



ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in
Behavioral
Sciences

Vocal learning: a language-relevant trait in need of a broad cross-species approach

Ella Z Lattenkamp^{1,2} and Sonja C Vernes^{1,3}

Although humans are unmatched in their capacity to produce speech and learn language, comparative approaches in diverse animal models are able to shed light on the biological underpinnings of language-relevant traits. In the study of vocal learning, a trait crucial for spoken language, passerine birds have been the dominant models, driving invaluable progress in understanding the neurobiology and genetics of vocal learning despite being only distantly related to humans. To date, there is sparse evidence that our closest relatives, nonhuman primates have the capability to learn new vocalisations. However, a number of other mammals have shown the capacity for vocal learning, such as some cetaceans, pinnipeds, elephants, and bats, and we anticipate that with further study more species will gain membership to this (currently) select club. A broad, cross-species comparison of vocal learning, coupled with careful consideration of the components underlying this trait, is crucial to determine how human speech and spoken language is biologically encoded and how it evolved. We emphasise the need to draw on the pool of promising species that have thus far been understudied or neglected. This is by no means a call for fewer studies in songbirds, or an unfocused treasure-hunt, but rather an appeal for structured comparisons across a range of species, considering phylogenetic relationships, ecological and morphological constraints, developmental and social factors, and neurogenetic underpinnings. Herein, we promote a comparative approach highlighting the importance of studying vocal learning in a broad range of model species, and describe a common framework for targeted cross-taxon studies to shed light on the biology and evolution of vocal learning.

Addresses

¹ Neurogenetics of Vocal Communication Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

² Division of Neurobiology, Ludwig-Maximilians University, Munich, Germany

³ Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands

Corresponding author: Vernes, Sonja C (Sonja.vernes@mpi.nl)

Current Opinion in Behavioral Sciences 2018, 21:xx-yy

This review comes from a themed issue on **The evolution of language**

Edited by **Christopher Petkov** and **William Marslen-Wilson**

<https://doi.org/10.1016/j.cobeha.2018.04.007>

2352-1546/© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Human speech and the need for animal models

Human speech and language acquisition is unparalleled within the animal kingdom. Human speech is dependent on the integration of auditory perception and vocal production, both during early development for babies learning their first words [1,2], and also later in life where audio-vocal feedback is essential for the maintenance of accurate speech production [3,4]. Although decades of research have elucidated many processes involved in human speech development, the full complexity of the mechanisms underlying human spoken language acquisition, like the integrated processes by which humans perceive, memorise, and compare auditory input to then guide vocal motor production, are still to be understood [5,6]. Deciphering the biological underpinnings of these abilities, such as the neuronal circuits or genetic mechanisms driving vocal learning, requires invasive *in vivo* studies that are largely technically and/or ethically impossible in humans. Such studies, including single cell electrophysiology, optogenetics, or genetic manipulations are essential to gain a thorough understanding of the biological processes underlying this trait. For this reason, animal models, in which controlled experiments can dissect the behavioural, neurological, and genetic components of language-relevant traits, are of great importance.

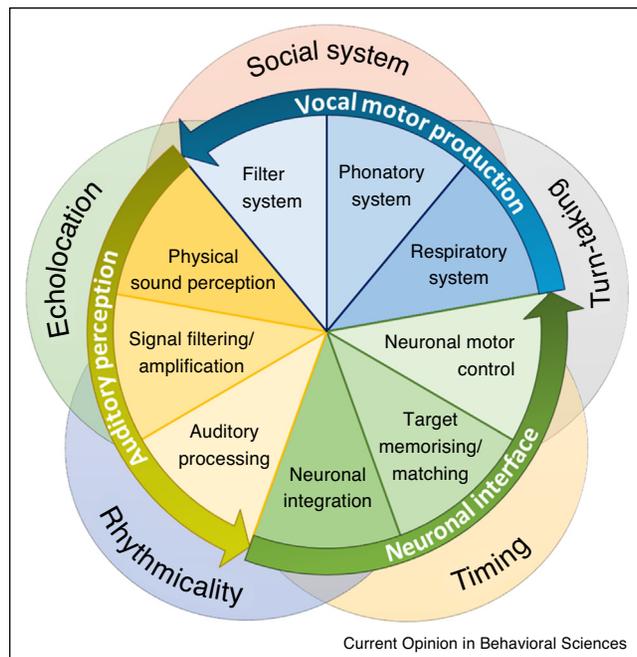
Vocal learning: a language-relevant trait

An essential component of human speech production and the acquisition of spoken language is the capacity for vocal production learning (herein referred to simply as ‘vocal learning’). Vocal learning is defined as the ability to acquire novel vocalisations or modify existing vocalisations as a result of auditory experience [7,8]. Vocal learning is the basis for the capacity to learn sounds and words used for human speech, but has also been identified in a handful of nonhuman species. This seemingly straightforward definition actually entails a complex multilevel process including auditory experience, neuronal integration, and vocal production (Figure 1).

Species-specific peculiarities of sound perception are a first limiting step for the ability to adjust vocalisations according to an auditory input. Auditory perception, including physical sound perception, signal filtering, and auditory processing, shapes and limits the input for subsequent neuronal processing of the sound (Figure 1). Once perceived, vocal signals need to be processed at the neural level, first being memorised as target sounds, and

2 The evolution of language

Figure 1



Conceptualisation of vocal learning. Three levels of organisation involved in vocal learning are illustrated (signal perception, neuronal integration, and vocal production) including their individual sublevels. The surroundings depict some of the hypothesised prerequisites for, or suggested traits correlated with, vocal learning – for which experimental evidence is still lacking or scarce.

in later recurrences compared to the auditory percept of self-emitted vocalisations. Neuro-motor control mechanisms are able to initialise sound production predominantly via the vocal apparatus. Sound generation can be achieved by vibrating the vocal folds (phonatory system) via exhalation (respiratory system), whereby the oscillation speed of the vocal folds determines the pitch of the vocalisation. The articulators in the vocal tract are used to filter the produced sound (filter system), which is a key step in the process of speech production (Figure 1; [9]). The accurate interplay between detailed sound perception, a precise neuronal interface, and a flexible vocal-motor production apparatus is essential to enable a dynamic adjustment of vocal emission to previously perceived and memorised auditory targets.

Vocal learning animals

Vocal learning has so far been studied most extensively in passerines or ‘songbirds’ and this research has provided invaluable insights into the physiological pathways, evolutionary factors, and gene expression patterns underlying vocal learning [10[•],11^{••},12,13]. Even though passerines are not the only vertebrate group, in fact not even the only order of birds, exhibiting vocal learning, they have some attributes that have made them a preferred model system.

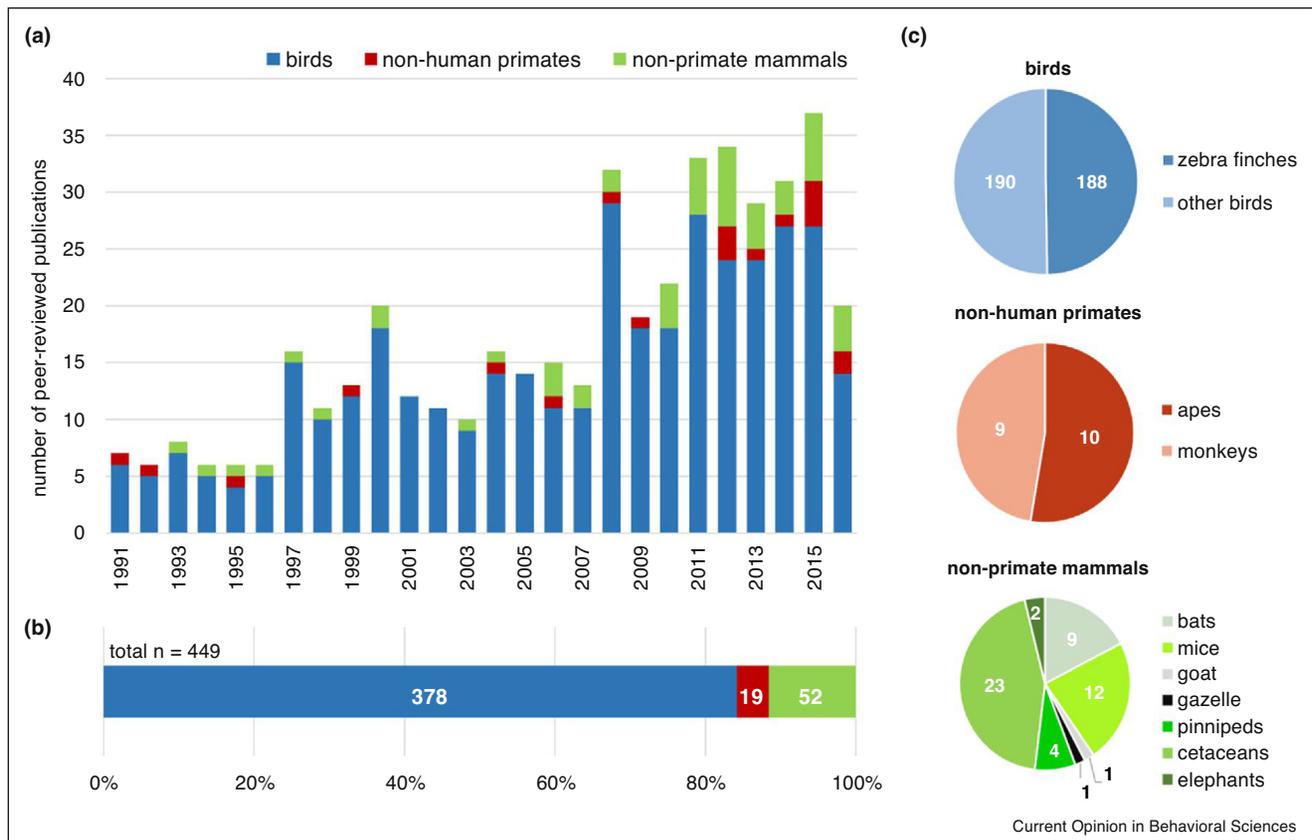
Songbirds are easy to keep, breed, and study and have a remarkable ability to copy detailed and complex song. Focusing on an easy and widely accessible study organism has certain benefits. The longstanding, combined focus of many labs around the world has advanced the study of vocal learning in songbirds immensely and provided detailed information about most levels of organisation from developmental processes and behavioural phenotypes to morphological structures, brain circuitry, and gene expression (e.g. [12,14–16,17^{••}]). Ease, accessibility, amenability to study, as well as historic factors, have led to a strong scientific focus on birds, such that when we performed a survey of published papers on vocal learning from the last 25 years, we found that papers on birds represented ~84% ($n = 378$) of all original research articles (Figure 2a and b). Despite vocal learning also being identified in bats, pinnipeds, cetaceans, and elephants, only 38 studies on vocal learning in these species (~8% of all articles) have been published over the same period (Figure 2c). Encouragingly, the field now seems to be expanding to explore these and other less traditional models that also show intriguing abilities to imitate structure, spectral content, and timing of acoustic signals. Several mammalian species found across varied habitats have shown evidence for vocal learning. Aquatic mammals (cetaceans, baleen whales, toothed whales, and dolphins) have shown vocal learning abilities both in the wild and in captive studies. Evidence for cetacean vocal learning includes adoption of new vocalisations from parents or conspecifics, copying new whistle sounds, and novel song learning and synchronisation [18^{••}]. In singular instances, pinnipeds (semiaquatic marine mammals), and elephants (terrestrial mammals) have demonstrated the ability to imitate human speech [19,20[•],21]. Moreover, several species of bats (the only mammals capable of self-powered flight) have shown indications of vocal learning [22^{••},23^{••}]. Bat species have shown individual and group signatures in their social calls, imitation of maternal calls by pups, and adjustment of spectral call parameters after transfer of animals between social groups [24–26].

Together these examples are already indicative for the existence of a broad spectrum of vocal learning species. However, only a small fraction of vertebrate species have thus far been tested, let alone recognised for their vocal learning abilities. For some mammals only one or a handful of individuals were ever reported to exhibit vocal learning. Nevertheless, if the observation is meaningful, it can be enough to establish a whole species as vocal learners with value for further study, as was the case for elephants [19]. These incidental observations should raise awareness for the large number of species that are possibly overlooked for their potential.

The vocal learning continuum

Despite a gradual increase in numbers of recognised vocal learning species, the method of classification of these

Figure 2



Primary research articles on 'vocal learning' from 1991 to 2016. **(a)** The presented numbers resulted from a search in the Web of Science Core Collection. The search term 'vocal learning' produced a total of 580 results in the section 'articles' (from 1970 to 2016), when searched for in 'topic'. Between 1970 and 1990 only 14 articles were published (precisely 2 studies on zebra finches, 8 on other bird species, 1 study on a monkey species, and 3 nonprimary research papers), which were excluded from the presented results. We further manually excluded a total of 117 papers, of which 93 did not present primary research (methodological, modelling, reviews, opinion paper, etc.), 22 were studies on humans, and 1 study each was on frogs and cichlid fish. **(b)** This led us to a total of 449 peer-reviewed articles between 1991 and 2016. The articles were categorised into three vertebrate groups (birds, nonhuman primates, and nonprimate mammals) and the total number of peer-reviewed articles for each category is given in the graphic. **(c)** The three categories were split to provide more detail about the studied taxa within each group. Note that different search platforms might produce slightly different results to these numbers, which were mined from Web of Science, however the same search performed using a different platform (Pubmed) returned very similar numbers.

species has only marginally changed over time. Vocal learning species are still frequently described bimodally as being either vocal learners or nonlearners. More valuable however is the idea of a vocal learning continuum — which has previously been proposed as a continuous distribution of species with increasing complexity of vocal learning ability [27]. Although the continuum has been represented as a two dimensional scale, it is probably better thought of as a multidimensional space in which each species has differing abilities and limitations. The vocal learning continuum must take into account capacities and constraints at all levels of organisation; auditory perception, neuronal interface, and vocal-motor production (Figure 1). And as noted previously, a large proportion of species are expected to cluster at the low end of the vocal learning continuum [27]. Considering where

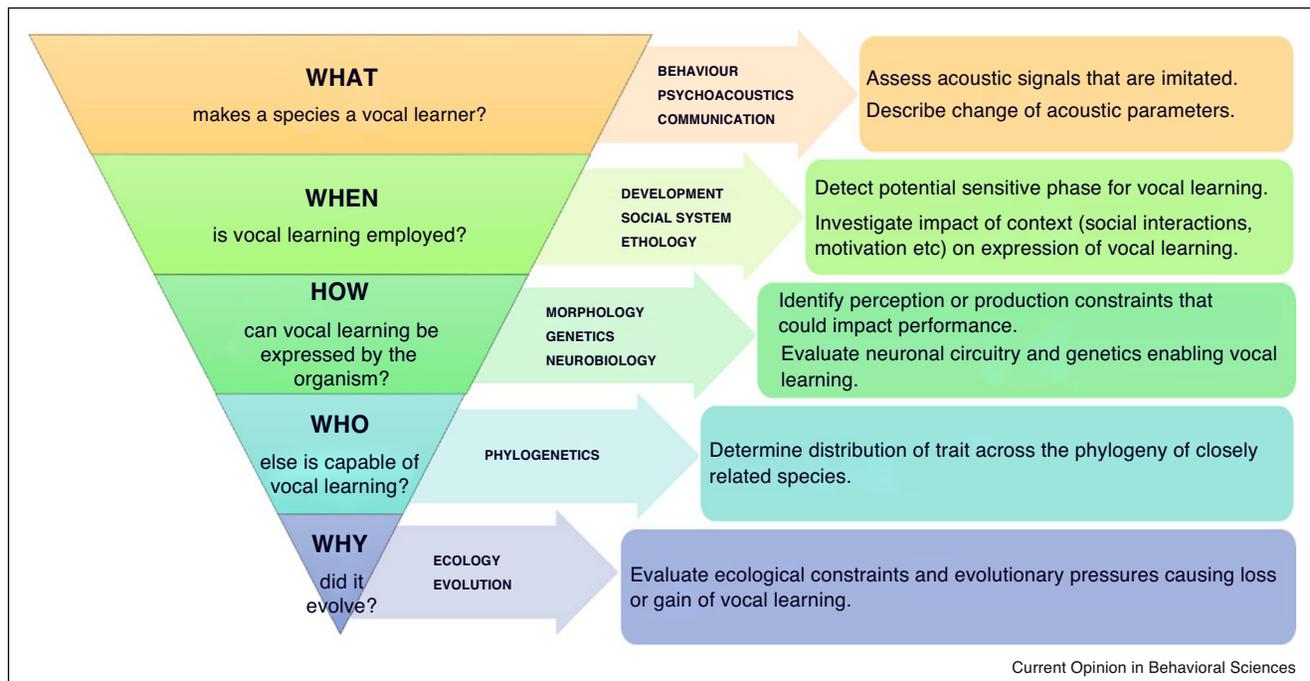
each species falls on this hypothetical continuum — including nonlearning species — will provide much greater insight into shared and divergent characteristics of each animal than a bimodal classification system, and thus greater potential for success in the comparative approach.

A proposed theoretical framework

To extract the most benefit from the comparative approach, a theoretical framework should be established, which enables the structured search for parallels and divergences between species at multiple levels and ultimately allows the full understanding of the communicative system and its evolution. Here, we attempt to develop a framework that could be built upon to facilitate

4 The evolution of language

Figure 3



A proposed theoretical framework: levels at which vocal learning can be studied and associated exemplary questions. The field of vocal learning is of interest for different research areas and is thus approached from different angles. The detailed study of one organism on all these levels (e.g. zebra finches) is important and leads to the in-depth understanding of a confined model system. However, comprehensive across species comparisons within each of these levels will advance the understanding of the evolution and transferable requirements for vocal learning as a whole.

a focused research approach when studying species for their vocal learning ability (Figure 3).

For each species the following could be considered and compared:

1. **WHAT** makes a species a vocal learner? Which acoustic signals do they imitate and how do they change their vocalisations?

Some animals are capable of copying only species-specific sounds, for example, zebra finches learn a short stereotyped song that they faithfully repeat throughout their lifetime [28,29]. By comparison, nightingales learn a massive repertoire of sounds that they combine into song in a flexible manner [30]. Not all forms of vocal learning involve learning from conspecifics as some species, for example, mockingbirds, are capable of mimicking the songs of other animals, and some particularly talented parrots, elephants, and seals can even mimic human speech (with varying degrees of success) [19,21,31,32]. Thus it is important to consider what type of signal an animal is capable of learning and to what degree the vocal repertoire can be modified. At one end of the spectrum vocal modifications may involve subtle changes to the structure and spectral content of calls that always retain species-specific

characteristics. At the other extreme are animals that can learn a dazzling array of vocalisations with seemingly few restrictions.

2. **WHEN** is vocal learning employed? Are they able to perform vocal learning during their whole life time or only in specific sensitive phases? Do they require a certain type of context, such as conspecific interaction (e.g. parental influence) or motivation (e.g. presence/absence of social input)?

Exploring vocal learning abilities at the appropriate ontogenic stage for each individual species is crucial to avoid false-negative classifications. Many animals have been shown to have a flexible learning period early in life, but lose the ability to imitate new sounds later on (open- vs. closed-ended learners) [14]. The developmental stage of a potential model species is thus often a sensitive and limiting factor in the study of vocal learning as it is possible to overlook vocal learning if an organism is investigated before it starts producing imitations (e.g. in the memorising phase), or likewise when the sensitive phase for auditory template learning is already over.

In many species only one of the sexes actually uses the learned vocalisations, for example, in zebra finches, juvenile male birds learn from an adult male tutor [14,28]. Social factors can also determine vocal

learning, for example, the superb starling shows no difference in male and female song, but rather differences in song are attributed to social rank [33]. In bats where male–male competition is often strongly expressed and male song is used for territorial defence and mate attraction [34^{*}], vocal learning might also be restricted to males, but this assumption should be rigorously tested in both sexes for bats, and for any other potential vocal learning species.

Motivation, influenced by the social environment as well as other factors could also influence the ability of an animal to learn vocalisations. The first reported instance of vocal learning in elephants was hypothesised to have occurred because of deprivation of conspecific company at a particularly sensitive time point during development [19]. Social interactions or a conspecific rich/poor environment might therefore have a strong influence on the expression, and thus visibility, of vocal learning. If the animal is not sufficiently motivated or there is not enough pressure on the animal to learn and modify their vocal outputs, an animal may appear unable to learn vocalisations, even though the required genetics, morphological, and neuronal structures are present.

3. HOW can vocal learning be expressed by the organism? Are they subject to physical constraints that impact their auditory perception? How do neural circuitry and gene expression facilitate the required complex processes? How precise is the control of the vocal production apparatus and are there morphological/production constraints?

Morphological (e.g. vocal and/or auditory apparatus), neurobiological and genetic factors shape the biology of an animal, allowing it to perform the complex task of vocal learning. Studies comparing these factors across songbirds and humans have suggested deep-homology in neural circuitry (e.g. cortico-basal ganglia circuits) and genetics (e.g. the *FOXP2* gene) that point to convergent evolutionary mechanisms [5,12,35]. However, investigation of these features in a broad range of species is needed to determine if these mechanisms are common to all vocal learners. Identifying morphology, gene pathways, and brain structures in one declared vocal learner group is not enough to declare them a prerequisite for vocal learning in general. For example, the notion of ‘the structure of a vocal learning brain’ is to be taken with caution as it is currently, by necessity, largely based on data from a limited number of closely related bird species [12,17^{**},36]. In order to find structures, circuits, and genes that enable vocal learning, these traits need to be compared at similar resolution across a range of vocal learners, preferably across distinct evolutionary clades.

Such studies can also point to reasons why some species are not vocal learners. For example, the morphological differences in the vocal tract of humans and nonhuman primates have long been

posited as a limiting factor in the ability of primates to produce human-like vocalisations. Recent work has shown however that a nonhuman primate vocal tract is capable of producing speech-like sounds showing that it is instead likely that differences in the neuronal wiring underlying vocal learning and vocal-motor control caused the gap in vocal learning capacity between humans and our closest primate relatives [37^{**},38].

4. WHO else is capable of vocal learning? Which of the close relatives of these species shows comparable vocal learning abilities? How is this trait distributed in the phylogeny of the focal species?

As can be seen from Figure 2, animals with evidence for vocal learning are spread across the evolutionary tree. However, outside of birds, evidence often comes from one or a few animal species within an order. For example, four species of bats have good evidence as vocal learners, another handful of species are thought to be promising for the study of vocal learning, but to date no bat species has been confirmed as a vocal nonlearner [22^{**},23^{**},34^{*}]. Identifying the distribution of this trait (in its different incarnations as discussed above), and determining which species do not display the trait, will be crucial to pinpointing the biological factors necessary and sufficient for vocal learning and for understanding how this trait evolved.

5. WHY did vocal learning evolve? Which ecological conditions and resulting evolutionary pressures cause a species to gain or lose the capacity for vocal learning? Previously several evolutionary benefits of vocal learning have been proposed [39]. However, which ecological conditions support sexual selection for the generation of a broader call repertoire or create the need for increased information-sharing will have to be investigated across a diverse range of species-specific habitat and living conditions. Taken together, understanding which species have the capacity for vocal learning, how they express the trait at a behavioural level, and the biological factors that underlie this trait will shed light on the evolutionary pressures that caused species to gain or lose this complex and evolutionarily expensive trait.

While this framework is by no means exhaustive, if these and related questions are addressed for every species studied for their ability to perform vocal learning, a comparative approach could yield a detailed and informed assignment of different species along the vocal learning continuum. We emphasise that not all of the above-mentioned levels of vocal learning need to be addressed for the comparative approach to work, but they could serve as a research guide. A framework such as this will be key to allow metaanalyses across multiple species, which will be crucial for a clear understanding of the evolution and biological encoding of vocal learning, human speech, and spoken language.

6 The evolution of language

Advantages of the broad comparative approach

With a nuanced approach to the multidimensional continuum of vocal learning abilities and more in-depth understanding of the contributing biological factors, will also come the ability to understand the larger forces shaping this trait over evolutionary time. Several traits have previously been identified as frequently co-occurring with, and thus hypothesised to be important for, the evolution of vocal learning (Figure 1), but comparative studies are absolutely essential to test such hypotheses. With a broader comparative approach that studies the array of vocal learning abilities in a wide range of species we can begin to answer questions like: Is rhythm perception and rhythmicity an evolutionary prerequisite for, or by-product of, vocal learning [20^{*},40,41^{**},42]? How important are the roles of neuronal and motor timing [11^{**},43,44]? Is it a coincidence that several established vocal learners use echolocation [45]? Is the co-occurrence of vocal learning and turn-taking incidental or required [46,47]? Is vocal learning facilitated by, or even a necessary consequence of, specific social structures in vertebrate groups [48,49^{*},50]? And finally what are the factors necessary and sufficient for the emergence of vocal learning? Questions like these cannot be answered by studying only one model system or a small number of phylogenetically closely related species, but must be investigated in many species across multiple different taxa.

Therefore, we here reiterate the need for a broader comparative approach applied to the study of vocal learning, in which species are compared depending on their evolutionary history, their ecological and morphological constraints, and their neuronal and genetic underpinnings. Only comparative studies between species with either, shared expression of vocal learning and different underlying mechanisms or shared evolutionary constraints and similar morphological and neurogenetic setup but varying degree of vocal learning, will lead to an in-depth understanding of all factors that are necessary and sufficient for the evolution of vocal learning.

Conflict of interest statement

Nothing declared.

Acknowledgements

This work was funded by a Marie Curie Career Integration Grant (PCIG12-GA-2012-333978), a Max Planck Research Group Award, and a Human Frontiers Science Program (HFSP) Research grant (RGP0058/2016), all awarded to S.C.V. We would like to thank Christina Bergmann for thoughtful comments on the manuscript.

References

1. Rasilo H, Räsänen O: **An online model for vowel imitation learning.** *Speech Commun* 2017, **86**:1-23.
The authors demonstrate a new computational model for infant vowel learning, using the virtual model of a child and human trainers to show how audio-vocal input can quickly lead to robust vowel imitation.

2. Lipkind D, Marcus GF, Bemis DK, Sasahara K, Jacoby N, Takahasi M, Suzuki K, Feher O, Ravbar P, Okanoya K *et al.*: **Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants.** *Nature* 2013, **498**:104-109.
3. Lee SH, Hsiao TY, Lee GS: **Audio-vocal responses of vocal fundamental frequency and formant during sustained vowel vocalizations in different noises.** *Hear Res* 2015, **324**:1-6.
4. Mitsuya T, Munhall KG, Purcell DW: **Modulation of auditory-motor learning in response to formant perturbation as a function of delayed auditory feedback.** *J Acoust Soc Am* 2017, **141**:2758-2767.
5. Enard W: **FOXP2 and the role of cortico-basal ganglia circuits of speech and language evolution.** *Curr Opin Neurobiol* 2011, **21**:415-424.
6. Fitch WT: **Empirical approaches to the study of language evolution.** *Psychon Bull Rev* 2017, **24**:3-33.
7. Boughman JW, Moss CF: **Social sounds: vocal learning and development of mammal and bird calls.** In *Acoustic Communication—Springer Handbook of Auditory Research*. Edited by Simmons AM, Fay RR, Popper AN. New York: Springer-Verlag; 2003:138-224.
8. Janik VM, Slater PJB: **The different roles of social learning in vocal communication.** *Anim Behav* 2000, **60**:1-11.
9. Fitch WT: **The evolution of speech: a comparative review.** *Trends Cogn Sci* 2000, **4**:258-267.
10. Wheatcroft D, Qvarnström A: **A blueprint for vocal learning: auditory predispositions from brains to genomes.** *Biol Lett* 2015, **11**:20150155.
11. Roberts TF, Hisey E, Tanaka M, Kearney MG, Chattree G, Yang CF, Shah NM, Mooney R: **Identification of a motor-to-auditory pathway important for vocal learning.** *Nat Neurosci* 2017, **20**:978-986.
12. Pfenning AR, Hara E, Whitney O, Rivas MV, Wang R, Rouhac PL, Howard JT, Wirthlin M, Lovell PV, Ganapathy G *et al.*: **Convergent transcriptional specializations in the brains of humans and song-learning birds.** *Science* 2014, **346**:1-13.
13. Wada K, Sakaguchi H, Jarvis ED, Hagiwara M: **Differential expression of glutamate receptors in avian neural pathways for learned vocalization.** *J Comp Neurol* 2004, **476**:44-64.
14. Brainard MS, Doupe AJ: **What songbirds teach us about learning.** *Nature* 2002, **417**:351-358.
15. Elemans CPH: **The singer and the song: the neuromechanics of avian sound production.** *Curr Opin Neurobiol* 2014, **28**:172-178.
16. Agate RJ, Scott BB, Haripal B, Lois C, Nottebohm F: **Transgenic songbirds offer an opportunity to develop a genetic model for vocal learning.** *Proc Natl Acad Sci U S A* 2009, **106**:17963-17967.
17. Elemans CPH, Rasmussen JH, Herbst CT, Düring DN, Zollinger SA, Brumm H, Srivastava K, Svane N, Ding M, Larsen ON *et al.*: **Universal mechanisms of sound production and control in birds and mammals.** *Nat Commun* 2015, **6**:1-13.

This study gives insights in the vocal motor processes and kinematics of avian vocal production and shows that syringeal sound production in birds is based on a myoelastic-aerodynamic system as it is in mammals.

18. Janik VM: **Cetacean vocal learning and communication.** *Curr Opin Neurobiol* 2014, **28**:60-65.

This review is a concise summary of vocal learning in cetaceans, combining case studies and conceptual relationships. A comparative approach to vocal learning is promoted as the author draws comparisons between cetacean and primate communication systems.

19. Stoeger AS, Mietchen D, Oh S, De Silva S, Herbst CT, Kwon S, Fitch WT: **An Asian elephant imitates human speech.** *Curr Biol* 2012, **22**:2144-2148.
20. Ravignani A, Fitch WT, Hanke FD, Heinrich T, Hurgitsch B, Kotz SA, Scharff C, Stoeger AS, Boer Bde: **What pinnipeds have to say about human speech, music, and the evolution of rhythm.** *Front Neurosci* 2016, **10**:1-9.
- The authors of this review summarise findings about pinnipeds as model species for rhythmicity and vocal learning. They draw comparisons to studies in other vertebrates and address the topic of the origin of speech and rhythmicity.
21. Schusterman RJ: **Vocal learning in mammals with special emphasis on pinnipeds.** In *The evolution of communicative flexibility: complexity, creativity, and adaptability in human and animal communication.* Edited by Oller DK, Gribel U. MIT Press; 2008:41-70.
22. Knörnschild M: **Vocal production learning in bats.** *Curr Opin Neurobiol* 2014, **28**:80-85.
- This article is an excellent review of vocal production learning in bats, summarizing findings of the past thirty years. Proximate and ultimate mechanisms for vocal production learning are discussed and bats are conclusively suggested as model species for vocal learning.
23. Prat Y, Taub M, Yovel Y: **Vocal learning in a social mammal: demonstrated by isolation and playback experiments in bats.** *Sci Adv* 2015:e1500019.
- The authors amass an extensive library of *Rousettus aegyptiacus* vocalisations. By following the vocal development of juvenile bats and using isolation and playback experiments, they suggest that these bats may perform vocal learning.
24. Boughman JW: **Vocal learning by greater spear-nosed bats.** *Proc R Soc B* 1998, **265**:227-233.
25. Knörnschild M, Nagy M, Metz M, Mayer F, Von Helversen O: **Complex vocal imitation during ontogeny in a bat.** *Biol Lett* 2010, **6**:156-159.
26. Knörnschild M, Nagy M, Metz M, Mayer F, Von Helversen O: **Learned vocal group signatures in the polygynous bat *Saccopteryx bilineata*.** *Anim Behav* 2012, **84**:761-769.
27. Petkov CI, Jarvis ED: **Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates.** *Front Evol Neurosci* 2012, **4**:1-24.
28. Tchernichovski O, Mitra PP, Lints T, Nottebohm F: **Dynamics of the vocal imitation process: how a zebra finch learns its song.** *Science* 2001, **291**:2564-2569.
29. Slater PJB, Eales LA, Clayton NS: **Song learning in zebra finches (*Taeniopygia guttata*): progress and prospects.** *Adv Study Behav* 1988, **18**:1-34.
30. Todt D, Hultsch H, Heike D: **Conditions affecting song acquisition in nightingales (*Luscinia megarhynchos* L.).** *Ethology* 1979, **51**:23-35.
31. Pepperberg IM: **Vocal learning in grey parrots: a brief review of perception, production, and cross-species comparisons.** *Brain Lang* 2010, **115**:81-91.
32. Gammon DE, Altizer CE: **Northern mockingbirds produce syntactical patterns of vocal mimicry that reflect taxonomy of imitated species.** *J Field Ornithol* 2011, **82**:158-164.
33. Pilowsky JA, Rubenstein DR: **Social context and the lack of sexual dimorphism in song in an avian cooperative breeder.** *Anim Behav* 2013, **85**:709-714.
34. Smotherman M, Knörnschild M, Smarsh G, Bohn K: **The origins and diversity of bat songs.** *J Comp Physiol A* 2016, **202**:535-554.
- This comprehensive review allows insight into singing behaviour in bats with a focus on the occurrence of this trait across taxa and evolutionary cost-benefit considerations, but it also discusses sampling and reporting difficulties. The authors draw parallels to bird song, referring to the similarities in life style and ecology.
35. Fisher SE, Scharff C: **FOXP2 as a molecular window into speech and language.** *Trends Genet* 2009, **25**:166-177.
36. Nelson DA, Marler P, Palleroni A: **A comparative approach to vocal learning: intraspecific variation in the learning process.** *Anim Behav* 1995, **50**:83-97.
37. Fitch WT, de Boer B, Mathur N, Ghazanfar AA: **Monkey vocal tracts are speech-ready.** *Sci Adv* 2016, **2**:1-7.
- With modern X-ray video techniques this study demonstrated that, contrary to long standing opinion, monkeys have a speech-ready vocal tract and their inability to perform vocal learning stems instead from a lack of neuronal potential to execute this task.
38. Boë L-J, Berthommier F, Legou T, Captier G, Kemp C, Sawallis TR, Becker Y, Rey A, Fagot J: **Evidence of a vocalic proto-system in the baboon (*Papio papio*) suggests pre-hominin speech precursors.** *PLOS ONE* 2017, **12**:e0169321.
39. Nowicki S, Searcy WA: **The evolution of vocal learning.** *Curr Opin Neurobiol* 2014, **28**:48-53.
40. Schachner A, Brady TF, Pepperberg IM, Hauser MD: **Spontaneous motor entrainment to music in multiple vocal mimicking species.** *Curr Biol* 2009, **19**:831-836.
41. Spierings MJ, Ten Cate C: **Zebra finches as a model species to understand the roots of rhythm.** *Front Neurosci* 2016, **10**:1-3.
- This concise review focuses on vocal timing in zebra finch vocalisations. Zebra finches are presented as suitable model species for the study of produced and perceived vocal rhythmicity.
42. Patel AD, Iversen JR, Bregman MR, Schulz I: **Experimental evidence for synchronization to a musical beat in a nonhuman animal.** *Curr Biol* 2009, **19**:827-830.
43. Prather JF, Peters S, Nowicki S, Mooney R: **Precise auditory-vocal mirroring in neurons for learned vocal communication.** *Nature* 2008, **451**:305-310.
44. Benichov JI, Benezra SE, Vallentin D, Globerson E, Long MA, Tchernichovski O: **The forebrain song system mediates predictive call timing in female and male zebra finches.** *Curr Biol* 2015, **26**:1-10.
45. Parker J, Tsagkogeorga G, Cotton Ja, Liu Y, Provero P, Stupka E, Rossiter SJ: **Genome-wide signatures of convergent evolution in echolocating mammals.** *Nature* 2013, **502**:1-9.
46. Vernes SC: **What bats have to say about speech and language.** *Psychon Bull Rev* 2017, **24**:111-117.
47. Chow CP, Mitchell JF, Miller CT: **Vocal turn-taking in a non-human primate is learned during ontogeny.** *Proc R Soc B Biol Sci* 2015, **282**.
48. Luo B, Huang X, Li Y, Lu G, Zhao J, Zhang K, Zhao H, Liu Y, Feng J: **Social call divergence in bats: a comparative analysis.** *Behav Ecol* 2017, **28**:533-540.
49. Chen Y, Matheson LE, Sakata JT: **Mechanisms underlying the social enhancement of vocal learning in songbirds.** *Proc Natl Acad Sci U S A* 2016, **113**:6641-6646.
- This study reports the enhancement of vocal learning performance in juvenile zebra finches as an effect of social interactions with a live tutor.
50. Sewall KB: **Social complexity as a driver of communication and cognition.** *Integr Comp Biol* 2015, **55**:384-395.