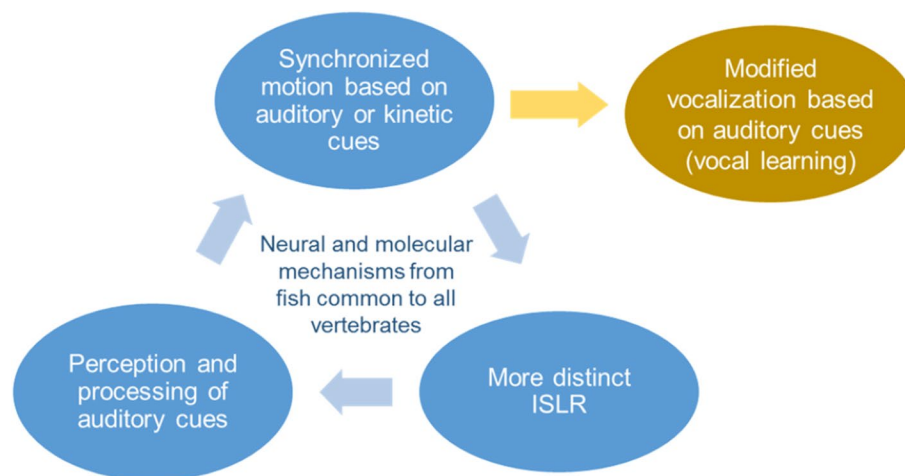
### Which Came First, Entrainment or Vocal Learning?

While it has been hypothesized that the ability to learn vocally is prerequisite to the ability to entrain to external acoustic cues, i.e. that vocal learning came first (Patel et al. 2009), entrainment has been observed in a large number of species that lack vocal learning abilities, including sea lions, frogs, and fireflies (Cook et al. 2013; Condro and White 2014; Wilson and Cook 2016; Ravnani et al. 2014). Additionally, auditory turn taking and call and response behaviors (arguably direct precursors to entrainment) are widespread in tetrapod lineages (Bee 2007; Borjon and Ghazanfar 2014; Louwerse et al. 2012). Before entrainment or vocal learning can occur, there must be awareness and responsiveness to the sonic environment (Fig. 1).

The ability to perceive and interpret complex sonic signals has a host of benefits, including facilitating social interactions (Wallin et al. 1999; Fitch 2010; Nagasaka et al. 2013; Bee 2007), and numerous individual and group acoustic advantages such as predator evasion and increased situational awareness (Larsson 2009, 2012b, 2012a, 2013, 2014; Vanesyan et al. 2015). To make the evolutionary leap from sonic awareness and motor entrainment to vocal learning (defined in the limited sense of being able to learn and

reproduce vocalizations), auditory-motor circuits would need to be co-opted for vocal control (Fig. 1). Molecular and neural mechanisms originally linking respiration with locomotion could have been readily exapted for this purpose because many structures in and surrounding the vocal tract and lungs are homologous to respiratory anatomy in fish, such as gill arches and the swim bladder (Fitch 2002; Bass and McKibben 2003). Indeed, both hearing (Ladich and Popper 2004) and the use of respiratory structures for vocalization extend back to fish, with some species using the swim bladder and other respiratory structures for sound production (Connaughton et al. 2002; Bass et al. 2008), potentially to locate and assemble with conspecifics (Kuznetsov 2009). In addition to the anatomical homology in vocal tracts, fish and tetrapods share deep molecular similarities. Genetic factors essential to vocalization and vocal learning in tetrapod lineages, such as the Forkhead box proteins FoxP1 and FoxP2 (see “Fact Sheet”; Lai et al. 2001; Wohlgemuth et al. 2014), are also present in fish, where they regulate analogous social and learning behaviors and development of the central nervous system (Condro and White 2014; Roberts et al. 2013; Scharff and Petri 2011; Bonkowsky et al. 2008).

Given the evidence of a possible ichthyic origin of entrainment and other precursors to vocal learning in tetrapods (Bass et al. 2008; Condro and White 2014; Roberts et al. 2013), we use the following sections to explore the selective pressures that could have led to the evolution of schooling and motor-respiratory coupling in fish. First, we explore genetic data relevant to the hypothesis, focusing on the transcription factor *FoxP2*, which has been investigated in a diverse range of species. Second, we consider the neuromolecular basis and selective advantages of fish



**Fig. 1** Conceptual model of the acoustic advantages hypothesis. The ability to synchronize movement based on external auditory cues (entrainment) could have evolved partially to minimize auditory masking in fish schools and confuse predator’s octovolateralis system. This synchronization results in organized ISLR, which improves

situational awareness of the individual and the group. The neural circuitry and genetic factors linking sound, movement, and respiration, which evolved in fish, could then have been co-opted for vocal synchronization in mammal and avian species capable of vocal learning

synchronization in general. Third, we examine the specific case of locomotor-respiratory coupling, which is particularly relevant to vocalization in air-breathing vertebrates (Bardy et al. 2015; Fabre et al. 2007). We adopt an aural approach in these investigations, in an attempt to understand how aquatic organisms such as fish experience their cacophonous environment. We use this auditory perspective to hypothesize origins of the genetic factors, neural circuits, anatomical structures, and social behaviors that underpin vocalization and vocal learning in modern tetrapod lineages.

### FoxP2 in Vocalization

A gene on chromosome 7 called Forkhead box P2, or FOXP2 (Lai et al. 2001) (see “Fact Sheet”) is essential to human language development, associated with structure of the striatum, cerebellum, and cortex (Lai et al. 2001; Vargha-Khadem et al. 2005; Mori and Wada 2015). Humans that are heterozygous for a non-functional FOXP2 allele experience impairment in speech and language development (Lai et al. 2001; Graham and Fisher 2015). One symptom is that they perceive rhythmic differences in stimuli less well and imitate rhythms less accurately than control subjects (Alcock et al. 2000). FOXP2 is proposed to play a crucial role for learned vocal-motor behavior such as singing and speech (Ayub et al. 2013).

FOXP2 affects dopamine-dependent learning processes in specific regions of the striatum (Schreiweis et al. 2014; Raghanti et al. 2016), though mice carrying only one functional *Foxp2*, show additional and partly opposite effects (Enard 2011), suggesting that FOXP2 has contributed to tuning cortico-basal ganglia circuits during the evolution of human speech and language (Enard et al. 2009; Schreiweis et al. 2014). Notably, the two amino-acid coding changes that differentiate the human sequence from that of chimpanzees were also present in Neandertal and Denisovan individuals (Krause et al. 2007; Enard 2016).

### Fact-Sheet

The name “Forkheadbox” was coined in 1989 when fruit fly researchers identified mutants with a «forked head» and they connected this phenotype to a specific stretch of DNA, called box (Weigel et al. 1989). This “forkhead box”, subsequently shortened to “Fox”, consists of more than forty genes organized into nineteen gene families: *FoxA*, *FoxB*, ...*FoxS*. \* These genes code for proteins regulating the function of other genes, in other words they are transcription factors that reduce or enhance activity by binding to specific regulatory regions of target genes. Several Fox transcription factors affect health and disease, but only the *P* gene family has been linked to speech and language (Wohlge-muth et al. 2014). Typically, genes have multiple roles at

numerous developmental time points or in different environmental contexts, for example, the ~20,000 protein-coding genes in the human genome are “re-used” in a number of different contexts in the brain and body (Fisher 2017). The pathways by which genetic factors influence neural circuitry and cognitive functions are indirect and require intermediate mechanisms such as proliferation, differentiation, connectivity, and plasticity. A gene cannot specify a precise behavior output, or a particular neural circuit, thus to discuss a “gene for language” is not particularly constructive (Fisher 2017). Likewise, *FoxP2* fills multiple roles. It influences the development of neural circuitry engaged in sensory guided motor learning, such as dendritic outgrowth and spine formation, and also the proper function of these circuits (Wohlge-muth et al. 2014). *FoxP2* is also important in lung development, thus the *FoxP2* gene is neither unique to humans nor exclusively associated with the central nervous system (Schatton and Scharff 2017).

\*When capitalized, *FOXP2* denotes the human gene, *Foxp2* = the gene in mice, and *FoxP2* = all other species (Kaestner et al. 2000). The genes are italicized (END OF FACT SHEET).

*FoxP2* exist in all 274 vertebrates studied so far (Scharff and Petri 2011), and is among the best conserved genes throughout vertebrate evolution, suggesting strong evolutionary pressure (Schatton and Scharff 2017). *FoxP2* has ancient roles in the growth and function of brain circuits in the cortex, basal ganglia and cerebellum, with bearing on sensorimotor integration and motor-skill learning (French et al. 2012; Scharff and Petri 2011; Fisher 2017). During embryonic development, *FoxP2* is expressed in homologous brain regions, including the basal ganglia, of human, monkeys, various rodent species, different bird species, frogs, and fish (Scharff and Petri 2011). In songbirds and mammals, the *FoxP2* protein is expressed in medium spiny neurons of the striatum, a region important for translating sensory stimuli (auditory in humans and birds, visual and tactile in mice) into motor acts (speech in humans, song in birds, and locomotion in mice; Schatton and Scharff 2017).

Beyond its developmental functions, *FoxP2* regulates the song circuitry in vocal-learning songbirds, humming-birds, and parrots (Feenders et al. 2008). These three orders are not linked by an immediate common ancestor (Hackett et al. 2008) and it is unclear whether they evolved the neural circuitry for vocal learning independently, or if a common ancestor to most extant birds possessed this trait that was more or less lost in non-vocal learners (Scharff and Petri 2011). *FoxP2* and its related molecular network have been proposed as evolutionary constraints contributing to convergent evolution of learned vocal communication in diverse taxa (Scharff and Petri 2011).

## Dopamine

In addition to its direct effects on brain function, *FoxP2* may be indirectly linked with vocal learning via interactions with the neurotransmitter dopamine. Experimental evidence in mice and birds suggests that *FoxP2* and its associated molecular network interact with dopamine signaling to regulate the strength of connections between particular sets of neurons, effectively fine tuning sensory-motor integration (Wohlgemuth et al. 2014). The basal ganglia have similar form and function in avian and mammalian brains, suggesting that sensory-motor learning in this circuit may rely on neurochemical rewards including dopamine transmission (Schatton and Scharff 2017). Auditory stimuli such as music listening stimulates dopamine release in the dorsal and ventral striatum in humans (Meehan et al. 2018; Zatorre 2015) and other animals (Panksepp and Bernatzky 2002; Sutoo and Akiyama 2004; Mavridis 2015). Dopamine may additionally aid in discrimination of important signals (Durstewitz et al. 1999), including footsteps and other ISLR, conveying adaptive value and contributing to the evolution of rhythmic abilities (Larsson 2014).

## Acoustic Benefits of Synchronized Movement for Fish

The occurrence of synchronized movements in all classes of vertebrates (Krause and Ruxton 2002; Larsson 2012a, 2013, 2014) suggests that the vertebrate brain has an innate capacity for kinetic coordination in groups, though this attribute may only be expressed under the right behavioral and ecological conditions (Larsson 2012a; Greenwood et al. 2013). Synchronized group movement in fish and other animals may have developed to improve foraging, hydro or aerodynamics, decision making, or predator evasion (Foster et al. 2001; Krause and Ruxton 2002; Svendsen et al. 2003; Parrish and Edelman-Keshet 1999; Berdahl et al. 2013). Fish experience an intensely aural world (Buerkle 1969; Slabbekoorn et al. 2010), sensing pressure waves with an advanced octavolateralis system consisting of the ear and lateral line, a pressure-sensitive organ running the length of their body (Coombs and Van Netten 2005). The octavolateralis system also provides mechanosensory input (i.e. sensing of water motion and electromagnetic cues for some fish), which is more or less analogous to hearing in different fish clades (Butler and Maruska 2016a, b). Consequently, acoustic advantages of schooling could have complimented or preceded other selective pressures, contributing to the emergence of sonic awareness and entrainment.

The majority of fish species school at some point in their development (Shaw 1978), though the type and synchrony of schooling depend on socioenvironmental conditions and species traits, including neurological and genetic factors

(Greenwood et al. 2013; Suriyampola et al. 2016). The ability to school requires two of the three prerequisites for vocal learning, i.e. perceiving and filtering auditory information and coupling that stimulus to complex movement (Fig. 1). While visual cues can strengthen schooling behavior (Rowland 1999), the molecular and neurological basis for schooling appears to be aural, since blind or blinded fish still school (Pitcher et al. 1976; Faucher et al. 2010; Kowalko et al. 2013) but fish with a temporarily disabled lateral line only do with great difficulty (Partridge and Pitcher 1980; Faucher et al. 2010).

One of the fundamental advantages of schooling is protection from predators (Shaw 1978; Greenwood et al. 2013). The protective benefits of schooling have primarily been attributed to visual confusion (Pavlov and Kasumyan 2000). However, many predatory fish rely heavily on their octavolateralis system (Moulton 1960; Larsson 2009) for prey detection (Montgomery and Bodznick 1999), particularly in the final stages of an attack (New et al. 2001). Consequently, to avoid being eaten, fish must thwart both the visual perception of a predator (Ruxton et al. 2007; Berdahl et al. 2013), and its auditory awareness (Larsson 2009). The latter could have been the primary or initial protective advantage of schooling in fish, i.e. to create complex, overlapping sound and pressure waves that confuse the predator's octavolateralis system (Larsson 2009, 2012b).

While sound from a school may confuse a predator, sounds from a solitary fish can attract predatory attack (New et al. 2001). One possible selective advantage of strongly synchronized ISLR, would be to reduce a predator's ability to identify the number of individuals based on sound alone. For example, in species that synchronize movement and respiratory sounds, a group may sound like an individual, and an individual like the group, potentially creating hesitation in predators and conveying secondary protection. Together, the benefits of being able to confuse a predator's inner ear, lateral line, and electrosensory system may have strengthened selection for individuals with the ability to synchronize their movements and respiration to external sound (Larsson 2009, 2012b).

Another advantage of synchronized movement is improved auditory perception of the school due to coordinated onset and cessation of ISLR (Larsson 2009). Noise from a fish's own movement or from external biotic and abiotic sources can drown out pertinent environmental sounds such as noise from predators or prey (Buerkle 1969; Slabbekoorn et al. 2010). Temporally coordinated movement and respiration in a school could reduce auditory masking because synchronous onset of noise improves auditory grouping (Bregman 1990) and synchronous cessation of movement creates momentary windows of silence, increasing situational awareness by allowing the group to

hear predators, prey, and signals from nearby conspecifics (Larsson 2009).

Within a school of fish, ISLR may also play a communicative function, broadcasting information about location and speed (Moulton 1960; Pitcher et al. 1976), and even fitness and social status (Butler and Maruska 2016b). ISLR produced by nearby school members may act as external auditory cues for group synchronization, linking auditory stimuli with motor control (Larsson 2012a, 2014). Since these are the building blocks for the capacity to entrain, the neural and genetic framework for schooling, which is present in cartilaginous and bony fish, could later increase likelihood of the evolution of synchronized vocalization and even vocal learning in some tetrapod species (Bee 2007; Roberts et al. 2013; Wohlgemuth et al. 2014). The communicative function of ISLR could result in an evolutionary positive feedback, with better synchronization resulting in more precise ISLR and in turn improving synchronization (Fig. 1). This feedback would rapidly increase the signal to noise ratio of ISLR, conveying selective advantage to individuals with neural circuitry and genes capable of deciphering complex auditory signals and linking this information with motor control (Lieberman 2001). The fact that reaction times to auditory stimuli are 14–40% faster than for visual stimuli in mammals and birds (Whitchurch and Takahashi 2006; Pain and Hibbs 2007; Shelton and Kumar 2010) could be a consequence of the highly developed auditory-motor circuits inherited from fish.

### Schooling Formations May Favor Acoustic Awareness

It has long been believed that that fish school to swim more efficiently (Weihs 1973). Faster swimming fish do group closer together and swim with greater synchrony (Ashraf et al. 2017). In a study of fish pairs, 90% of the pairs were synchronized when swimming at high speed (Ashraf et al. 2016). This tendency of fish to swim closer and synchronize movements more at high speed has been attributed solely to energy gain, though acoustic advantages could provide a complimentary advantage. Indeed, the energy gain hypothesis has been based more on theoretical modeling than observations from nature (Ashraf et al. 2017), and there are several lines of evidence that school architecture is not optimized for swimming efficiency. In several species, schooling fish do not swim in appropriate positions to maximize hydrodynamic advantage (Partridge and Pitcher 1979). Schools of the cohesive fish species *Hemigrammus bleheri* tend to gather in a phalanx configuration (side by side, similar to a military formation) when swimming fast, though an echelon formation would be energetically more efficient (Ashraf et al. 2017). We suggest that an alternative reason for swimming in phalanx could be improved perception of

neighbors' ISLR. Individual's ears and lateral lines will be relatively close and water-movements and sound will be relatively symmetrically propagated when swimming side by side. Moving close to one another, ISLR will be transmitted faster between fish and ISLR will be of larger amplitude, two things that are likely to reinforce synchronization based on acoustic cues. This hypothesis is supported by observations that phalanx formation is associated with local kinematic synchronization of each swimmer with its nearest neighbors (Ashraf et al. 2017).

### Evolution and Perceptual Consequences of Locomotor-Respiratory Coupling

While many of the physiological, molecular, and neural functions used during entrainment and vocal learning in tetrapods are present in fish, several notable changes in auditory-motor coupling have occurred since these groups diverged some 370 million years ago. Specifically, the lateral line, which is ineffectual in air, was lost in terrestrial tetrapods, and mouth breathing resulted in increased coupling of sound production and respiration (Bass and Chagnaud 2012). While locomotor-respiratory coupling was already present in fish (Tytell and Alexander 2007; Hoffman et al. 2016), air breathing necessitated deeper integration of central vocal motor commands into the respiratory cycle as well as somatosensory feedbacks in respiratory and vocal motor pathways (Smotherman et al. 2006). For example, in modern birds, the vocal motor system may occasionally take control over the respiratory musculature through central pathways, but pulmonary and laryngeal proprioceptive feedbacks usually allow interweaving vocalizations with breathing (Suthers et al. 2002). The parabrachial nucleus is a central constituent of the mammalian vocal motor pathway (Smotherman et al. 2006), with the lateral region strongly influencing respiratory rhythms and locomotor-respiratory coupling in mammals and other non-mammalian tetrapods (Bramble and Carrier 1983). Coordination of laryngeal and respiratory activity is synchronized through a network of interacting medullary and pontine feedback pathways—the same network that regulates breathing, chewing, swallowing, drinking, vomiting, coughing, and sneezing (Sakamoto et al. 1997). These fundamental functions long predate the exaptation of the system for vocal communication in terrestrial tetrapods (Smotherman 2007), in some cases extending back to fish ancestors (Perry et al. 2001). Together, the neural circuits regulating locomotor-respiratory pathways in fish provide a plausible neurological basis for the association between forelimb motor processing and vocalization in tetrapods (Bass and Chagnaud 2012).

### Locomotor-Respiratory Coupling in Individual Fish

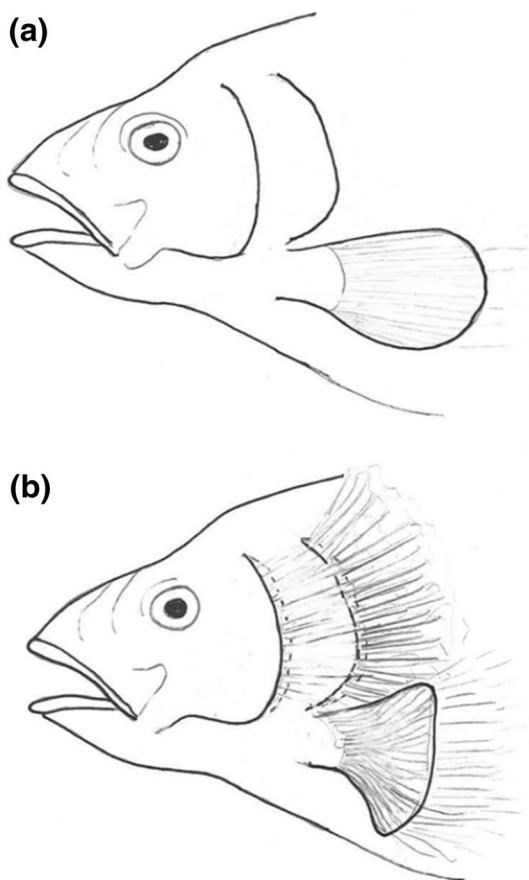
The synchronization of respiratory rhythms with locomotor patterns is evident in all classes of vertebrates (Bramble and Carrier 1983; Funk et al. 1992) and has been hypothesized to reduce energy consumption by avoiding competing muscle groups from working against one another (Bramble and Carrier 1983). However, energy gain by locomotor-respiratory coupling seems to be insignificant in humans (Wilke et al. 1975) and fish (Tytell and Alexander 2007), though there is some evidence to the contrary (Akanyeti et al. 2016). A non-mutually exclusive hypothesis is that acoustic benefits contributed to the evolution of locomotor-respiratory coupling. *Lepomis macrochirus* (Bluegill) tend to ventilate the gills every second or third pectoral fin beat, with a regular phase relationship between locomotion and ventilation (Tytell and

Alexander 2007). During pectoral fin abduction, the pumping operculum (the hard bony flap covering the gills) creates a jet of water that stops immediately after the fin is fully abducted (Fig. 2).

Abduction of the pectoral fin is likely to produce more intense masking sound, particularly during fast swimming, while in contrast, adduction of the fin will result in a more streamline fish body and consequently reduced turbulence. The synchronization of these two movements reduces interaction between flow from the operculum and flow over the pectoral fin, minimizing turbulence (Tytell and Alexander 2007) and likely reducing noise (Larsson 2012a). Notably, the pectoral fin operates adjacent to the inner ear. Altogether, locomotor-respiratory coupling in bluegill results in concurrent respiratory and movement turbulence (Tytell and Alexander 2007) and consequent noise, which prolongs quite intervals, reducing masking problems for an individual fish. There is recent evidence that salmonids (and likely other species) coordinate respiration and motion to improve auditory perception (Akanyeti et al. 2016), supporting the hypothesis that acoustic advantages played a role in the evolution of locomotor-respiratory coupling. Additionally, mechanosensory neurons in teleost fish and second-order electrosensory neurons in elasmobranch fish can regulate to cancel the effects of stimuli that are tied with fish respiratory movements (Montgomery and Bodznick 1994).

### Locomotor-Respiratory Coupling in the School

Little is known about respiratory behavior in schools, and it is unknown if fish synchronize respiration to elongate relatively quiet intervals. Operculum jet turbulence and other ISLR from individual fish are likely to be perceived by nearby conspecifics, potentially masking other environmental sounds. At the same time, the jet may communicate important information, about distance, speed, and position of nearby fishes, which may help fish maintain accurate distance and avoid collisions (Butler and Maruska 2016b). Since the operculum jet is phase locked with the sound of pectoral fins, the operculum jet may serve as an auditory cue reinforcing synchronization of fin movements, and vice versa. This locomotor-respiratory neural circuitry could potentially have been co-opted for auditory entrainment from external cues in tetrapod descendants. Notably, vocal and pectoral systems seem to have a shared developmental origin. So far all investigated tetrapods have forelimb motor neurons that serve in both sonic and gestural signalling (e.g. the pectoral fin of midshipman fish) (Bass and Chagnaud 2012). For some species this could be directly connect with vocalization such as for many Gobiidae species that use a pectoral-girdle-based mechanism for sound production (Parmentier et al. 2017). Similarly, herring and whiting emit



**Fig. 2** Drawing of locomotor-respiratory synchronization in bluegill (*Lepomis macrochirus*) during pectoral fin abduction (a) and during pectoral fin movement and operculum pumping (b). Lines in (b) represent turbulence and sound associated with respiration and fin movement. Bluegill and other fish synchronize locomotion and ventilation, potentially for auditory reasons. The synchronization of these two movements minimizes interaction between the flow from the operculum and flow over the pectoral fins, creating windows of silence and improving situational awareness

sound signals, ostensibly from the swim bladder, before rapid maneuvers (Gray and Denton 1991).

### Locomotor-Respiratory Coupling in Other Vertebrates

There is some evidence from other tetrapods that auditory advantages are associated with ISLR and synchronization of respiration and locomotion. Flocks of geese vocalize intensely before departure; their movements and vocal signals seem to act in combination as communicative cues (Ramseyer et al. 2009). Goose vocalization is amplified breathing sound and since breathing is phase-locked with wing flapping (Funk et al. 1993), these vocal calls inherently carry rhythmic information about location and trajectory, potentially reinforcing synchronization and the link between entrainment and external auditory cues. Vocalization also crescendos during take-off and once in flight before climbing, supporting the idea that this respiratory noise plays a communicative role. Many bird species, such as sparrows (*Passer domesticus*) and jackdaws (*Corvus monedula*), vocalize continuously when flying in formation. Breathing in dolphins is necessarily explosive (Fahlman et al. 2015), producing considerable turbulence and noise, and both diving dolphins (Hastie et al. 2003) and whales (Senigaglia and Whitehead 2012; Aoki et al. 2013) display synchronized breathing, potentially to reduce problems of auditory masking. Humans tend to synchronize breathing in groups unconsciously, a tendency that is reinforced by external rhythmic stimulation such as music listening (Codrons et al. 2014) or hearing a metronome (Bardy et al. 2015).

## Discussion

We hypothesized that the neural circuitry, genetic factors, and anatomy originally used in fish synchronization provided constraints on the evolution of entrainment and vocal learning in tetrapods, and specifically that locomotor-respiratory coupling in fish is linked with forelimb motor processing and vocalization in tetrapod descendants. Acoustic benefits of locomotor-respiratory coordination in fish may have selected for genetic factors and brain circuitry capable of synchronizing respiratory and limb movements, predisposing tetrapod lines to synchronized movement, vocalization, and vocal learning. Neuromolecular commonalities among tetrapods and fish and the strong link between forelimb motor processing and vocal communication in all investigated vertebrates support these hypotheses (Bass and Chagnaud 2012; Feenders et al. 2008; Scharff and Petri 2011).

Acoustic advantages may have also been one of the original benefits of social interaction for fish (Larsson 2012b), another factor increasing the likelihood of the evolution of auditory entrainment. Synchronization of movements during schooling may: (1) reduce masking problems caused by ISLR, (2) improve the signal function of ISLR within the school, and (3) confuse the lateral line and electrosensory perception of predators (Larsson 2009, 2012b, 2013). Our analysis of the literature indicated that synchronization of breathing and locomotion in individuals (i.e. locomotor-respiratory coupling) may also result in improved auditory perception—an alternative mechanism to the energetic efficiency hypothesis (Wilke et al. 1975; Tytell and Alexander 2007).

### Why are There So Few Vocal Learners?

If the roots of vocal learning extend all the way back to fish, another question emerges: why do so few vertebrates express vocal learning (Fitch 2011)? All tetrapods are descended from fish, however only a small fraction appear to have the capacity for vocal learning or entrainment based on external rhythmic cues (Patel 2014; Merker et al. 2015). Darwin's hypothesis that movement in response to rhythmic auditory stimuli was likely universal (Darwin 1871/1981) has not been borne out by the data, though only a fraction of all species have been tested for these abilities (Scharff and Petri 2011). A telencephalic explanation of the infrequency of vocal learning does little to resolve the vocal learning conundrum, because many highly encephalized vertebrates lack complex vocal learning abilities. Given the potential benefits of vocal learning and the diversity of taxa that show some version of call imitation, it is puzzling that some degree of vocal learning has not evolved in more species (Sewall et al. 2016), particularly in cognitively advanced, social mammals such as dogs, hyenas, lions, and non-human primates.

One null solution to the vocal learning conundrum is that entrainment and vocal learning are simply much wider spread than currently believed (Condro and White 2014; Patel 2014). Testing more species with more appropriate methods may reveal a spectrum of abilities better distributed across lineages (Scharff and Petri 2011) in line with the continuum hypothesis (Arriaga and Jarvis 2013). However, if entrainment and vocal learning prove to be rare after more exhaustive study, multiple selective pressures could account for this sparse distribution. One possibility, is that the specialized neural mechanisms that underpin vocal learning (Bolhuis and Wynne 2009; Chakraborty and Jarvis 2015; Jarvis 2004) may simple be difficult to evolve on a structural or technical level (Chakraborty and Jarvis 2015; Isler and van Schaik 2006; Mink et al. 1981). Another possibility is that tradeoffs associated with vocal learning could have prevented the evolution of this trait in certain groups.



The functional costs of learning processes, including time and social retaliation for making errors (Akçay et al. 2012; Sewall et al. 2016), can be avoided if unlearned vocalizations are sufficient for mediating social dynamics, such as when species live in social groups that are small, stable, and genetically homogenous. In such situations, unlearned calls, or calls learned during a single critical period, can function to mediate social interactions within and among groups without incurring costs associated with vocal learning (Seyfarth and Cheney 2014). It has been proposed that vocal learners could be more vulnerable to predation by standing out from the crowd (Jarvis 2006), though a non-typical call could also reduce identification as a prey item by the predator, counterbalancing this effect. Another potential selective pressure against vocal learning is that while flexible vocal expression can allow more advanced communication, it can also result in accidental social separation. For example, rapid divergence of vocalization, such as an isolated African elephant learning to imitate truck sounds (Poole et al. 2005), could negatively affect the fitness of individuals and prevent reintegration into the group. On the other hand, this vocal divergence could introduce cultural barriers or shibboleths that increase the likelihood of allopatric speciation, potentially conferring adaptive advantages to the lineage by increasing the rate of divergent evolution. Additional sexual selection mechanisms are also likely. For example, in songbirds, fish and whales, vocal learning is primarily used during courtship and breeding (reproductive advertisement), with many songbird and fish species losing vocal plasticity seasonally (Nottebohm and Liu 2010; Forlano et al. 2015).

Taken together, the tradeoffs outlined in the previous paragraph suggest that the ability to remember, recognize, and reproduce calls may only become beneficial in large, dynamic social groups (Sewall et al. 2016). This large-group hypothesis is complementary to our acoustic advantages hypothesis, because large groups are more likely to need efficient signaling systems to coordinate movements (e.g. complex flight formations of birds or bats), providing an additional selective pressure towards entrainment and vocal learning. More generally, acoustic advantages of grouping and synchronized movement in fish could have contributed to overcoming the challenges and tradeoffs of social living, a critical step in the evolution of vocal learning. Social group bonding and unconscious structuring (synchronization) of group movement during migration or long-distance travel could have contributed to the evolution of vocal learning by stimulating responsiveness to rhythmic auditory input.

Finally, it is important to note again that we are not proposing that the neural circuitry and gene factors used to synchronize movement necessarily lead to vocal learning, but that the neural circuitry and genetic factors used in synchronization, which are ancient and widespread, makes the convergent evolution towards vocal learning more likely

(Fig. 1). Neuroanatomical structures and genetic factors that may have evolved in fish due to acoustic advantages are part of the deep homology shared by all tetrapods, which could account for the strong similarities in modern avian and mammalian entrainers and vocal learners (Jarvis et al. 2000; Feenders et al. 2008).

### From Rudimentary Entrainment to High-Level Vocal Learning

Even if there is a deep homology going back to fish or beyond for the gene structures, social interactions, and neural circuits predisposing the evolution of entrainment and vocal learning (Scharff and Petri 2011), the modern expression of vocal learning in mammals and birds depends on direct, monosynaptic projection from telencephalic vocal nuclei (Merker et al. 2015) that exceeds analogous capacities in fish (Forlano et al. 2014; Goodson and Bass 2002; Kittelberger and Bass 2013). There is theoretical evidence suggesting that these telencephalic connections that functionally define the current understanding of vocal learning did not arise *de novo* (Arbib 2011), and general evidence that some of the framework for sensory-guided motor learning has deep roots (Forlano et al. 2014; Goodson and Bass 2002; Kittelberger and Bass 2013; Scharff and Petri 2011). In the case of hominids (Arbib 2011) and other vocal learners (Feenders et al. 2008), gestural control circuits may have collateralized vocal control regions. Since other tetrapods have a sensory-motor arrangement similar to the human “homunculus”—in other words that the brain area representing the forelimb is directly adjacent to the area that controls facial and vocal articulation (larynx/syrinx)—highly developed auditory-motor circuits that evolved partially for their acoustic advantages could have been easily co-opted for vocal control. Songbirds, parrots, and hummingbirds have cerebral vocal learning nuclei adjacent to discrete brain areas that are active during limb and body movements (Feenders et al. 2008).

According to the acoustic advantages hypothesis, synchronized locomotion based on ISLR processing includes mutual adjustments based on auditory cues (e.g. finbeats, wing-flapping, footsteps). Mutual adjustments of movements are similar, or even functionally identical, to mimicking. Because neurons in the telencephalon are plastic and may change function (Haubensak et al. 2004; Temple 2001), circuits initially involved in the processing of ISLR could have switched functions to serve sound production machinery (i.e. mouth, beak, tongue, larynx, syrinx) instead of the forelimbs. This hypothesis is compatible with the fact that new structures unique to vocal learners are telencephalic, providing a possible explanation for how telencephalic areas achieved direct connections with vocalization centers in vocal learners (Jarvis 2004). This hypothesized switch of auditory-motor control circuits is more parsimonious

than previous suggestions of transmission of neurons from visual tasks (gestures) to the auditory channel (Arbib 2011), because coupling would remain within auditory channels.

## Future Studies

It appears that ISLR allows human dyads and groups to unconsciously synch their locomotion (Nessler and Gilliland 2009; Nessler et al. 2011, 2013) and facilitates coordinated movement in fish (Pitcher et al. 1976; Pitcher 2001). Despite strong observational and theoretical evidence that sound plays a role in synchronization in many animal groups, there has been little experimental testing of this mechanism (Ravignani et al. 2014). Based on the evidence presented in this paper, ISLR may serve a signaling function to achieve synch in multiple animals groups including frogs, cetaceans, bats, and birds (Cook et al. 2013; Condro and White 2014; Wilson and Cook 2016; Ravignani et al. 2014). One way to test this hypothesis could be to investigate how animal groups synchronize movements when they are deprived of auditory cues, *e.g.* by using masking sound. Additionally, while synchronized movement is clearly beneficial for the group, it could incur costs to individuals, and there are likely internal physiological rewards such as dopamine release that reinforce this behavior (Larsson 2012a; Pinker 1997). In zebrafish, significant increases in the brain levels of dopamine have been reported during the period of gradually increasing schooling-related behavior (see Roberts et al. 2013 for review). Monitoring gene expression or hormonal response of animals in groups when exposed to ISLR would shed light on the generality of this phenomenon (*i.e.* do the same genetic factors and hormonal mechanisms encourage synchronization in phylogenetically distant lineages). There is already evidence that hearing music has measurable effects on animal behavior and brain chemistry (Panksepp and Bernatzky 2002; Meehan et al. 2018) and it would be of interest to investigate if rhythmic and un-rhythmic ISLR influence the release of dopamine in humans and other animals in similar or dissimilar ways.

Another testable hypothesis is that fish species that are particularly skilled at synching their locomotion (*e.g.* obligate schooling fish such as herring) may express genes that converge in some way with vocal learning species. As described earlier, there is evidence of common genetic factors (*e.g.* *FOXP2*) among zebra fish and vocal learners (Condro and White 2014; Lai et al. 2001; Wohlgenuth et al. 2014; Roberts et al. 2013; Scharff and Petri 2011), and a broader analysis of genetic factors in schooling and non-schooling fish within and among species could more definitively test this hypothesis. The hypothesis that fish near one another would synchronize gill-breathing could be observationally tested with visual or sound recordings. We predict that when situational awareness is the primary

pressure (*e.g.* when fish are at rest), nearby fish would synchronize respiration, particularly if predators were nearby. Concerning the communicative function of ISLR, if the pre-movement vocalization of herring or the vocalization of migratory birds is used for synchronization rather than just a neutral consequence of respiration, we would predict that individuals would vocalize less in isolation than in a group for the same level of exertion. A comparative study of vocal learning abilities in bird species in relation to their abilities to fly in synchronized flocks would also be of interest. Evaluating the acoustic advantages from locomotor-respiratory coupling (such as possible reduction of masking) could elucidate associations between forelimb motor processing and vocal communication in vertebrates. Acoustic principles in moving animal groups are scarcely investigated and we suggest that further study of associated physics and neural processes is critical to understanding how synchronized behavior originally evolved and also to recognize how and why extant animal groups and species differ in this respect.

In conclusion, there are multiple lines of evidence that acoustic advantages influenced the evolution of respiratory-locomotor linkages and complex synchronized movement (schooling), and that these capacities contributed to the emergence of social grouping, in vertebrates, and eventually in a few lineages to vocal learning. This acoustic advantages hypothesis is not mutually exclusive to the hypothesis that synchronization of movement increases social connection in groups of animals. Rather it complements the social connection hypothesis and suggests a possible evolutionary mechanism selecting for social behavior and occasionally vocal learning.

**Acknowledgements** We would like to thank our sons Filip and Caspian, whose highly-developed talents in making sound increased our awareness of the information its presence and absence can convey. We also thank Bill Thompson, for insightful input on the manuscript. B. Abbott was supported by the European Union's Seventh Framework Program for research, technological development and demonstration under Grant Agreement no. 607150 (FP7-PEOPLE-2013-ITN-INTERFACES—Ecohydrological interfaces as critical hotspots for transformations of ecosystem exchange fluxes and biogeochemical cycling).

## Compliance with Ethical Standards

**Conflict of interest** The authors have no conflict of interest to declare.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Akanyeti, O., Thornycroft, P. J., Lauder, G. V., Yanagitsuru, Y. R., Peterson, A. N., & Liao, J. C. (2016). Fish optimize sensing and respiration during undulatory swimming. *Nature Communications*, 7, 11044. <https://doi.org/10.1038/ncomms11044>.
- Akçay, E., Meirowitz, A., Ramsay, K. W., & Levin, S. A. (2012). Evolution of cooperation and skew under imperfect information. *Proceedings of the National Academy of Sciences of the United States of America*, 109(37), 14936–14941. <https://doi.org/10.1073/pnas.1212925109>.
- Alcock, K. J., Passingham, R. E., Watkins, K., & Vargha-Khadem, F. (2000). Pitch and timing abilities in inherited speech and language impairment. *Brain and Language*, 75(1), 34–46. <https://doi.org/10.1006/brln.2000.2323>.
- Aoki, K., Sakai, M., Miller, P. J. O., Visser, F., & Sato, K. (2013). Body contact and synchronous diving in long-finned pilot whales. *Behavioural Processes*, 99, 12–20. <https://doi.org/10.1016/j.beproc.2013.06.002>.
- Arbib, M. A. (2011). From mirror neurons to complex imitation in the evolution of language and tool use. *Annual Review of Anthropology*, 40, 257–273. <https://doi.org/10.1146/annurev-anthro-081309-145722>.
- Arriaga, G., & Jarvis, E. D. (2013). Mouse vocal communication system: Are ultrasounds learned or innate? *Brain and Language*, 124(1), 96–116. <https://doi.org/10.1016/j.bandl.2012.10.002>.
- 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. <https://doi.org/10.1371/journal.pone.0046610>.
- Ashraf, I., Bradshaw, H., Ha, T. T., Halloy, J., Godoy-Diana, R., & Thiria, B. (2017). Simple phalanx pattern leads to energy saving in cohesive fish schooling. *Proceedings of the National Academy of Sciences of the United States of America*, 114(36), 9599–9604. <https://doi.org/10.1073/pnas.1706503114>.
- Ashraf, I., Godoy-Diana, R., Halloy, J., Collignon, B., & Thiria, B. (2016). Synchronization and collective swimming patterns in fish (*Hemigrammus bleheri*). *Journal of the Royal Society, Interface*. <https://doi.org/10.1098/rsif.2016.0734>.
- Ayub, Q., Yngvadottir, B., Chen, Y., Xue, Y., Hu, M., Vernes, S. C., et al. (2013). FOXP2 targets show evidence of positive selection in European populations. *American Journal of Human Genetics*, 92(5), 696–706. <https://doi.org/10.1016/j.ajhg.2013.03.019>.
- Bardy, B. G., Hoffmann, C. P., Moens, B., Leman, M., & Dalla Bella, S. (2015). Sound-induced stabilization of breathing and moving. *Annals of the New York Academy of Sciences*, 1337, 94–100. <https://doi.org/10.1111/nyas.12650>.
- Bass, A. H., & Chagnaud, B. P. (2012). Shared developmental and evolutionary origins for neural basis of vocal-acoustic and pectoral-gestural signaling. *Proceedings of the National Academy of Sciences of the United States of America*, 109(Suppl 1), 10677–10684. <https://doi.org/10.1073/pnas.1201886109>.
- Bass, A. H., Gilliland, E. H., & Baker, R. (2008). Evolutionary origins for social vocalization in a vertebrate hindbrain-spinal compartment. *Science*, 321(5887), 417–421. <https://doi.org/10.1126/science.1157632>.
- Bass, A. H., & McKibben, J. R. (2003). Neural mechanisms and behaviors for acoustic communication in teleost fish. *Progress in Neurobiology*, 69(1), 1–26. [https://doi.org/10.1016/s0301-0082\(03\)00004-2](https://doi.org/10.1016/s0301-0082(03)00004-2).
- Bee, M. A. (2007). Sound source segregation in grey treefrogs: spatial release from masking by the sound of a chorus. *Animal Behaviour*, 74, 549–558. <https://doi.org/10.1016/j.anbehav.2006.12.012>.
- Berdahl, A., Torney, C. J., Ioannou, C. C., Faria, J. J., & Couzin, I. D. (2013). Emergent sensing of complex environments by mobile animal groups. *Science*, 339(6119), 574–576. <https://doi.org/10.1126/science.1225883>.
- Bolhuis, J. J., & Wynne, C. D. (2009). Can evolution explain how minds work? *Nature*, 458(7240), 832–833. <https://doi.org/10.1038/458832a>.
- Bonkowski, J. L., Wang, X., Fujimoto, E., Lee, J. E., Chien, C. B., & Dorsky, R. I. (2008). Domain-specific regulation of foxP2 CNS expression by lef1. *BMC Developmental Biology*, 8, 103. <https://doi.org/10.1186/1471-213x-8-103>.
- Borjon, J. I., & Ghazanfar, A. A. (2014). Convergent evolution of vocal cooperation without convergent evolution of brain size. *Brain, Behavior and Evolution*, 84(2), 93–102. <https://doi.org/10.1159/000365346>.
- Bramble, D. M., & Carrier, D. R. (1983). Running and breathing in mammals. *Science*, 219(4582), 251–256.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge: MIT Press.
- Buerkle, U. (1969). Auditory masking and the critical band in Atlantic Cod (*Gadus morhua*). *Journal of the Fisheries Research Board of Canada*, 26, 1113–1119.
- Butler, J. M., & Maruska, K. P. (2016a). The mechanosensory lateral line system mediates activation of socially-relevant brain regions during territorial interactions. *Frontiers in Behavioral Neuroscience*, 10(93), 1–18. <https://doi.org/10.3389/fnbeh.2016.00093>.
- Butler, J. M., & Maruska, K. P. (2016b). Mechanosensory signaling as a potential mode of communication during social interactions in fishes. *Journal of Experimental Biology*, 219(Pt 18), 2781–2789. <https://doi.org/10.1242/jeb.133801>.
- Chakraborty, M., & Jarvis, E. D. (2015). Brain evolution by brain pathway duplication. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. <https://doi.org/10.1098/rstb.2015.0056>.
- Codrons, E., Bernardi, N. F., Vandoni, M., & Bernardi, L. (2014). Spontaneous group synchronization of movements and respiratory rhythms. *PLoS ONE*, 9(9), e107538. <https://doi.org/10.1371/journal.pone.0107538>.
- Condro, M. C., & White, S. A. (2014). Recent advances in the genetics of vocal learning. *Comparative Cognition & Behavior Reviews*, 9, 75–98. <https://doi.org/10.3819/ccbr.2014.900003>.
- Connaughton, M. A., Fine, M. L., & Taylor, M. H. (2002). Weakfish sonic muscle: Influence of size, temperature and season. *The Journal of Experimental Biology*, 205(Pt 15), 2183–2188.
- Cook, P., Rouse, A., Wilson, M., & Reichmuth, C. (2013). A California sea lion (*Ziphius Californianus*) can keep the beat: Motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *Journal of Comparative Psychology*, 127(4), 412–427. <https://doi.org/10.1037/a0032345>.
- Coombs, S., & Van Netten, S. (2005). The hydrodynamics and structural mechanics of the lateral line system. In R. Shadwick & G. V. Lauder (Eds.), *Fish physiology* (Vol. 23, pp. 103–139). New York: Academic Press.
- Darwin, C. (1871/1981). *The descent of man, and selection in relation to sex*. London: John Murray (first edition reprinted by Princeton University Press, Princeton, NJ).
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567–631. <https://doi.org/10.1146/annurev.neuro.22.1.567>.
- Durstewitz, D., Kelc, M., & Güntürkün, O. (1999). A neurocomputational theory of the dopaminergic modulation of working memory functions. *Journal of Neuroscience*, 19(7), 2807–2822.
- Enard, W. (2011). FOXP2 and the role of cortico-basal ganglia circuits in speech and language evolution. *Current Opinion in Neurobiology*, 21(3), 415–424. <https://doi.org/10.1016/j.conb.2011.04.008>.

- Enard, W. (2016). The molecular basis of human brain evolution. *Current Biology*, 26(20), R1109–R1117. <https://doi.org/10.1016/j.cub.2016.09.030>.
- Enard, W., Gehre, S., Hammerschmidt, K., Holter, S. M., Blass, T., Somel, M., et al. (2009). A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell*, 137(5), 961–971. <https://doi.org/10.1016/j.cell.2009.03.041>.
- Fabre, N., Perrey, S., Arbez, L., & Rouillon, J. D. (2007). Neuro-mechanical and chemical influences on locomotor respiratory coupling in humans. *Respiratory Physiology & Neurobiology*, 155(2), 128–136. <https://doi.org/10.1016/j.resp.2006.04.015>.
- Fahlman, A., Loring, S. H., Levine, G., Rocho-Levine, J., Austin, T., & Brodsky, M. (2015). Lung mechanics and pulmonary function testing in cetaceans. *Journal of Experimental Biology*, 218(Pt 13), 2030–2038. <https://doi.org/10.1242/jeb.119149>.
- Faucher, K., Parmentier, E., Becco, C., Vandewalle, N., & Vandewalle, P. (2010). Fish lateral system is required for accurate control of shoaling behaviour. *Animal Behaviour*, 79(3), 679–687. <https://doi.org/10.1016/j.anbehav.2009.12.020>.
- Feenders, G., Liedvogel, M., Rivas, M., Zapka, M., Horita, H., Hara, E., et al. (2008). Molecular mapping of movement-associated areas in the avian brain: A motor theory for vocal learning origin. *PLoS ONE*, 3(3), e1768. <https://doi.org/10.1371/journal.pone.0001768>.
- Fisher, S. E. (2017). Evolution of language: Lessons from the genome. *Psychonomic Bulletin & Review*, 24(1), 34–40. <https://doi.org/10.3758/s13423-016-1112-8>.
- Fitch, W. T. (2002). Comparative vocal production and the evolution of speech: Reinterpreting the descent of the larynx. The transition to language (pp. 21–45). Oxford: Oxford University Press.
- Fitch, W. T. (2010). *The evolution of language*. Cambridge: Cambridge University Press.
- Fitch, W. T. (2011). The biology and evolution of rhythm: unravelling a paradox. In P. Rebuschat, M. Rohmeier, J. A. Hawkins & I. Cross (Eds.), *Language and music as cognitive systems* (pp. 73–95). Oxford: Oxford University Press.
- Forlano, P. M., Kim, S. D., Krzyminska, Z. M., & Sisneros, J. A. (2014). Catecholaminergic connectivity to the inner ear, central auditory, and vocal motor circuitry in the plainfin midshipman fish *porichthys notatus*. *Journal of Comparative Neurology*, 522(13), 2887–2927. <https://doi.org/10.1002/cne.23596>.
- Forlano, P. M., Sisneros, J. A., Rohmann, K. N., & Bass, A. H. (2015). Neuroendocrine control of seasonal plasticity in the auditory and vocal systems of fish. *Frontiers in Neuroendocrinology*, 37, 129–145. <https://doi.org/10.1016/j.yfrne.2014.08.002>.
- Foster, E. G., Ritz, D. A., Osborn, J. E., & Swadlow, K. M. (2001). Schooling affects the feeding success of Australian salmon (*Arripis trutta*) when preying on mysid swarms (*Paramesopodopsis rufa*). *Journal of Experimental Marine Biology and Ecology*, 261(1), 93–106.
- French, C. A., Jin, X., Campbell, T. G., Gerfen, E., Groszer, M., Fisher, S. E., et al. (2012). An aetiological Foxp2 mutation causes aberrant striatal activity and alters plasticity during skill learning. *Molecular Psychiatry*, 17(11), 1077–1085. <https://doi.org/10.1038/mp.2011.105>.
- Funk, G. D., Milsom, W. K., & Steeves, J. D. (1992). Coordination of wingbeat and respiration in the Canada goose. I. Passive wing flapping. *Journal of Applied Physiology*, 73(3), 1014–1024.
- Funk, G. D., Sholomenko, G. N., Valenzuela, I. J., Steeves, J. D., & Milsom, W. K. (1993). Coordination of wing beat and respiration in Canada Geese during free flight. *Journal of Experimental Biology*, 175, 317–323.
- Goodson, J. L., & Bass, A. H. (2002). Vocal-acoustic circuitry and descending vocal pathways in teleost fish: Convergence with terrestrial vertebrates reveals conserved traits. *Journal of Comparative Neurology*, 448(3), 298–322. <https://doi.org/10.1002/cne.10258>.
- Graham, S. A., & Fisher, S. E. (2015). Understanding language from a genomic perspective. *Annual Review of Genetics*, 49, 131–160. <https://doi.org/10.1146/annurev-genet-120213-092236>.
- Gray, J. A. B., & Denton, E. J. (1991). Fast pressure pulses and communication between fish. *Journal of the Marine Biological Association of the United Kingdom*, 71(1), 83–106.
- Greenwood, A. K., Wark, A. R., Yoshida, K., & Peichel, C. L. (2013). Genetic and neural modularity underlie the evolution of schooling behavior in threespine sticklebacks. *Current Biology*, 23(19), 1884–1888. <https://doi.org/10.1016/j.cub.2013.07.058>.
- Hackett, S. J., Kimball, R. T., Reddy, S., Bowie, R. C., Braun, E. L., Braun, M. J., et al. (2008). A phylogenomic study of birds reveals their evolutionary history. *Science*, 320(5884), 1763–1768. <https://doi.org/10.1126/science.1157704>.
- Hastie, G., Wilson, B., Tufft, L. H., & Thompson, P. M. (2003). Bottlenose dolphins increase breathing synchrony in response to boat traffic. *Marine Mammal Science*, 19(1), 74–84.
- Haubensak, W., Attardo, A., Denk, W., & Huttner, W. B. (2004). Neurons arise in the basal neuroepithelium of the early mammalian telencephalon: A major site of neurogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 101(9), 3196–3201. <https://doi.org/10.1073/pnas.0308600100>.
- Hoffman, M., Taylor, B. E., & Harris, M. B. (2016). Evolution of lung breathing from a lungless primitive vertebrate. *Respiratory Physiology & Neurobiology*, 224, 11–16. <https://doi.org/10.1016/j.resp.2015.09.016>.
- Isler, K., & van Schaik, C. (2006). Costs of encephalization: the energy trade-off hypothesis tested on birds. *Journal of Human Evolution*, 51(3), 228–243. <https://doi.org/10.1016/j.jhevol.2006.03.006>.
- Janik, V. M., & Slater, P. J. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, 60(1), 1–11. <https://doi.org/10.1006/anbe.2000.1410>.
- Jarvis, E. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences*, 1016, 749–777. <https://doi.org/10.1196/annals.1298.038>.
- Jarvis, E. (2006). Selection for and against vocal learning in birds and mammals. *Ornithological Science*, 5, 5–14.
- Jarvis, E., Ribeiro, S., da Silva, M. L., Ventura, D., Vielliard, J., & Mello, C. V. (2000). Behaviourally driven gene expression reveals song nuclei in hummingbird brain. *Nature*, 406(6796), 628–632. <https://doi.org/10.1038/35020570>.
- Kaestner, K. H., Knochel, W., & Martinez, D. E. (2000). Unified nomenclature for the winged helix/forkhead transcription factors. *Genes and Development*, 14(2), 142–146.
- Kittelberger, J. M., & Bass, A. H. (2013). Vocal-motor and auditory connectivity of the midbrain periaqueductal gray in a teleost fish. *Journal of Comparative Neurology*, 521(4), 791–812. <https://doi.org/10.1002/cne.23202>.
- Kowalko, J. E., Rohner, N., Rompani, S. B., Peterson, B. K., Linden, T. A., Yoshizawa, M., et al. (2013). Loss of schooling behavior in cavefish through sight-dependent and sight-independent mechanisms. *Current Biology*, 23(19), 1874–1883. <https://doi.org/10.1016/j.cub.2013.07.056>.
- Krause, J., Lalueza-Fox, C., Orlando, L., Enard, W., Green, R. E., Burbano, H. A., et al. (2007). The derived FOXP2 variant of modern humans was shared with Neandertals. *Current Biology*, 17(21), 1908–1912. <https://doi.org/10.1016/j.cub.2007.10.008>.
- Krause, J., & Ruxton, G. D. (2002). *Living in groups*. Oxford: Oxford University Press.
- Kroodsma, D. E., & Konishi, M. (1991). A subsong bird (Eastern Phoebe, *Sayornis phoebe*) develops normal song without auditory-feedback. *Animal Behaviour*, 42, 477–487. [https://doi.org/10.1016/s0003-3472\(05\)80047-8](https://doi.org/10.1016/s0003-3472(05)80047-8).

- Kuznetsov, M. Y. (2009). Traits of acoustic signalization and generation of sounds by some schooling physostomous fish. *Acoustical Physics*, 55(6), 866–875. <https://doi.org/10.1134/s1063771009060219>.
- Ladich, F., & Popper, A. N. (2004). Parallel evolution in fish hearing organs. In G. A. Manley, R. R. Fay & A. N. Popper (Eds.), *Evolution of the vertebrate auditory system* (pp. 95–127). New York: Springer.
- Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, 413(6855), 519–523. <https://doi.org/10.1038/35097076>.
- Larsson, M. (2009). Possible functions of the octavolateralis system in fish schooling. *Fish and Fisheries*, 10, 344–355.
- Larsson, M. (2012a). Incidental sounds of locomotion in animal cognition. *Animal Cognition*, 15(1), 1–13. <https://doi.org/10.1007/s10071-011-0433-2>.
- Larsson, M. (2012b). Why do fish school? *Current Zoology*, 58(1), 116–128.
- Larsson, M. (2013). Schooling fish: a multisensory approach. In S. Elias (Ed.), *Reference module in earth systems and environmental sciences*. New York: Elsevier.
- Larsson, M. (2014). Self-generated sounds of locomotion and ventilation and the evolution of human rhythmic abilities. *Animal Cognition*, 17(1), 1–14. <https://doi.org/10.1007/s10071-013-0678-z>.
- Lieberman, P. (2001). Human language and our reptilian brain: The subcortical bases of speech, syntax, and thought. *Perspectives in Biology and Medicine*, 44(1), 32–51. <https://doi.org/10.1353/pbm.2001.0011>.
- Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive science*, 36(8), 1404–1426. <https://doi.org/10.1111/j.1551-6709.2012.01269.x>.
- Mavridis, I. N. (2015). Music and the nucleus accumbens. *Surgical and Radiologic Anatomy*, 37(2), 121–125. <https://doi.org/10.1007/s00276-014-1360-0>.
- Meehan, A. D., Abbott, B. W., & Larsson, M. (2018). Movement is the song of the body: Reflections on the evolution of rhythm and music and its possible significance for the treatment of Parkinson's disease. *Evolutionary Studies in Imaginative Culture*, 1, 73–86.
- Merker, B., Madison, G. S., & Eckerdel, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, 45(1), 4–17. <https://doi.org/10.1016/j.cortex.2008.06.011>.
- Merker, B., Morley, I., & Zuidema, W. (2015). Five fundamental constraints on theories of the origins of music. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 370(1664), 20140095. <https://doi.org/10.1098/rstb.2014.0095>.
- Mink, J. W., Blumenschine, R. J., & Adams, D. B. (1981). Ratio of central nervous system to body metabolism in vertebrates: Its constancy and functional basis. *American Journal of Physiology*, 241(3), R203–212.
- Montgomery, J. C., & Bodznick, D. (1994). An adaptive filter that cancels self-induced noise in the electrosensory and lateral line mechanosensory systems of fish. *Neuroscience Letters*, 174(2), 145–148.
- Montgomery, J. C., & Bodznick, D. (1999). Signals and noise in the elasmobranch electrosensory system. *Journal of Experimental Biology*, 202(10), 1349–1355.
- Mori, C., & Wada, K. (2015). Songbird: A unique animal model for studying the molecular basis of disorders of vocal development and communication. *Experimental Animals*, 64(3), 221–230. <https://doi.org/10.1538/expanim.15-0008>.
- Moulton, J. M. (1960). Swimming sounds and the schooling of fish. *Biological Bulletin*, 119, 210–223.
- Nagasaka, Y., Chao, Z. C., Hasegawa, N., Notoya, T., & Fujii, N. (2013). Spontaneous synchronization of arm motion between Japanese macaques. *Scientific Reports*, 3, 1151. <https://doi.org/10.1038/srep01151>.
- Nessler, J. A., & Gilliland, S. J. (2009). Interpersonal synchronization during side by side treadmill walking is influenced by leg length differential and altered sensory feedback. *Human Movement Science*, 28(6), 772–785. <https://doi.org/10.1016/j.humov.2009.04.007>.
- Nessler, J. A., Gonzales, T., Rhoden, E., Steinbrick, M., & De Leone, C. J. (2011). Stride interval dynamics are altered when two individuals walk side by side. *Motor Control*, 15(3), 390–404.
- Nessler, J. A., McMillan, D., Schoulten, M., Shallow, T., Stewart, B., & De Leone, C. (2013). Side by side treadmill walking with intentionally desynchronized gait. *Annals of Biomedical Engineering*, 41(8), 1680–1691. <https://doi.org/10.1007/s10439-012-0657-6>.
- New, J. G., Fewkes, L. A., & Khan, A. N. (2001). Strike feeding behavior in the muskellunge, *Esox masquinongy*: Contributions of the lateral line and visual sensory systems. *Journal of Experimental Biology*, 204(6), 1207–1221.
- Nottebohm, F., & Liu, W. C. (2010). The origins of vocal learning: New sounds, new circuits, new cells. *Brain and Language*, 115(1), 3–17. <https://doi.org/10.1016/j.bandl.2010.05.002>.
- Pain, M. T., & Hibbs, A. (2007). Sprint starts and the minimum auditory reaction time. *Journal of Sports Sciences*, 25(1), 79–86. <https://doi.org/10.1080/02640410600718004>.
- Panksepp, J., & Bernatzky, G. (2002). Emotional sounds and the brain: The neuro-affective foundations of musical appreciation. *Behavioural Processes*, 60(2), 133–155.
- Parmentier, E., Petrinisec, M., Fonseca, P. J., & Amorim, M. C. P. (2017). Sound-production mechanism in *Pomatoschistus pictus*. *Journal of Experimental Biology*, 220(Pt 23), 4374–4376. <https://doi.org/10.1242/jeb.164863>.
- Parrish, J. K., & Edelman-Keshet, L. (1999). Complexity, pattern, and evolutionary trade-offs in animal aggregation. *Science*, 284(5411), 99–101.
- Partridge, B. L., & Pitcher, T. J. (1979). Evidence against a hydrodynamic function for fish schools. *Nature*, 279(5712), 418–419.
- Partridge, B. L., & Pitcher, T. J. (1980). The sensory basis of fish schools: Relative roles of lateral line and vision. *Journal of Comparative Physiology*, 135(4), 315–325.
- Patel, A. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception*, 24(1), 99–103. <https://doi.org/10.1525/mp.2006.24.1.99>.
- Patel, A. (2014). The evolutionary biology of musical rhythm: Was Darwin wrong? *PLoS Biology*, 12(3), e1001821. <https://doi.org/10.1371/journal.pbio.1001821>.
- Patel, A., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, 19(10), 827–830. <https://doi.org/10.1016/j.cub.2009.03.038>.
- Pavlov, D. S., & Kasumyan, A. O. (2000). Patterns and mechanisms of schooling behavior in fish: A review. *Journal of Ichthyology*, 40(Suppl.2), S163–S231.
- Perry, S. F., Wilson, R. J., Straus, C., Harris, M. B., & Remmers, J. E. (2001). Which came first, the lung or the breath? *Comparative Biochemistry and Physiology. Part A, Molecular and Integrative Physiology*, 129(1), 37–47.
- Petkov, C. I., & Jarvis, E. D. (2012). Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, 4, 12. <https://doi.org/10.3389/fnevo.2012.00012>.
- Pfenning, A. R., Hara, E., Whitney, O., Rivas, M. V., Wang, R., Roulhac, P. L., et al. (2014). Convergent transcriptional

- specializations in the brains of humans and song-learning birds. *Science*, 346(6215), 1256846. <https://doi.org/10.1126/science.1256846>.
- Pinker, S. (1997). The meaning of life. In *How the mind works* (pp. 526–538). New York: Norton.
- Pitcher, T. J. (2001). Fish schooling: Implications for pattern in the oceans and impacts on human fisheries. In J. H. Steele, K. K. Turekian & S. A. Thorpe (Eds.), *Encyclopedia of ocean sciences* (pp. 975–987). UK: Academic Press.
- Pitcher, T. J., Partridge, B. L., & Wardle, C. S. (1976). Blind fish can school. *Science*, 194(4268), 963–965.
- 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. <https://doi.org/10.1038/434455a>.
- Raghanti, M. A., Edler, M. K., Stephenson, A. R., Wilson, L. J., Hopkins, W. D., Ely, J. J., et al. (2016). Human-specific increase of dopaminergic innervation in a striatal region associated with speech and language: A comparative analysis of the primate basal ganglia. *Journal of Comparative Neurology*, 524(10), 2117–2129. <https://doi.org/10.1002/cne.23937>.
- Ramseyer, A., Petit, O., & Thierry, B. (2009). Decision-making in group departures of female domestic geese. *Behaviour*, 146, 351–371. <https://doi.org/10.1163/156853909x410955>.
- Ravignani, A., Bowling, D. L., & Fitch, W. T. (2014). Chorusing, synchrony, and the evolutionary functions of rhythm. *Frontiers in Psychology*, 5, 1118. <https://doi.org/10.3389/fpsyg.2014.01118>.
- Roberts, A. C., Bill, B. R., & Glanzman, D. L. (2013). Learning and memory in zebrafish larvae. *Frontiers in neural circuits*, 7, 126. <https://doi.org/10.3389/fncir.2013.00126>.
- Rowland, W. J. (1999). Studying visual cues in fish behavior: A review of ethological techniques. *Environmental Biology of Fishes*, 56(3), 285–305. <https://doi.org/10.1023/a:1007517720723>.
- Ruxton, G. D., Jackson, A. L., & Tosh, C. R. (2007). Confusion of predators does not rely on specialist coordinated behavior. *Behavioral Ecology*, 18(3), 590–596. <https://doi.org/10.1093/beheco/arm009>.
- Sakamoto, T., Nonaka, S., & Katada, A. (1997). Control of respiratory muscles during speech and vocalization. In A. Miller, A. L. Bianchi & B. P. Bishop (Eds.), *Neural control of the respiratory muscles* (pp. 249–258). New York: CRC Press.
- Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology*, 19(10), 831–836.
- 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 of London. Series B: Biological Sciences*, 366(1574), 2124–2140. <https://doi.org/10.1098/rstb.2011.0001>.
- Schatton, A., & Scharff, C. (2017). Next stop: Language The ‘FOXP2’ gene’s journey through time. *Metode Science Studies Journal*. <https://doi.org/10.7203/metode.7.7248>.
- Schreiweis, C., Bornschein, U., Burguiere, E., Kerimoglu, C., Schreiter, S., Dannemann, M., et al. (2014). Humanized Foxp2 accelerates learning by enhancing transitions from declarative to procedural performance. *Proceedings of the National Academy of Sciences of the United States of America*, 111(39), 14253–14258. <https://doi.org/10.1073/pnas.1414542111>.
- Senigaglia, V., & Whitehead, H. (2012). Synchronous breathing by pilot whales. *Marine Mammal Science*, 28(1), 213–219. <https://doi.org/10.1111/j.1748-7692.2011.00465.x>.
- Sewall, K. B., Young, A. M., & Wright, T. F. (2016). Social calls provide novel insights into the evolution of vocal learning. *Animal Behaviour*, 120, 163–172. <https://doi.org/10.1016/j.anbehav.2016.07.031>.
- Seyfarth, R. M., & Cheney, D. L. (2014). The evolution of language from social cognition. *Current Opinion in Neurobiology*, 28, 5–9. <https://doi.org/10.1016/j.conb.2014.04.003>.
- Shaw, E. (1978). Schooling fishes. *American Scientist*, 66(2), 166–175.
- Shelton, J., & Kumar, G. (2010). Comparison between auditory and visual simple reaction times. *Neuroscience and Medicine*, 1, 30–32. <https://doi.org/10.4236/nm.2010.11004>.
- Slabbekoorn, H., Bouton, N., van Opzeeland, I., Coers, A., ten Cate, C., & Popper, A. N. (2010). A noisy spring: The impact of globally rising underwater sound levels on fish. *Trends in Ecology & Evolution*, 25(7), 419–427. <https://doi.org/10.1016/j.tree.2010.04.005>.
- Smotherman, M. S. (2007). Sensory feedback control of mammalian vocalizations. *Behavioural Brain Research*, 182(2), 315–326. <https://doi.org/10.1016/j.bbr.2007.03.008>.
- Smotherman, M. S., Kobayasi, K., Ma, J., Zhang, S., & Metzner, W. (2006). A mechanism for vocal-respiratory coupling in the mammalian parabrachial nucleus. *Journal of Neuroscience*, 26(18), 4860–4869. <https://doi.org/10.1523/jneurosci.4607-05.2006>.
- Soma, M., & Mori, C. (2015). The songbird as a percussionist: Syntactic rules for non-vocal sound and song production in Java sparrows. *PLoS ONE*, 10(5), e0124876. <https://doi.org/10.1371/journal.pone.0124876>.
- Suriyampola, P. S., Shelton, D. S., Shukla, R., Roy, T., Bhat, A., & Martins, E. P. (2016). Zebrafish social behavior in the wild. *Zebrafish*, 13(1), 1–8. <https://doi.org/10.1089/zeb.2015.1159>.
- Suthers, R. A., Goller, F., & Wild, J. M. (2002). Somatosensory feedback modulates the respiratory motor program of crystallized birdsong. *Proceedings of the National Academy of Sciences of the United States of America*, 99(8), 5680–5685. <https://doi.org/10.1073/pnas.042103199>.
- Sutoo, D. e., & Akiyama, K. (2004). Music improves dopaminergic neurotransmission: Demonstration based on the effect of music on blood pressure regulation. *Brain Research*, 1016(2), 255–262. <https://doi.org/10.1016/j.brainres.2004.05.018>.
- Svendsen, J. C., Skov, J., Bildsoe, M., & Steffensen, J. F. (2003). Intra-school positional preference and reduced tail beat frequency in trailing positions in schooling roach under experimental conditions. *Journal of Fish Biology*, 62(4), 834–846. <https://doi.org/10.1046/j.1095-8649.2003.00068.x>.
- Temple, S. (2001). The development of neural stem cells. *Nature*, 414(6859), 112–117. <https://doi.org/10.1038/35102174>.
- Tytell, E. D., & Alexander, J. K. (2007). Bluegill *Lepomis macrochirus* synchronize pectoral fin motion and opercular pumping. *Journal of Fish Biology*, 70(4), 1268–1279. <https://doi.org/10.1111/j.1095-8649.2007.01416.x>.
- Vanesyan, A., Rodd, F. H., & Ryu, W. S. (2015). Tracking the startle response of guppies *Poecilia reticulata* in three dimensions. *Journal of Fish Biology*, 87(4), 981–999. <https://doi.org/10.1111/jfb.12773>.
- Vargha-Khadem, F., Gadian, D. G., Copp, A., & Mishkin, M. (2005). FOXP2 and the neuroanatomy of speech and language. *Nature Reviews Neuroscience*, 6(2), 131–138. <https://doi.org/10.1038/nrn1605>.
- Wallin, N. L., Merker, B., & Brown, S. (Eds.), (1999). *The origins of music*. London: The MIT Press.
- Weigel, D., Jürgens, G., Küttner, F., Seifert, E., & Jäckle, H. (1989). The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the Drosophila embryo. *Cell*, 57(4), 645–658. [https://doi.org/10.1016/0092-8674\(89\)90133-5](https://doi.org/10.1016/0092-8674(89)90133-5).
- Weihs, D. (1973). Hydrodynamics of fish schooling. *Nature*, 241, 290–291. <https://doi.org/10.1038/241290a0>.
- Whitchurch, E. A., & Takahashi, T. T. (2006). Combined auditory and visual stimuli facilitate head saccades in the barn owl (*Tyto alba*). *Journal of Neurophysiology*, 96(2), 730–745. <https://doi.org/10.1152/jn.00072.2006>.

- Wilke, J. T., Lansing, R. W., & Rogers, C. A. (1975). Entrainment of respiration to repetitive finger tapping. *Physiological Psychology*, 3(4), 345–349.
- Wilson, M., & Cook, P. F. (2016). Rhythmic entrainment: Why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed. *Psychonomic Bulletin & Review*, 23(6), 1647–1659. <https://doi.org/10.3758/s13423-016-1013-x>.
- Wohlgemuth, S., Adam, I., & Scharff, C. (2014). FoxP2 in songbirds. *Current Opinion in Neurobiology*, 28, 86–93. <https://doi.org/10.1016/j.conb.2014.06.009>.
- Zatorre, R. J. (2015). Musical pleasure and reward: Mechanisms and dysfunction. *Annals of the New York Academy of Sciences*, 1337, 202–211. <https://doi.org/10.1111/nyas.12677>.