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

Pedro Tiago Martins* and Cedric Boeckx¹,²,³,⁴,*

¹Universitat de Barcelona
²Universitat de Barcelona Institute of Neurosciences
³Universitat de Barcelona Institute of Complex Systems
⁴Catalan 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 learn-
ing 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. Feen-
ders 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 con-
trols 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 as-
pects 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, in-
cluding 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 Knörschild, 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 neurobi-
ological barrier against the evolution of vocal learning, the latter phenotype
constitutes a complex behavior that imposes additional demands on the organ-
ism, which species are likely to meet in non-uniform ways. Here we list some
conditions that facilitate the development and manifestation of the vocal learn-
ing 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 (Elema\ns et al., 2008), and
in bats for the laryngeal control required for their extremely high-frequency
echolocating calls (Elema\ns 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.,
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). 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 inte-
gration 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 neces-
sary 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 devel-
opmental 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

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 Falotico 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 (Gal-
lardo 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 pri-
mates 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 exam-
ple, 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


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.

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

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 Ad-
vancement of Science Section: Research Article.

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

Iwaniuk, A. N., Pellis, S. M., and Whishaw, I. Q. (2000). The relative im-
portance of body size, phylogeny, locomotion, and diet in the evolution of
forelimb dexterity in fissiped carnivores (Carnivora). Canadian Journal of

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


mals revisited. Philosophical Transactions of the Royal Society B: Biological

Janik, V. M. and Sayigh, L. S. (2013). Communication in bottlenose dolphins:
50 years of signature whistle research. Journal of Comparative Physiology A,


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., Bredeney, 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,


