# The trade-off between vocal learning and dexterity: a balancing act

Pedro Tiago Martins[*] and Cedric Boeckx[1,2,3,4,*]

[1]Universitat de Barcelona
[2]Universitat de Barcelona Institute of Neurosciences
[3]Universitat de Barcelona Institute of Complex Systems
[4]Catalan Institute for Research and Advanced Studies (ICREA)
[*]Correspondence: `ptsgmartins@gmail.com,cedric.boeckx@ub.edu`

May 29, 2023

**Abstract**

Uncontroversial evidence of vocal production learning, the capacity to modify vocal output on the basis of experience, is sparsely distributed in the animal kingdom. We suggest that this is in large part due to a trade-off between vocal learning complexity and a much more widely distributed trait—non-vocal dexterity. We argue that given some generally required neural and anatomical conditions for vocal production learning, species lacking both the (manual) appendages fine-tuned for grasping, manipulation, etc. and the neural control of those structures are more likely to display complex vocal learning. Conversely, the presence and control of these (manual) structures relegates the vocal apparatus to simpler vocal, feeding, and manipulation behaviors in other species. In other words, vocal learners tend to be flyers or swimmers. We also address the obvious exception to this generalization: humans are both highly dexterous and complex vocal learners. We hypothesize that the degree of bipedalism in land species and its connection with locomotion and breathing control is also a factor that helps shape the distribution of the vocal learning phenotype.

**Keywords:** vocal learning, dexterity, bipedalism, locomotion

## 1 A rare trait

Vocal (production) learning is the capacity to modify vocal output on the basis of experience (Martins and Boeckx, 2020). Although evidence accumulates in favor of treating vocal learning as a multidimensional ability, where behavior, neural circuitry and functional pressures all help shape the phenotype (Lameira, 2017; Ghazanfar et al., 2019; Wirthlin et al., 2019; Martins and Boeckx, 2020; Fischer et al., 2020), in its popular characterization as vocal imitation subserved by a direct cortico-laryngeal/syringeal connection (Petkov and Jarvis, 2012), vocal (production) learning remains a rare ability, and has been identified in a few species: humans, three orders of birds (songbirds, parrots, hummingbirds)

(Petkov and Jarvis, 2012), cetaceans (Janik, 2014), pinnipeds (Ravignani et al., 2016), bats (Vernes and Wilkinson, 2020), and elephants (Stoeger and Manger, 2014). Perhaps due to this, vocal learning is often still treated as a binary trait (e.g. Christmas et al., 2023), even if there is promise of vocal learning behavior in more species, including primates (Lameira, 2017; Takahashi et al., 2017; Martins and Boeckx, 2020).

In this paper we isolate a property of canonical vocal production learners that we put forward as a heretofore underappreciated constraint on the emergence of the phenotype. After stating the key observation, we formulate a hypothesis that explains the constraint, and examine (potential) exceptions.

## 2 The key observation

There is one aspect that connects most canonical vocal learners: they are flyers or swimmers. Species whose vocal learning abilities are typically considered less impressive tend to be neither. This relationship was hinted at in Janik and Slater (1997), who pointed out that by virtue of their being flyers and/or swimmers, vocal learners inhabit and navigate 3D spaces, leading to different pressures that foster the emergence of some kind of vocal learning complexity: an increase in vocal versus visual communication signals, signal noise adding to signal diversity, and mate choice and sexual selection in such open-ended environments.

This insight has gone largely unexplored in the vocal learning literature, with the exception of Verpooten (2021), who discussed two key observations regarding selective pressures. One is that sexual selection mechanisms depend on the dimensionality of the mating environment, and monopolization is harder in vast 3D environments (*Dimensionality hypothesis*; Puts, 2010). The three-dimensionality of water, sea, and trees favours mate monopolization methods other than force, with the latter being favored in two-dimensional environments where direct contests are the main method of excluding competitors. The other is that the possibility of escape by females from coercive males fosters ornamental versus weaponized methods of mate attraction (*Coercion-avoidance hypothesis*; Pradhan and Van Schaik, 2009). In vast three-dimensional environments, where females have more behavioral freedom and male monopolization is harder, sexually dimorphic ornaments are favored, while terrestrial species in two-dimensional environments are more dependent on dimorphic weaponry, which can be more readily used to fend off male competitors and coerce female mates. Indeed, as Verpooten (2021) points out, the dynamics of sexual selection in 3D environments, where ornamentation and methods other than force are highly advantageous, might be a very important factor in the emergence of vocal learning complexity.

Valuable as these observations are, we would like to point out that flyers and swimmers have more in common than the vastness and dimensionality of their environment and how they navigate it. In particular, they don't have prehensile forelimbs. Their forelimbs are anatomically and mechanically specialized for swimming and/or flying. We put forth the following generalization: Species that display robust vocal learning abilities tend to lack dexterous forelimbs.

Even non-flyers/swimmers who display complex vocal learning behavior tend not to have forelimb dexterity (elephants), while those species whose vocal learning behavior is more elusive tend to have a great deal of prehensile dexterity. In the next sections, we articulate why we think that this relationship holds. We also turn our attention to apparent counter-examples to this correlation, with the most obvious one being humans, who are both complex vocal learners and have high forelimb dexterity.

# 3 The hypothesis

Dexterity can be understood as the ability to perform fine motor movements: to reach for, grasp and manipulate objects with anatomically suitable appendages such as the hands/paws, digits, or other structures. Much like vocal learning, dexterity has evolved independently in different clades (Nowicki and Searcy, 2014; Iwaniuk and Whishaw, 2000), which highlights the existence of selective pressures leading to its emergence. Also like vocal learning, dexterity is not a monolithic ability; there are different ways in which it can be manifested, which involve different degrees of grasping and object manipulation, which can be employed in a variety of behaviors. Anatomical, allometric, biomechanical and functional factors influence the different grasping skills species will display (Iwaniuk et al., 2000). Importantly, dexterity is not limited to forelimbs. For example, forelimbs in birds are not prehensile and are used almost exclusively for flight, forcing dexterous behavior to the tongue and beak, and in several cases the feet (Sustaita et al., 2013; Gutiérrez-Ibáñez et al., 2023). For instance, parrots, which are notable vocal imitators, have specialized tongue muscles that allow them to perform dexterous tasks (Homberger, 2003). In cetaceans, forelimbs lack dexterity and indeed hand musculature and innervation, and are used chiefly for swimming (Cooper et al., 2007). However, in some cetaceans, the hyolingual apparatus (hyoid bone and tongue) is highly prehensile and allows for complex feeding behavior in an aquatic environment, in the absence of any other dexterous extremity (Werth, 2007).

It seems that, in the absence of available forelimbs, dexterity is to be found in other structures that are anatomically, mechanically and neurally suitable, with the rostrum and adjacent structures taking center-stage. This is also the case for terrestrial species, such as elephants, for example, which have a highly prehensile trunk, which they use for object manipulation and for feeding (Racine, 1980; Kaufmann et al., 2022, 2023).

We think that all this is highly relevant for the emergence of vocal learning. Succinctly put: the lack of anatomically suitable forelimbs has the net effect of relegating dexterity (i.e., fine motor control) to the rostral structures. This leads to a higher degree of development and control of these structures, making it easier for them to be recruited for complex behavior characteristic of vocal production learning. That is to say, low forelimb dexterity pushes towards the elaboration of structures that are beneficial for vocal learning. Species that lack dexterous forelimbs rely on their rostral features (mouth, beak, trunk) for dexterity-related duties, whose musculature, innervations and neural control will develop disproportionally and thus (we hypothesize) be more easily available for (complex) vocal learning. Using the image of adaptive landscape, one could say that forelimb dexterity pushes species away from adaptive peaks associated with vocal production learning.

If correct, our hypothesis adds to our understanding of the relation between neural structures for vocal learning and those used for skilled movement. Feenders et al. (2008) already pointed out that in vocal learning birds, the cerebral vocal learning nuclei are adjacent to discrete brain areas active during limb and body movements, and hypothesized that the brain areas specialized for vocal learning evolved as a specialization of a pre-existing motor pathway that controls movement (possibly via a mechanism of duplication and divergence, Jarvis (2019)). Indeed, anatomically, areas relevant for the control of vocal learning structures are adjacent to portions of the motor cortex devoted to other aspects of dexterity (Simonyan, 2014). And molecular mechanisms underlying the circuit formation of critical forebrain to muscle pathways for vocal learning (Wang et al., 2015) exploit the very same molecular toolkit as the one leading to the fine-motor control of hand movements (Lemon, 2008). Our hypothesis is that the very structures that may be recruited for the development of the neurobiology required for vocal learning skills could act as a constraint on the development of these skills if they themselves require behavioral elaborations, as is the case among non-flyers/swimmers, particularly those species occupying an an arboreal niche, where selective pressures for forelimb manipulation, including arboreal locomotion, digging and prey handling are strong (Iwaniuk and Whishaw, 2000; Whishaw, 2003; Sustaita et al., 2013; Gutiérrez-Ibáñez et al., 2023; Schwartz et al., 2023).

## 4 Necessary refinements

An important consideration is that lack of forelimb dexterity alone does not immediately confer vocal learning ability. After all, although they collectively represent more than half of all bird species, only 3 of about almost 30 bird orders are canonical vocal learners (Petkov and Jarvis, 2012). Also, in mammalian orders and families canonically considered to be vocal learners, not all members have been possess this ability. For example, about half of the bat families are consistent with vocal learning, with around a quarter showing direct evidence (Vernes and Wilkinson, 2020). Similarly, only some cetaceans and pinnipeds have shown vocal learning ability or promise (Janik and Knörnschild, 2021; Ravignani et al., 2016). The case of mammals can, however, more easily be attributed to difficulty in both direct observation and experimental testing, and it is reasonable to assume that the list of accepted vocal learning species will increase for these orders and families. Still, something more than lack of forelimb dexterity must be at work.

Although the absence of forelimb dexterity lifts off an important neurobiological barrier against the evolution of vocal learning, the latter phenotype constitutes a complex behavior that imposes additional demands on the organism, which species are likely to meet in non-uniform ways. Here we list some conditions that facilitate the development and manifestation of the vocal learning phenotype in those flyers and swimmers considered canonical vocal learners. One is the availability of superfast muscles, which in songbirds allow for the syringeal control required for rapid and precise calls (Elemans et al., 2008), and in bats for the laryngeal control required for their extremely high-frequency echolocating calls (Elemans et al., 2011). Though work in still underway in determining bat echolocation ontology (Nojiri et al., 2021), it could be that all bat species are predisposed to develop echolocating abilities (Wang et al., 2017). Hummingbirds also possess superfast muscles, which not only allow for their impressive flight and hover abilities (Reiser et al., 2013), but are also likely at the level of songbirds for syringeal control (Monte et al., 2020). Cetaceans are also echolocating (Janik and Sayigh, 2013), and use their sound production apparatuses for dynamically to modulate both vocal communication signals and echolocation (Madsen et al., 2023). Furthermore, they have functional aspects of their breathing under voluntary control, much more so than other mammals (Fahlman et al., 2017). And while pinnipeds do not echolocate, they have evolved whisker control to accomplish the same goal. They orient, retract and protract their mouth-adjacent whiskers in a rhythmic fashion to sense, locate, and even coerce small prey (Milne et al., 2020; Adachi et al., 2022).

These traits, presumably evolved independently of vocal learning, can be thought of as extremely valuable precursors that can be recruited for complex vocal learning. There are likely to be additional neurobiological requirements for the establishment of complex vocal learning behavior, such as overall encephalization (which may have taken different routes among birds, Ksepka et al. (2020), and which app fears to leave molecular footprints in the motor cortex, Kaplow et al. (2023)), or regional expansion of motor-relevant regions such as the cerebellum (Smaers et al., 2018; Ströckens et al., 2022; Sol et al., 2022), or increase in sheer neuron numbers (Olkowicz et al., 2016). Incidentally, and in line with our hypothesis, manual dexterity appears to impose similar "encephalization"-related requirements (Heldstab et al., 2020).

# 5 The obvious exception(s)

*Homo sapiens* doesn't seem to easily fit the picture presented so far: we are land mammals (i.e., neither flyers nor swimmers), but unlike elephants, we have highly dexterous forelimbs *and* display complex vocal learning behavior. However, there is one key feature that we contend is key in understanding how humans fit in the context of our hypothesis, related to the mode of locomotion: bipedalism. In virtue of our being obligatory bipeds, and having reduced the engagement of the thorax in locomotion, our species has evolved a decoupling of respiratory and locomotive rhythms, with several possible phase ratios depending on the activity (Raßler and Kohl, 2000). This freeing of respiratory rhythm from gait rhythm allows for vocalizations that are mechanically independent from locomotion and the phase ratios it imposes (e.g. Provine, 2017). This independence is a requirement for volitional vocalizations, which functionally rely on controlled modification of acoustic aspects such as frequency, amplitude and duration. Indeed, there is work showing that the human laryngeal motor cortex integrates both laryngeal and respiratory motor control (Belyk and Brown, 2017; Belyk et al., 2021), as opposed to it being strictly laryngeal. Moreover, work on regulatory regions of mammalian genomes showing convergent evolution associated with vocal learning has highlighted the role of genes like *TSHZ3* (Wirthlin et al., 2022) (avian vocal learners display signals of accelerated evolution around the same gene, Zhang et al. (2014)). As Wirthlin et al. (2022) observes, disruptive mutations affecting *TSHZ3* in humans impacts respiratory rhythms and cortico-striatal circuits, critical for learned motor behavior (Caubit et al., 2010, 2016).

A rigid 1:1 coupling between breathing and locomotion as usually observed in quadrupeds (Bramble and Carrier, 1983) would either not allow this integration between larynx and respiration, or lead to much more limited vocal production ability. Interestingly, this kind of locomotive-respiratory decoupling has an analog in avian species: non-vocal learning birds have phase-locked wing and respiratory (and concomitantly vocalization) cycles, while the vocalizations of vocal-learning birds are "emancipated" from respiratory constraints (Berg et al., 2019). While motor control has been highlighted as a plausible necessary evolutionary step towards vocal learning, it seems that, after this system is place, decoupling from the constraints it imposes on vocal behavior might be important to expand vocal learning complexity. Of note, the locomotor developmental program appears to promote vocal learning in juvenile zebra finches (Liu et al., 2022).

These observations suggest that increased breathing control in our lineage as a consequence of bipedalism may have played a major role in the evolution of our vocal production learning capacity, in line with the suggestions in Maclarnon and Hewitt (2004).

Bipedalism also has implications for manual dexterity. As humans became bipedal species, expanding on arboreal locomotive behavior that is still found in living apes (Thorpe et al., 2007), the forelimbs, well developed owing to that same earlier arboreal niche (Crompton et al., 2010; Sustaita et al., 2013), became available for more dexterous behaviors, in line with Falótico and Ottoni (2023). It is this privileged position that we think has allowed humans to be capable of elaborate vocal learning while possessing dexterous forelimbs, circumventing the ecological and anatomical pressures that we claim here place vocal learning and dexterity at odds. The higher the degree of bipedalism, or the more varied the scenarios in which bipedal locomotion or stance can be employed, the higher the potential for more complex dexterous behavior.

It is quite likely that this privileged position went hand-in-hand with the well-attested encephalization trend in the *Homo* lineage (Püschel et al., 2021), allowing for the maintenance of multiple parallel circuits rooted in the motor cortex (Gavrilov and Nieder, 2021), and the expansion of "Broca's" region (Gallardo et al., 2023). The implications of this for multi-modal communication, and the entanglement (Pouw and Fuchs, 2022) of gestures and speech is an important topic for future research.

In light of its purported role in the decoupling of respiration and locomotion and facilitator of forelimb dexterity, we suspect that some degree of bipedalism might be relevant in understanding why other species, besides humans, show some degree of vocal learning. Non-human primates occasionally walk, and while this is done in specific scenarios, (Duarte et al., 2012), with gibbons being the most proficient biped (Vereecke et al., 2006), it seems to be the case that primates that show promise of vocal learning ability also show some non-trivial degree of bipedalism. We believe this could help explain why non-human primates show potential for vocal learning (Lameira, 2017) even though they are dexterous species, on the one hand, but also why this vocal learning behavior is harder to elicit in these species than in humans, on the other.

Likewise, rodents that show some degree of theoretical probability of vocal learning ability also have some degree of bipedalism/bipedal stance. For example, in their study of the relationship between vocal learning acoustic allometry, Ravignani and Garcia (2022) performed phylogenetic regressions and as one of their results identified a handful of rodent species as promising vocal learners (yet unstudied as such). As far as we can tell, all of them show some degree of biped locomotion or stance. Interestingly, many rodents are also dexterous species (e.g., Whishaw and Coles, 1996), though to a lesser extent compared to other mammals such as primates (Gu et al., 2017).

## 6 Conclusion

To sum up, we have highlighted a relationship between forelimb dexterity and vocal production learning and put forward the claim that because of their similar ecological and neurobiological requirements one acts as a constraint on the emergence of the other. Curiously, then, the very same neurobiological resources recruited for vocal learning (Feenders et al., 2008) may act as a developmental barrier if ecological factors impose pressure leading to the elaboration of manual dexterity.

The intuition behind our explanation is reminiscent of the “neurobiological” real-estate conflict (neural Darwinism) put forward by Deacon (1997) to account for the origins of learned vocal behavior. Whereas Deacon posited a competition between the circuits responsible for innate and learned vocalizations (a conflict that he claimed is mitigated by encephalization), we claim that the conflict lies rather in the neurobiological requirements imposed for fine-grained motor control with the hands or with the vocal apparatus.

To be very clear, this is only one of the constraints that make canonical vocal learners so rare in the animal kingdom. We have listed other conditions important for vocal learning, and given the very nature of biology, where complexity reigns (Lewontin, 2000; Wimsatt, 2007), we do not expect a uniform way in which all attested vocal production learners meet them. In the case of humans, and with implications for non-human mammals, we have argued that bipedalism and attendant breathing control was a major factor.

## Funding statement

CB acknowledges support from the Spanish Ministry of Science and Innovation (grant PID2019-107042GB-I00) and Generalitat de Catalunya (2021-SGR-313).

## References

Adachi, T., Naito, Y., Robinson, P. W., Costa, D. P., Hückstädt, L. A., Holser, R. R., Iwasaki, W., and Takahashi, A. (2022). Whiskers as hydrodynamic prey sensors in foraging seals. *Proceedings of the National Academy of Sciences*, 119(25):e2119502119, doi:10.1073/pnas.2119502119.

Belyk, M., Brown, R., Beal, D. S., Roebroeck, A., McGettigan, C., Guldner, S., and Kotz, S. A. (2021). Human larynx motor cortices coordinate respiration for vocal-motor control. *NeuroImage*, 239:118326, doi:10.1016/j.neuroimage.2021.118326.

Belyk, M. and Brown, S. (2017). The origins of the vocal brain in humans. *Neuroscience & Biobehavioral Reviews*, 77(Supplement C):177–193, doi:10.1016/j.neubiorev.2017.03.014.

Berg, K. S., Delgado, S., and Mata-Betancourt, A. (2019). Phylogenetic and kinematic constraints on avian flight signals. *Proceedings of the Royal Society B: Biological Sciences*, 286(1911):20191083, doi:10.1098/rspb.2019.1083. Publisher: Royal Society.

Bramble, D. M. and Carrier, D. R. (1983). Running and Breathing in Mammals. *Science*, 219(4582):251–256, doi:10.1126/science.6849136.

Caubit, X., Gubellini, P., Andrieux, J., Roubertoux, P. L., Metwaly, M., Jacq, B., Fatmi, A., Had-Aissouni, L., Kwan, K. Y., Salin, P., Carlier, M., Liedén, A., Rudd, E., Shinawi, M., Vincent-Delorme, C., Cuisset, J.-M., Lemaitre, M.-P., Abderrehamane, F., Duban, B., Lemaitre, J.-F., Woolf, A. S., Bockenhauer, D., Severac, D., Dubois, E., Zhu, Y., Sestan, N., Garratt, A. N., Kerkerian-Le Goff, L., and Fasano, L. (2016). TSHZ3 deletion causes an autism syndrome and defects in cortical projection neurons. *Nature Genetics*, 48(11):1359–1369, doi:10.1038/ng.3681.

Caubit, X., Thoby-Brisson, M., Voituron, N., Filippi, P., Bevengut, M., Faralli, H., Zanella, S., Fortin, G., Hilaire, G., and Fasano, L. (2010). Teashirt 3 Regulates Development of Neurons Involved in Both Respiratory Rhythm and Airflow Control. *Journal of Neuroscience*, 30(28):9465–9476, doi:10.1523/JNEUROSCI.1765-10.2010.

Christmas, M. J., Kaplow, I. M., Genereux, D. P., Dong, M. X., Hughes, G. M., Li, X., Sullivan, P. F., Hindle, A. G., Andrews, G., Armstrong, J. C., Bianchi, M., Breit, A. M., Diekhans, M., Fanter, C., Foley, N. M., Goodman, D. B., Goodman, L., Keough, K. C., Kirilenko, B., Kowalczyk, A., Lawless, C., Lind, A. L., Meadows, J. R. S., Moreira, L. R., Redlich, R. W., Ryan, L., Swofford, R., Valenzuela, A., Wagner, F., Wallerman, O., Brown, A. R., Damas, J., Fan, K., Gatesy, J., Grimshaw, J., Johnson, J., Kozyrev, S. V., Lawler, A. J., Marinescu, V. D., Morrill, K. M., Osmanski, A., Paulat, N. S., Phan, B. N., Reilly, S. K., Schäffer, D. E., Steiner, C., Supple, M. A., Wilder, A. P., Wirthlin, M. E., Xue, J. R., Zoonomia Consortium§, Birren, B. W., Gazal, S., Hubley, R. M., Koepfli, K.-P., Marques-Bonet, T., Meyer, W. K., Nweeia, M., Sabeti, P. C., Shapiro, B., Smit, A. F. A., Springer, M. S., Teeling, E. C., Weng, Z., Hiller, M., Levesque, D. L., Lewin, H. A., Murphy, W. J., Navarro, A., Paten, B., Pollard, K. S., Ray, D. A., Ruf, I., Ryder, O. A., Pfenning, A. R., Lindblad-Toh, K., Karlsson, E. K., Andrews, G., Armstrong, J. C., Bianchi, M., Birren, B. W., Bredemeyer, K. R., Breit, A. M., Christmas, M. J., Clawson, H., Damas, J., Di Palma, F., Diekhans, M., Dong, M. X., Eizirik, E., Fan, K., Fanter, C., Foley, N. M., Forsberg-Nilsson, K., Garcia, C. J., Gatesy, J., Gazal, S., Genereux, D. P., Goodman, L., Grimshaw, J., Halsey, M. K., Harris, A. J., Hickey, G., Hiller, M., Hindle, A. G., Hubley, R. M., Hughes, G. M., Johnson, J., Juan, D., Kaplow, I. M., Karlsson, E. K., Keough, K. C., Kirilenko, B., Koepfli, K.-P., Korstian, J. M., Kowalczyk, A., Kozyrev, S. V., Lawler, A. J., Lawless, C., Lehmann, T., Levesque, D. L., Lewin, H. A., Li, X., Lind, A., Lindblad-Toh, K., Mackay-Smith, A., Marinescu, V. D., Marques-Bonet, T., Mason, V. C., Meadows, J. R. S., Meyer, W. K., Moore, J. E., Moreira, L. R., Moreno-Santillan, D. D., Morrill, K. M., Muntané, G., Murphy, W. J., Navarro, A., Nweeia, M., Ortmann, S., Osmanski, A., Paten, B., Paulat, N. S., Pfenning, A. R., Phan, B. N., Pollard, K. S., Pratt, H. E., Ray, D. A., Reilly, S. K., Rosen, J. R., Ruf, I., Ryan, L., Ryder, O. A., Sabeti, P. C., Schäffer, D. E., Serres, A., Shapiro, B., Smit, A. F. A., Springer, M., Srinivasan, C., Steiner, C., Storer, J. M., Sullivan, K. A. M., Sullivan, P. F., Sundström, E., Supple, M. A., Swofford, R., Talbot, J.-E., Teeling, E., Turner-Maier, J., Valenzuela, A., Wagner, F., Wallerman, O., Wang, C., Wang, J., Weng, Z., Wilder, A. P., Wirthlin, M. E., Xue, J. R., and Zhang, X. (2023). Evolutionary constraint and innovation across hundreds of placental mammals. *Science*, 380(6643):eabn3943, doi:10.1126/science.abn3943.

Cooper, L. N., Dawson, S. D., Reidenberg, J. S., and Berta, A. (2007). Neuromuscular Anatomy and Evolution of the Cetacean Forelimb. *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology*, 290(9):1121–1137, doi:10.1002/ar.20571.

Crompton, R. H., Sellers, W. I., and Thorpe, S. K. S. (2010). Arboreality, terrestriality and bipedalism. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1556):3301–3314, doi:10.1098/rstb.2010.0035.

Deacon, T. W. (1997). *The Symbolic Species*. Norton, New York.

Duarte, M., Hanna, J., Sanches, E., Liu, Q., and Fragaszy, D. (2012). Kinematics of bipedal locomotion while carrying a load in the arms in bearded capuchin monkeys (Sapajus libidinosus). *Journal of Human Evolution*, 63(6):851–858, doi:10.1016/j.jhevol.2012.10.002.

Elemans, C. P. H., Mead, A. F., Jakobsen, L., and Ratcliffe, J. M. (2011). Superfast Muscles Set Maximum Call Rate in Echolocating Bats. *Science*, 333(6051):1885–1888, doi:10.1126/science.1207309.

Elemans, C. P. H., Mead, A. F., Rome, L. C., and Goller, F. (2008). Superfast Vocal Muscles Control Song Production in Songbirds. *PLoS ONE*, 3(7):e2581, doi:10.1371/journal.pone.0002581.

Fahlman, A., Moore, M. J., and Garcia-Parraga, D. (2017). Respiratory function and mechanics in pinnipeds and cetaceans. *Journal of Experimental Biology*, 220(10):1761–1773, doi:10.1242/jeb.126870.

Falótico, T. and Ottoni, E. B. (2023). Greater tool use diversity is associated with increased terrestriality in wild capuchin monkeys. *American Journal of Biological Anthropology*, page ajpa.24740, doi:10.1002/ajpa.24740.

Feenders, G., Liedvogel, M., Rivas, M., Zapka, M., Horita, H., Hara, E., Wada, K., Mouritsen, H., and Jarvis, E. D. (2008). Molecular Mapping of Movement-Associated Areas in the Avian Brain: A Motor Theory for Vocal Learning Origin. *PLOS ONE*, 3(3):e1768, doi:10.1371/journal.pone.0001768.

Fischer, J., Wegdell, F., Trede, F., Dal Pesco, F., and Hammerschmidt, K. (2020). Vocal convergence in a multi-level primate society: insights into the evolution of vocal learning. *Proceedings of the Royal Society B: Biological Sciences*, 287(1941):20202531, doi:10.1098/rspb.2020.2531. Publisher: Royal Society.

Gallardo, G., Eichner, C., Sherwood, C. C., Hopkins, W. D., Anwander, A., and Friederici, A. D. (2023). Uncovering the Morphological Evolution of Language-Relevant Brain Areas. preprint, Neuroscience.

Gavrilov, N. and Nieder, A. (2021). Distinct neural networks for the volitional control of vocal and manual actions in the monkey homologue of Broca's area. *eLife*, 10:e62797, doi:10.7554/eLife.62797.

Ghazanfar, A. A., Liao, D. A., and Takahashi, D. Y. (2019). Volition and learning in primate vocal behaviour. *Animal Behaviour*, 151:239–247, doi:10.1016/j.anbehav.2019.01.021.

Gu, Z., Kalambogias, J., Yoshioka, S., Han, W., Li, Z., Kawasawa, Y. I., Pochareddy, S., Li, Z., Liu, F., Xu, X., Wijeratne, H. R. S., Ueno, M., Blatz, E., Salomone, J., Kumanogoh, A., Rasin, M.-R., Gebelein, B., Weirauch, M. T., Sestan, N., Martin, J. H., and Yoshida, Y. (2017). Control of species-dependent cortico-motoneuronal connections underlying manual dexterity. *Science*, 357(6349):400–404, doi:10.1126/science.aan3721.

Gutiérrez-Ibáñez, C., Amaral-Peçanha, C., Iwaniuk, A. N., Wylie, D. R., and Baron, J. (2023). The evolution of skilled hindlimb movements in birds: A citizen science approach. preprint, Animal Behavior and Cognition.

Heldstab, S. A., Isler, K., Schuppli, C., and Schaik, C. P. v. (2020). When ontogeny recapitulates phylogeny: Fixed neurodevelopmental sequence of manipulative skills among primates. *Science Advances*, 6(30):eabb4685, doi:10.1126/sciadv.abb4685. Publisher: American Association for the Advancement of Science Section: Research Article.

Homberger, D. G. (2003). The comparative biomechanics of a prey-predator relationship: the adaptive morphologies of the feeding apparatus of Australian Black-Cockatoos and their foods as a basis for the reconstruction of the evolutionary history of the Psittaciformes. In Bels, V. L., Gasc, J.-P., and Casinos, A., editors, *Vertebrate Biomechanics and Evolution*, pages 203–228. BIOS, Oxford. Publisher: BIOS Scientific Publishers Oxford.

Iwaniuk, A. N., Pellis, S. M., and Whishaw, I. Q. (2000). The relative importance of body size, phylogeny, locomotion, and diet in the evolution of forelimb dexterity in fissiped carnivores (Carnivora). *Canadian Journal of Zoology*, 78(7):1110–1125, doi:10.1139/z00-023.

Iwaniuk, A. N. and Whishaw, I. Q. (2000). On the origin of skilled forelimb movements. *Trends in Neurosciences*, 23(8):372–376, doi:10.1016/S0166-2236(00)01618-0.

Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology*, 28:60–65, doi:10.1016/j.conb.2014.06.010.

Janik, V. M. and Knörnschild, M. (2021). Vocal production learning in mammals revisited. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1836):20200244, doi:10.1098/rstb.2020.0244.

Janik, V. M. and Sayigh, L. S. (2013). Communication in bottlenose dolphins: 50 years of signature whistle research. *Journal of Comparative Physiology A*, 199(6):479–489, doi:10.1007/s00359-013-0817-7.

Janik, V. M. and Slater, P. J. (1997). Vocal Learning in Mammals. In *Advances in the Study of Behavior*, volume 26, pages 59–99. Elsevier.

Jarvis, E. D. (2019). Evolution of vocal learning and spoken language. *Science*, 366(6461):50–54, doi:10.1126/science.aax0287.

Kaplow, I. M., Lawler, A. J., Schäffer, D. E., Srinivasan, C., Sestili, H. H., Wirthlin, M. E., Phan, B. N., Prasad, K., Brown, A. R., Zhang, X., Foley, K., Genereux, D. P., Zoonomia Consortium**, Karlsson, E. K., Lindblad-Toh, K., Meyer, W. K., Pfenning, A. R., Andrews, G., Armstrong, J. C., Bianchi, M., Birren, B. W., Bredemeyer, K. R., Breit, A. M., Christmas, M. J., Clawson, H., Damas, J., Di Palma, F., Diekhans, M., Dong, M. X., Eizirik, E., Fan, K., Fanter, C., Foley, N. M., Forsberg-Nilsson, K., Garcia, C. J., Gatesy, J., Gazal, S., Genereux, D. P., Goodman, L., Grimshaw, J., Halsey, M. K., Harris, A. J., Hickey, G., Hiller, M., Hindle, A. G., Hubley, R. M., Hughes, G. M., Johnson, J., Juan, D., Kaplow, I. M., Karlsson, E. K., Keough, K. C., Kirilenko, B., Koepfli, K.-P., Korstian, J. M., Kowalczyk, A., Kozyrev, S. V., Lawler, A. J., Lawless, C., Lehmann, T., Levesque, D. L., Lewin, H. A., Li, X., Lind, A., Lindblad-Toh, K., Mackay-Smith, A., Marinescu, V. D., Marques-Bonet, T., Mason, V. C., Meadows, J. R. S., Meyer, W. K., Moore, J. E., Moreira, L. R., Moreno-Santillan, D. D., Morrill, K. M., Muntané, G., Murphy, W. J., Navarro, A., Nweeia, M., Ortmann, S., Osmanski, A., Paten, B., Paulat, N. S., Pfenning, A. R., Phan, B. N., Pollard, K. S., Pratt, H. E., Ray, D. A., Reilly, S. K., Rosen, J. R., Ruf, I., Ryan, L., Ryder, O. A., Sabeti, P. C., Schäffer, D. E., Serres, A., Shapiro, B., Smit, A. F. A., Springer, M., Srinivasan, C., Steiner, C., Storer, J. M., Sullivan, K. A. M., Sullivan, P. F., Sundström, E., Supple, M. A., Swofford, R., Talbot, J.-E., Teeling, E., Turner-Maier, J., Valenzuela, A., Wagner, F., Wallerman, O., Wang, C., Wang, J., Weng, Z., Wilder, A. P., Wirthlin, M. E., Xue, J. R., and Zhang, X. (2023). Relating enhancer genetic variation across mammals to complex phenotypes using machine learning. *Science*, 380(6643):eabm7993, doi:10.1126/science.abm7993.

Kaufmann, L. V., Becker, R., Ochs, A., and Brecht, M. (2023). Elephant banana peeling. *Current Biology*, 33(7):R257–R258, doi:10.1016/j.cub.2023.02.076.

Kaufmann, L. V., Schneeweiß, U., Maier, E., Hildebrandt, T., and Brecht, M. (2022). Elephant facial motor control. *Science Advances*, 8(43):eabq2789, doi:10.1126/sciadv.abq2789.

Ksepka, D. T., Balanoff, A. M., Smith, N. A., Bever, G. S., Bhullar, B.-A. S., Bourdon, E., Braun, E. L., Burleigh, J. G., Clarke, J. A., Colbert, M. W., Corfield, J. R., Degrange, F. J., De Pietri, V. L., Early, C. M., Field, D. J., Gignac, P. M., Gold, M. E. L., Kimball, R. T., Kawabe, S., Lefebvre, L., Marugán-Lobón, J., Mongle, C. S., Morhardt, A., Norell, M. A., Ridgely, R. C., Rothman, R. S., Scofield, R. P., Tambussi, C. P., Torres, C. R., Van Tuinen, M., Walsh, S. A., Watanabe, A., Witmer, L. M., Wright, A. K., Zanno, L. E., Jarvis, E. D., and Smaers, J. B. (2020). Tempo and Pattern of Avian Brain Size Evolution. *Current Biology*, 30(11):2026–2036.e3, doi:10.1016/j.cub.2020.03.060.

Lameira, A. R. (2017). Bidding evidence for primate vocal learning and the cultural substrates for speech evolution. *Neuroscience & Biobehavioral Reviews*, 83:429–439, doi:10.1016/j.neubiorev.2017.09.021.

Lemon, R. N. (2008). Descending Pathways in Motor Control. *Annual Review of Neuroscience*, 31(1):195–218, doi:10.1146/annurev.neuro.31.060407.125547.

Lewontin, R. (2000). *The Triple Helix: Gene, Organism, and Environment*. Harvard University Press, Cambridge, MA.

Liu, W.-c., Landstrom, M., Cealie, M., and MacKillop, I. (2022). A juvenile locomotor program promotes vocal learning in zebra finches. *Communications Biology*, 5(1):573, doi:10.1038/s42003-022-03533-3.

Maclarnon, A. and Hewitt, G. (2004). Increased breathing control: Another factor in the evolution of human language. *Evolutionary Anthropology: Issues, News, and Reviews*, 13(5):181–197, doi:10.1002/evan.20032. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/evan.20032.

Madsen, P. T., Siebert, U., and Elemans, C. P. H. (2023). Toothed whales use distinct vocal registers for echolocation and communication. *Science*, 379(6635):928–933, doi:10.1126/science.adc9570.

Martins, P. T. and Boeckx, C. (2020). Vocal learning: Beyond the continuum.

*PLOS Biology*, 18(3):e3000672, doi:10.1371/journal.pbio.3000672.

Milne, A. O., Smith, C., Orton, L. D., Sullivan, M. S., and Grant, R. A. (2020). Pinnipeds orient and control their whiskers: a study on Pacific walrus, California sea lion and Harbor seal. *Journal of Comparative Physiology A*, 206(3):441–451, doi:10.1007/s00359-020-01408-8.

Monte, A., Cerwenka, A. F., Ruthensteiner, B., Gahr, M., and Düring, D. N. (2020). The hummingbird syrinx morphome: a detailed three-dimensional description of the black jacobin's vocal organ. *BMC Zoology*, 5(1):7, doi:10.1186/s40850-020-00057-3.

Nojiri, T., Wilson, L. A., López-Aguirre, C., Tu, V. T., Kuratani, S., Ito, K., Higashiyama, H., Son, N. T., Fukui, D., Sadier, A., Sears, K. E., Endo, H., Kamihori, S., and Koyabu, D. (2021). Embryonic evidence uncovers convergent origins of laryngeal echolocation in bats. *Current Biology*, 31(7):1353–1365.e3, doi:10.1016/j.cub.2020.12.043.

Nowicki, S. and Searcy, W. A. (2014). The evolution of vocal learning. *Current Opinion in Neurobiology*, 28:48–53, doi:10.1016/j.conb.2014.06.007.

Olkowicz, S., Kocourek, M., Lučan, R. K., Porteš, M., Fitch, W. T., Herculano-Houzel, S., and Němec, P. (2016). Birds have primate-like numbers of neurons in the forebrain. *Proceedings of the National Academy of Sciences*, 113(26):7255–7260, doi:10.1073/pnas.1517131113.

Petkov, C. I. and Jarvis, E. D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, 4, doi:10.3389/fnevo.2012.00012.

Pouw, W. and Fuchs, S. (2022). Origins of vocal-entangled gesture. *Neuroscience & Biobehavioral Reviews*, 141:104836, doi:10.1016/j.neubiorev.2022.104836.

Pradhan, G. R. and Van Schaik, C. P. (2009). Why do females find ornaments attractive? The coercion-avoidance hypothesis: Why do females find ornaments attractive? *Biological Journal of the Linnean Society*, 96(2):372–382, doi:10.1111/j.1095-8312.2008.01131.x.

Provine, R. R. (2017). Laughter as an approach to vocal evolution: The bipedal theory. *Psychonomic Bulletin & Review*, 24(1):238–244, doi:10.3758/s13423-016-1089-3.

Puts, D. A. (2010). Beauty and the beast: mechanisms of sexual selection in humans. *Evolution and Human Behavior*, 31(3):157–175, doi:10.1016/j.evolhumbehav.2010.02.005.

Püschel, H. P., Bertrand, O. C., O'Reilly, J. E., Bobe, R., and Püschel, T. A. (2021). Divergence-time estimates for hominins provide insight into encephalization and body mass trends in human evolution. *Nature Ecology & Evolution*, doi:10.1038/s41559-021-01431-1.

Racine, R. N. (1980). Behavior Associated with Feeding in Captive African and Asian Elephants. *Elephant*, 1(5):57–71, doi:10.22237/elephant/1521731845.

Ravignani, A., Fitch, W. T., Hanke, F. D., Heinrich, T., Hurgitsch, B., Kotz, S. A., Scharff, C., Stoeger, A. S., and de Boer, B. (2016). What Pinnipeds Have to Say about Human Speech, Music, and the Evolution of Rhythm. *Frontiers in Neuroscience*, 10, doi:10.3389/fnins.2016.00274.

Ravignani, A. and Garcia, M. (2022). A cross-species framework to identify vocal learning abilities in mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1841):20200394, doi:10.1098/rstb.2020.0394.

Publisher: Royal Society.

Raßler, B. and Kohl, J. (2000). Coordination-related changes in the rhythms of breathing and walking in humans. *European Journal of Applied Physiology*, 82(4):280–288, doi:10.1007/s004210000224.

Reiser, P. J., Welch, K. C., Suarez, R. K., and Altshuler, D. L. (2013). Very low force-generating ability and unusually high temperature-dependency in hummingbird flight muscle fibers. *Journal of Experimental Biology*, page jeb.068825, doi:10.1242/jeb.068825.

Schwartz, E., Nenning, K.-H., Heuer, K., Jeffery, N., Bertrand, O. C., Toro, R., Kasprian, G., Prayer, D., and Langs, G. (2023). Evolution of cortical geometry and its link to function, behaviour and ecology. *Nature Communications*, 14(1):2252, doi:10.1038/s41467-023-37574-x.

Simonyan, K. (2014). The laryngeal motor cortex: its organization and connectivity. *Current opinion in neurobiology*, 28:15–21, doi:10.1016/j.conb.2014.05.006.

Smaers, J. B., Turner, A. H., Gómez-Robles, A., and Sherwood, C. C. (2018). A cerebellar substrate for cognition evolved multiple times independently in mammals. *eLife*, 7:e35696, doi:10.7554/eLife.35696.

Sol, D., Olkowicz, S., Sayol, F., Kocourek, M., Zhang, Y., Marhounová, L., Osadnik, C., Corssmit, E., Garcia-Porta, J., Martin, T. E., Lefebvre, L., and Němec, P. (2022). Neuron numbers link innovativeness with both absolute and relative brain size in birds. *Nature Ecology & Evolution*, 6(9):1381–1389, doi:10.1038/s41559-022-01815-x.

Stoeger, A. S. and Manger, P. (2014). Vocal learning in elephants: neural bases and adaptive context. *Current Opinion in Neurobiology*, 28:101–107, doi:10.1016/j.conb.2014.07.001.

Ströckens, F., Neves, K., Kirchem, S., Schwab, C., Herculano-Houzel, S., and Güntürkün, O. (2022). High associative neuron numbers could drive cognitive performance in corvid species. *Journal of Comparative Neurology*, 530(10):1588–1605, doi:10.1002/cne.25298.

Sustaita, D., Pouydebat, E., Manzano, A., Abdala, V., Hertel, F., and Herrel, A. (2013). Getting a grip on tetrapod grasping: form, function, and evolution: Grasping in tetrapods. *Biological Reviews*, 88(2):380–405, doi:10.1111/brv.12010.

Takahashi, D. Y., Liao, D. A., and Ghazanfar, A. A. (2017). Vocal Learning via Social Reinforcement by Infant Marmoset Monkeys. *Current Biology*, 27(12):1844–1852.e6, doi:10.1016/j.cub.2017.05.004.

Thorpe, S. K. S., Holder, R. L., and Crompton, R. H. (2007). Origin of Human Bipedalism As an Adaptation for Locomotion on Flexible Branches. *Science*, 316(5829):1328–1331, doi:10.1126/science.1140799.

Vereecke, E. E., D'Aoû t, K., and Aerts, P. (2006). Locomotor versatility in the white-handed gibbon (Hylobates lar): A spatiotemporal analysis of the bipedal, tripedal, and quadrupedal gaits. *Journal of Human Evolution*, 50(5):552–567, doi:10.1016/j.jhevol.2005.12.011.

Vernes, S. C. and Wilkinson, G. S. (2020). Behaviour, biology and evolution of vocal learning in bats. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1789):20190061, doi:10.1098/rstb.2019.0061.

Verpooten, J. (2021). Complex vocal learning and three-dimensional mating environments. *Biology & Philosophy*, 36(2):12, doi:10.1007/s10539-021-09786-2.

Wang, R., Chen, C.-C., Hara, E., Rivas, M. V., Roulhac, P. L., Howard, J. T., Chakraborty, M., Audet, J.-N., and Jarvis, E. D. (2015). Convergent differential regulation of SLIT-ROBO axon guidance genes in the brains of vocal learners. *The Journal of Comparative Neurology*, 523(6):892–906, doi:10.1002/cne.23719.

Wang, Z., Zhu, T., Xue, H., Fang, N., Zhang, J., Zhang, L., Pang, J., Teeling, E. C., and Zhang, S. (2017). Prenatal development supports a single origin of laryngeal echolocation in bats. *Nature Ecology & Evolution*, 1(2):1–5, doi:10.1038/s41559-016-0021. Number: 2 Publisher: Nature Publishing Group.

Werth, A. J. (2007). Adaptations of the cetacean hyolingual apparatus for aquatic feeding and thermoregulation. *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology*, 290(6):546–568, doi:10.1002/ar.20538.

Whishaw, I. Q. (2003). Did a change in sensory control of skilled movements stimulate the evolution of the primate frontal cortex? *Behavioural Brain Research*, 146(1-2):31–41, doi:10.1016/j.bbr.2003.09.027.

Whishaw, I. Q. and Coles, B. L. (1996). Varieties of paw and digit movement during spontaneous food handling in rats: Postures, bimanual coordination, preferences, and the effect of forelimb cortex lesions. *Behavioural Brain Research*, 77(1-2):135–148, doi:10.1016/0166-4328(95)00209-X.

Wimsatt, W. C. (2007). *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Harvard University Press.

Wirthlin, M., Chang, E. F., Knörnschild, M., Krubitzer, L. A., Mello, C. V., Miller, C. T., Pfenning, A. R., Vernes, S. C., Tchernichovski, O., and Yartsev, M. M. (2019). A Modular Approach to Vocal Learning: Disentangling the Diversity of a Complex Behavioral Trait. *Neuron*, 104(1):87–99, doi:10.1016/j.neuron.2019.09.036.

Wirthlin, M. E., Schmid, T. A., Elie, J. E., Zhang, X., Shvareva, V. A., Rakuljic, A., Ji, M. B., Bhat, N. S., Kaplow, I. M., Schäffer, D. E., Lawler, A. J., Annaldasula, S., Lim, B., Azim, E., Zoonomia Consortium, Meyer, W. K., Yartsev, M. M., and Pfenning, A. R. (2022). Vocal learning-associated convergent evolution in mammalian proteins and regulatory elements. preprint, Neuroscience.

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., Storz, J. F., Antunes, A., Greenwold, M. J., Meredith, R. W., Ödeen, A., Cui, J., Zhou, Q., Xu, L., Pan, H., Wang, Z., Jin, L., Zhang, P., Hu, H., Yang, W., Hu, J., Xiao, J., Yang, Z., Liu, Y., Xie, Q., Yu, H., Lian, J., Wen, P., Zhang, F., Li, H., Zeng, Y., Xiong, Z., Liu, S., Zhou, L., Huang, Z., An, N., Wang, J., Zheng, Q., Xiong, Y., Wang, G., Wang, B., Wang, J., Fan, Y., Da Fonseca, R. R., Alfaro-Núñez, A., Schubert, M., Orlando, L., Mourier, T., Howard, J. T., Ganapathy, G., Pfenning, A., Whitney, O., Rivas, M. V., Hara, E., Smith, J., Farré, M., Narayan, J., Slavov, G., Romanov, M. N., Borges, R., Machado, J. P., Khan, I., Springer, M. S., Gatesy, J., Hoffmann, F. G., Opazo, J. C., Håstad, O., Sawyer, R. H., Kim, H., Kim, K.-W., Kim, H. J., Cho, S., Li, N., Huang, Y., Bruford, M. W., Zhan, X., Dixon, A., Bertelsen, M. F., Derryberry, E., Warren, W., Wilson, R. K., Li, S., Ray, D. A., Green, R. E., O'Brien, S. J., Griffin, D., Johnson, W. E., Haussler, D., Ryder, O. A., Willerslev, E., Graves, G. R., Alström, P., Fjeldså, J., Mindell, D. P., Edwards, S. V., Braun, E. L., Rahbek, C., Burt, D. W., Houde, P., Zhang, Y., Yang, H., Wang, J., Avian Genome Consortium, Jarvis, E. D., Gilbert, M. T. P., Wang, J., Ye, C., Liang, S., Yan, Z., Zepeda, M. L., Campos, P. F., Velazquez, A. M. V., Samaniego, J. A., Avila-Arcos, M., Martin, M. D., Barnett, R., Ribeiro, A. M., Mello, C. V., Lovell, P. V., Almeida, D., Maldonado, E., Pereira, J., Sunagar, K., Philip, S., Dominguez-Bello, M. G., Bunce, M., Lambert, D., Brumfield, R. T., Sheldon, F. H., Holmes, E. C., Gardner, P. P., Steeves, T. E., Stadler, P. F., Burge, S. W., Lyons, E., Smith, J., McCarthy, F., Pitel, F., Rhoads, D., and Froman, D. P. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, 346(6215):1311–1320, doi:10.1126/science.1251385.