Electronic Supplementary Material
Acoustic allometry and vocal learning in mammals
Garcia & Ravignani

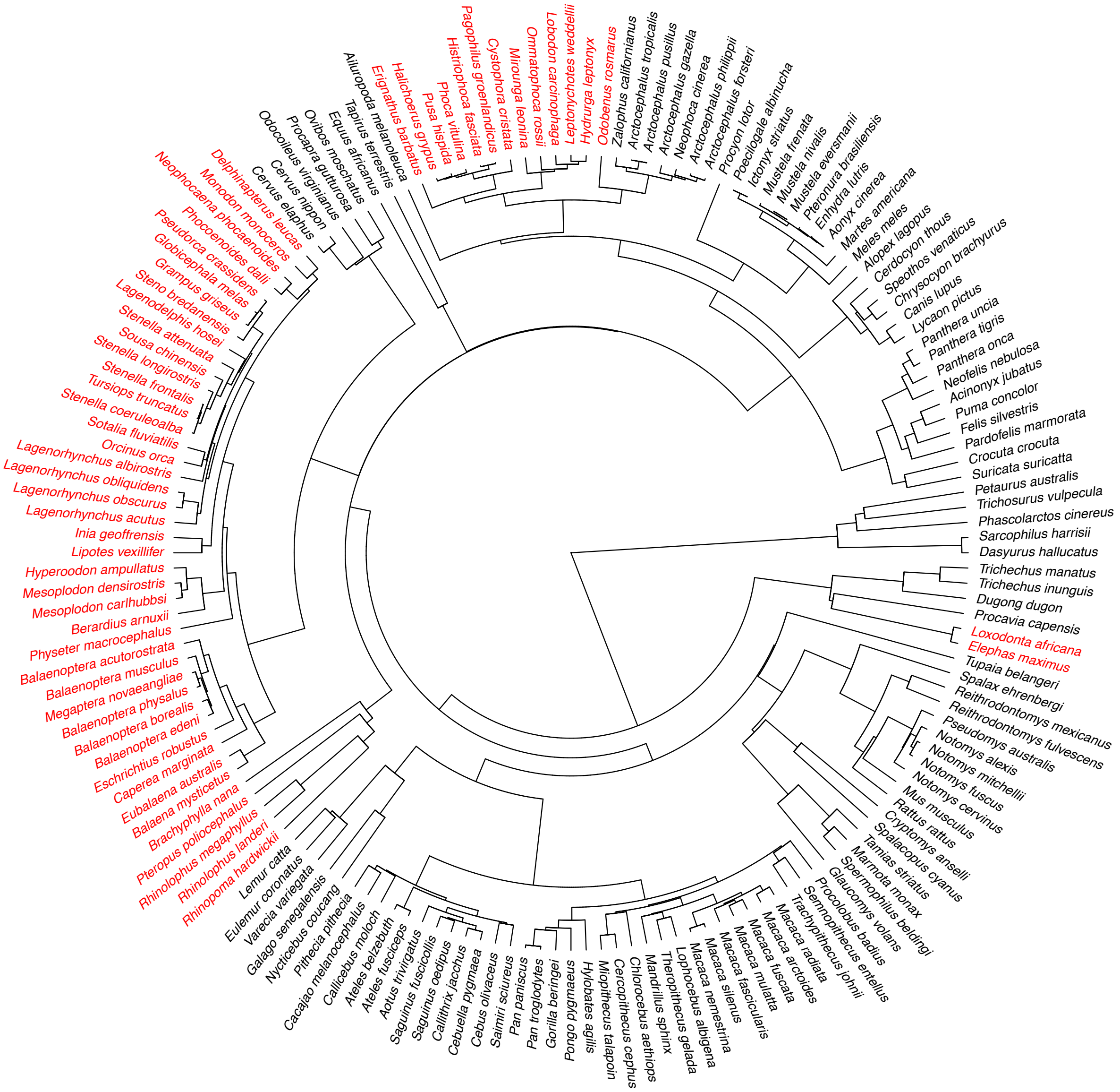
Supplementary Methods

Choice of dataset

Although this is an Opinion paper, we thought it would be important to sketch how one possible test of our hypotheses could be performed. We therefore reviewed available empirical papers comparing vocalization features across mammalian species, searching for suitable cross-species datasets. We aimed at selecting a previous study for which: 1) results had already been published in a peer-reviewed journal, as opposed for instance to conference proceedings; 2) the dataset was readily available, for instance as supplementary material; 3) the dataset included as many mammalian species as possible for which acoustic measures were available; 4) these measures featured a good proportion of species from the 4 mammalian vocal learners clades; 5) the purpose of the study and dataset was unrelated to vocal learning, so to avoid biases and cherry-picking of data. Keeping all these factors in mind, we found the dataset by Martin and colleagues [1] to be the most suitable for our purposes. Martin and colleagues identified, in their measurements, 'the signal with the lowest, and the signal with the highest peak frequency' ([1], pg.251). These chosen parameters, we believe, are best described as minimum and maximum dominant frequencies across the whole repertoire. Martin and colleagues [1] reported these parameters, which directly motivated our choice for using them. Under some assumptions, for tonal signals, the dominant frequency generally matches the fundamental frequency [2]; however, for chaotic and harsh calls, the dominant frequency more likely matches a formant frequency, since f_0 is lacking in such vocalizations. This feature, in this context, thus represents a compromise between the choice for either source-related or filter-related parameters to investigate allometry. We acknowledge that, in these dominant (peak) frequency measurements, the formants and harmonics may potentially collide spectrographically, but we do not think there is a VPL-specific bias in this respect, making our preliminary results in principle promising. Here we only offer a sketch of potential future empirical work, which will be needed to test our hypothesis more thoroughly using datasets consisting of source-only (e.g. f_0) and filter-only acoustic features (e.g. formant spacing, the most commonly used feature in acoustic allometry studies (e.g. [3])).

Characterizations of taxa as VPL

When choosing which animals to characterize as capable of VPL, we were faced with two main options. We could have chosen only the few species for which VPL had been shown (VPL species: VPLs). Alternatively, we could have chosen animal clades (VPL clades) containing multiple VPL species. The former choice, we believe, would have been more controversial, as experiments on different species have provided evidence of VPL of different strengths. We therefore decided to opt for the latter option (see Supplementary Figure 1), considering VPL clades the four non-human mammalian clades often characterized as such: cetaceans, bats, elephants and non-otariid pinnipeds [4]. The choice of excluding otariid pinnipeds is supported by the fact that, when challenged with similar tasks, phocid and otariid pinnipeds will show similar general learning abilities (e.g. vocal usage learning) but, despite their trainability, otariid pinnipeds seem incapable of VPL [5, 6]. Categorizing species within a VPL clade as VPLs involves generalizing this ability to a varying number of species given the clade considered; we recognize this broad classification constitutes a limitation of our preliminary empirical approach, which should be improved as more and more species are shown capable of VPL (or at least appropriately tested for such abilities). Note that our framework is echoed by the study of VPL in songbirds, a historically valued model in this field [7]. We chose to focus on mammals only because of the different anatomical and neural structures involved in VPL in mammals and songbirds [7, 8]. However, songbirds typically show sexually selected vocalizations and extreme frequency modulation [9], as well as varying degrees of acoustic allometry scaling [10]. Thus, songbirds provide an ideal comparative basis for future research testing our hypotheses.



Supplementary Figure 1. Phylogenetic tree of all species included in our preliminary analyses. VPL clades are highlighted in red and are defined as the animal clades containing multiple VPL species.

Data acquisition and analysis

To investigate the relationship between body mass and two acoustic parameters (maximum dominant frequency and minimum dominant frequency) we used previously published datasets for both acoustic data (https://doi.org/10.5061/dryad.289kh ; [1], see our supplementary_data for our edited version of the dataset) and phylogenies (http://doi.org/10.1016/j.ympev.2014.11.001 ; [11]). We retained 164 species for which we had overlapping acoustic and phylogenetic data, and after log-transforming acoustic variables and body mass, we conducted Phylogenetic Generalized Least Square (PGLS) regressions using the 'nlme' package of the R software [12]. We then extracted the residuals from these regressions and qualified as outliers all species whose residual values exceeded the threshold of at least 2.5 times that of the Median Absolute Deviation (following [13]). This step led to 1) cross-tabulation via contingency tables counting species based on whether they were either from VPL clades or non-VPL clades, and either outliers or non-outliers (Supplementary Table 1), and 2) a residual value for each species, quantifying the magnitude of deviation from allometric scaling (see supplementary data). Barnard's exact tests [14] were then used to test for the statistical association between the cross-tabulated number of species in VPL clades vs. outliers of acoustic allometry. In parallel, residual values were compared between VPL clades and non-VPL clades. Given that the normality assumption did not hold for the dataset of the residuals, Mann-Whitney U-tests were used to test for a difference in distribution of the magnitude of deviation from allometric scaling (i.e. the absolute values of residuals) between VPL clades and non-VPL clades (main text Table 1). Two-tailed statistics were used, and significance levels were set at alpha = 0.05.

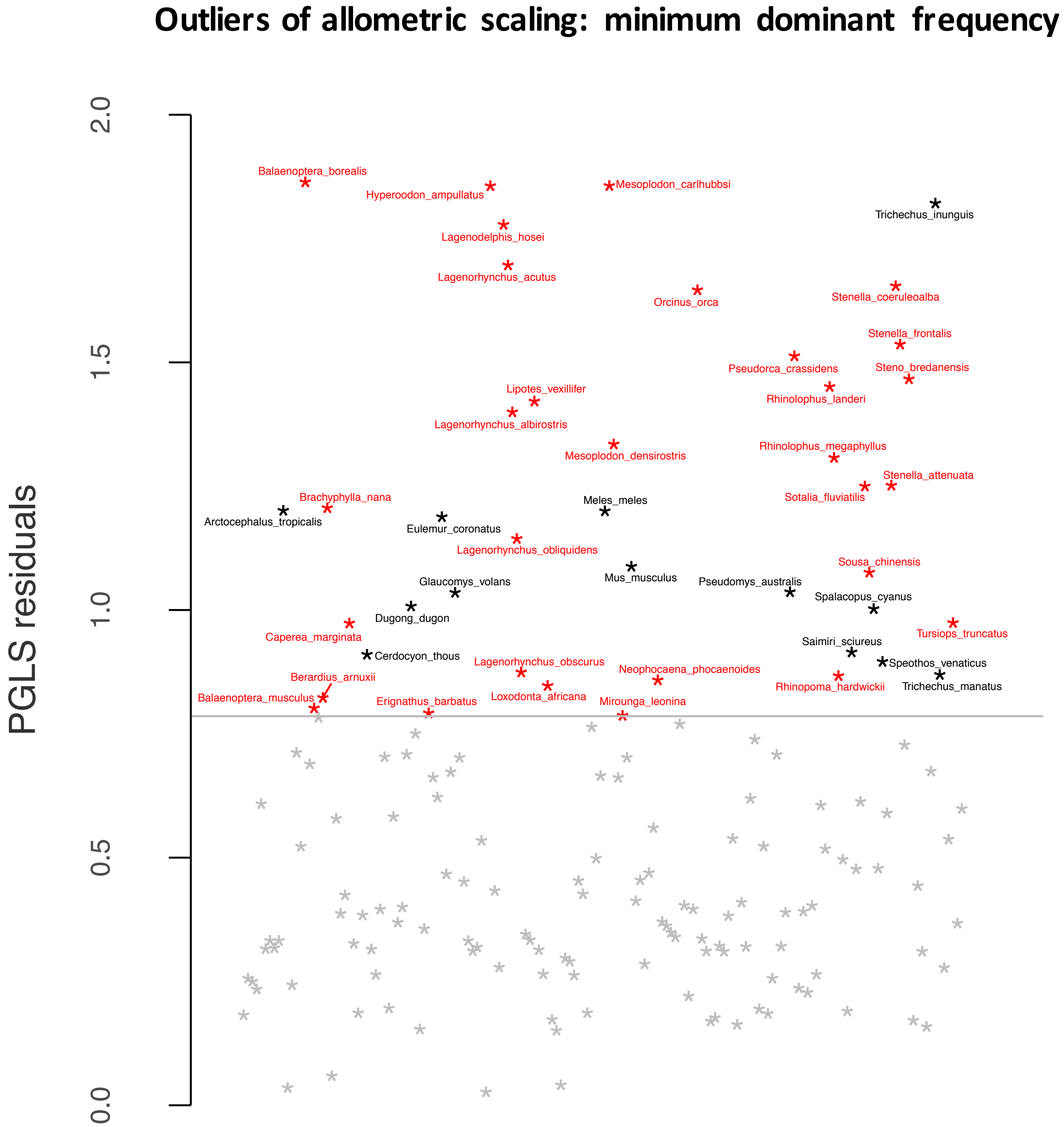
Supplementary Table 1. Number of species in VPL clades and outlier species among 164 mammals.

		VPL	**non-VPL**	**Total**
Minimum frequency	**Outliers**	30	13	43
	Non-outliers	28	93	121
	Total	58	106	164
		VPL	**non-VPL**	**Total**
Maximum frequency	**Outliers**	15	19	34
	Non-outliers	43	87	130
	Total	58	106	164

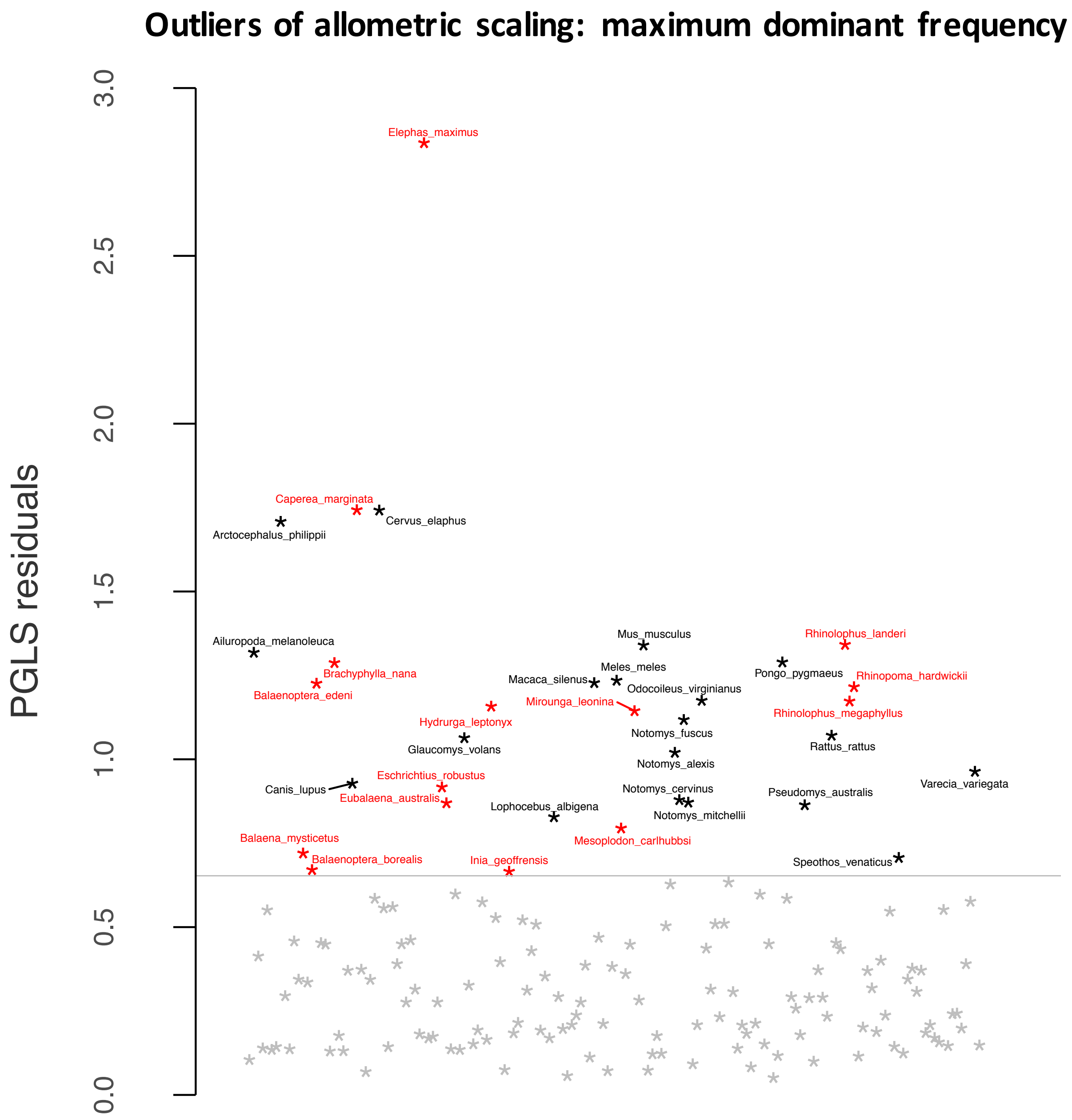
In order to highlight the predictive value of our approach, we complemented the analyses described above with a visual representation of the residuals' distribution in Supplementary Figures 2 & 3. As in the main text's Figure 2, we see an association between being an outlier to allometric scaling (i.e. having large residuals in a PGLS regression between body size and an acoustic trait) and belonging to a VPL clade (highlighted in red in Supplementary Figures 2 & 3). In particular, this association is significant when using minimum dominant frequency (Supplementary Figure 2), while the proportion of outliers when using maximum dominant frequency is similar between VPL clades and non-VPL clades (Supplementary Figure 3). This difference, and the fact that significant results are found in both tests (main text Figure 2 and Table 1) on the minimum dominant frequency, dovetails with the idea that minimum frequency may be more appropriate than maximum frequency to investigate acoustic allometric relationships [15, 16]. In this particular case, the dominant frequency from [1] may be more reliable when using the minimum dominant frequency, as this parameter (unlike maximum dominant frequency) would generally coincide with f_0 in tonal calls and with F_1 in harsh vocalisations. Note that this does not preclude finding significant results for both minimum and maximum dominant frequencies (Table 1 in the main text) when comparing species-wise residuals in VPL clades and non-VPL clades, because significant differences between residuals might be driven by the magnitude of outliers in the maximum dominant frequency regression (see Supplementary Figure 3).

A confounding factor in our approach lies in that outliers to allometry (species above the grey line in Supplementary Figures 2 & 3) can either correspond to species with anatomical adaptations or species with enhanced volitional vocal control (see main text's Figure 1e for a theoretical perspective). This could partly explain the difference of significance when testing for association (between belonging to a VPL clade and being an allometric outlier) via cross-tabulation between minimum dominant frequency (main text Figure 2, Supplementary Figure 2) and maximum dominant frequency (main text Figure 2, Supplementary Figure 3).

Bearing in mind that outliers can be explained in at least two ways (VPL and anatomical adaptation), our framework emphasizes intriguing, possibly promising, cases of species currently assigned to non-VPL clades but for which controversies and debates exist. Indeed, among the non-VPL species found to be allometric outliers (see main text Table 1 for the five species with the largest residuals; see Supplementary Figures 2 & 3, black stars, for the full set of outliers in non-VPL clades), we find species whose VPL abilities have been questioned, such as *Pongo pygmaeus* [17] in Supplementary Figure 3, and *Mus musculus* [18, 19] in Supplementary Figures 2 & 3. Other interesting cases are phylogenetically close to recognized VPL clades. For instance, *Arctocephalus tropicalis* (Supplementary Figure 2) and *Arctocephalus philippii* (Supplementary Figure 3) are both otariid pinnipeds and related to phocid pinnipeds (Supplementary Figure 1); *Trichechus inunguis, Trichechus manatus* and *Dugong dugon* (Supplementary Figure 2) are all sirenia and related to elephantidae (Supplementary Figure 1). One can also note in Supplementary Figure 3 the presence of all members (among the 164 species in our dataset) of the murinae subfamily (*Mus musculus; Notomys alexis; Notomys cervinus; Notomys fuscus; Notomys mitchellii; Pseudomys australis; Rattus rattus*). This suggests a potential convergence towards deviation from allometry scaling, which could or could not have to do with VPL, but at the very least calls for comparative investigation of selective pressures over sound production in this clade. Overall, while the framework and approaches we propose do not allow to draw direct conclusions on whether a species possesses VPL abilities, they do provide valuable tools to identify candidate VPLs or reconsider the status of debated cases.



Supplementary Figure 2. Relationship between VPL and deviation from allometric scaling in 164 mammal species using body mass and minimum dominant frequency. Outlier species to standard allometric scaling are those above the threshold indicated by the grey line. All species below the grey line are not outliers and are indicated by grey stars. Among the outliers, species in VPL clades are indicated with a red star and those in non-VPL clades with a black star. Species are arranged alphabetically along the horizontal axis.



Supplementary Figure 3. Relationship between VPL and deviation from allometric scaling in 164 mammal species using body mass and maximum dominant frequency. Outlier species to standard allometric scaling are those above the threshold indicated by the grey line. All species below the grey line are not outliers and are indicated by grey stars. Among the outliers, species in VPL clades are indicated with a red star and those in non-VPL clades with a black star. Species are arranged alphabetically along the horizontal axis.

Supplementary discussion

Are there no allometric constraints in VPL species?

In the theoretical scenario we described, we emphasize the role of flexibility within allometric constraints in VPLs; however, these anatomical constraints (on e.g. the fundamental frequency and formants of vocalisations) seen in many species still apply to vocal learners. In other words, VPLs still do have limits, as we actually show in Figure 1a-d, where the boundary whiskers are always present and bound sound production. Nevertheless, VPLs may have developed a tendency to vocalize both at the edge of the boundary and far from it, flexibly across communicative interactions, which may be a key difference from species lacking VPL abilities. Thus, although signals of VPLs could in principle reflect body size, our suggestion is that the enhanced capacity for vocal modulation makes VPLs stand out from these strict "size give-away" boundaries.

Besides offering a predictive framework for the identification of new VPLs, our analysis also provides a window into the relationship between size variation of the vocal apparatus, and increased volitional motor control. For candidate species where sexual selection for allometry-cheating could have led to VPL, a promising approach will consist in investigating the association, or potential tradeoff, between the magnitude of size variation of the larynx or vocal tract to shift f_o or F_1, and the degree of volitional control. Going one step further, and in line with the VPL contiguum idea [4], one could even try to test hypotheses about the relationship between vocal tract scaling and neuroanatomical scaling underlying volitional motor control. One empirical approach to this would be to relate 5 sorts of measures, namely: 1) body size, 2) vocal tract measurements, 3) a quantification of neural projections (e.g. fiber length [20]) towards the larynx and upper vocal tract, 4) f_o, $F_1$and other classical acoustic measures, 5) acoustic measures of range and variability of source-related and filter-related features.

Within-species allometry vs. between-species allometry

When talking about allometry and violations from allometry, it is worth remembering there is a difference between within-species allometry and between-species allometry. In general, one could imagine one species for which the sounds produced by larger individuals have features (statistically) equal to those of smaller individuals, while the species as a whole perfectly fits an allometric cross-species scaling law. This would illustrate a case of between-species only (but not within-species) allometry. Conversely, one could imagine one species which violates the cross-species allometric law (e.g. a squeaky manatee sounds unexpectedly high-pitched when compared to a similar-sized cow) where, however, the sound features of an individual strongly scale with its body size when compared to other conspecifics. This would be a case of within-species only (but not between-species) allometry. Acoustic allometry research can deal with both types of analyses and, needless to say, the animal kingdom shows cases of within-species only and between-species only allometry, apart from cases where both or neither are present. This distinction is quite relevant for our reasoning. In fact, when detecting scaling laws between acoustic features and body measures at a between-species level, one could hypothesize that different species or clusters of species offer snapshots in time of potential evolutionary paths.

Why would selective pressures sometimes lead to VPL?

If we hypothesize that VPLs can volitionally modify the fundamental frequency and formant features of their vocalizations without resorting to additional anatomical modifications, a question may arise: Could natural selection have made exaggerated acoustic features possible without volitional control over vocal production, resorting instead to specialized anatomical modifications? In fact, natural selection did create cases of exaggerated acoustic features through anatomical modifications, with no need for volitional control over vocal production; examples of these anatomical modifications include the extremely elongated trachea of some bird species [21] or the laryngeal descent seen in red deer [22].

Natural selection has resorted multiple times to specialized anatomical modifications, instead of volitional control over vocal production, to produce exaggerated acoustic features. However, escaping standard acoustic allometry is one ‘target’ with many possible ‘solutions’: one of these consists in producing exaggerated acoustic features via physical adaptations (e.g. the examples given above), while another consists in flexible deviations from allometry as precursors to VPL. Often, witnessing the outcome of selection is only possible when all individuals have converged towards an evolutionary stable strategy [23, 24]. In the first case, the strategy was, for instance, to elongate the vocal tract, possibly triggering runaway selection towards even more elongation, resulting in today’s descended larynges in red deer or coiled tracheas in trumpet manucodes [21, 22]. In the second case, the strategy may have been to deploy allometry-cheating vocalizations only opportunistically and in specific situations. This probabilistic reach of the extremes (as opposed to always hitting anatomical boundaries) is what makes dishonest signalling an evolutionary stable strategy; e.g. by ‘sounding low’ only sometimes, runaway selection for anatomical modifications is not triggered and these modifications are not pushed to an extreme (unlike the first case) [23-25].

The difference between the two scenarios may be due to the strength of selective pressures. If selection for dishonest signaling is strong, a trait (an acoustic feature in the present case) may be pushed to one of its extremes (via e.g. runaway selection), leading to anatomical adaptations such as those seen in red deer [22] and koalas [26]. These bio-physical boundaries provide constraints for all species, both VPLs and non-VPLs alike. However, relaxed selection can pave the way to phenotypic plasticity [27], where (in our case) sound production does not need to always reach extrema and can instead explore the whole space of deviations from allometry. This space can include size-exaggeration but does not coincide with it: it features, more in general, various potential forms of ‘dishonest’ signaling. Note that, for the sake of argument, the above rationale relies on a size exaggeration mechanism, but it could in theory equally rely on a size reduction mechanism.

Besides the conventional scenarios described above, a third bio-physical mechanism could explain allometry-cheating. Indeed, we ask: could selective pressures lead to exaggerated acoustic features via allometry-cheating while resorting to neither volitional control over vocal production nor specialized anatomical modifications? If such a mechanism existed, it would likely take the form of specific brain wiring that does not require volitional control. This circuit would 'automatically' send a signal (e.g., via the autonomic nervous system) to the phonatory apparatus to emit sounds which are further away from the average frequency typically predicted by allometric scaling for that individual or species. Similar to the argument above, this may be an evolutionary stable strategy as long as this non-volitional behaviour is only deployed sporadically. While this third mechanism cannot be excluded, it is also not more likely or parsimonious than the other two scenarios.

In addition to these arguments, there is some evidence that learning can push the boundaries of sound production. Humans receiving training to sing expand the range of fundamental frequencies they can produce [28, 29]. In other words, those humans who train their voice can actively push the boundaries of their low and high frequencies, hence effectively pushing their individual 'whiskers' in our Figure 1. Obviously, humans trained to sing are still vocalizing within their physical limits; however, training seems to let them reach a section of the boundary, which may not be reached without learning. While, to our knowledge, no such long-term experiment involving training has been done in other species, it is important to notice that training can only work if there is an underlying capability for learning.

Supplementary references

[1] Martin, K., Tucker, M.A. & Rogers, T.L. 2017 Does size matter? Examining the drivers of mammalian vocalizations. *Evolution* **71**, 249-260. (doi:10.1111/evo.13128).
[2] Bowling, D., Garcia, M., Dunn, J., Ruprecht, R., Stewart, A., Heinz-Frommolt, K. & Fitch, W. 2017 Body size and vocalization in primates and carnivores. *Sci. Rep.* **7**, 41070.
[3] Fitch, W.T. 1997 Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *J. Acoust. Soc. Am.* **102**. (doi:10.1121/1.421048).
[4] Martins, P.T. & Boeckx, C. 2020 Vocal learning: Beyond the continuum. *PLoS Biol.* **18**, e3000672. (doi:10.1371/journal.pbio.3000672).
[5] Cook, P., Rouse, A., Wilson, M. & Reichmuth, C. 2013 A California sea lion (*Zalophus californianus*) can keep the beat: Motor entrainment to rhythmic auditory stimuli in a non vocal mimic. (pp. 412-427. US, American Psychological Association.
[6] Schusterman, R.J. 2008 Vocal learning in mammals with special emphasis on pinnipeds. In *Evolution of communicative flexibility: Complexity, creativity, and adaptability in human and animal communication* (eds. D.K. Oller & U. Griebel), pp. 41-70, MIT Press.
[7] Jarvis, E.D. 2007 Neural systems for vocal learning in birds and humans: a synopsis. *Journal of Ornithology* **148**, 35-44. (doi:10.1007/s10336-007-0243-0).

[8] Tyack, P.L. 2020 A taxonomy for vocal learning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20180406. (doi:10.1098/rstb.2018.0406).

[9] Marler, P.R. & Slabbekoorn, H. 2004 *Nature's music: the science of birdsong*. San Diego, CA, Elsevier Academic Press.

[10] Patel, R., Mulder, R.A. & Cardoso, G.C. 2010 What makes vocalisation frequency an unreliable signal of body size in birds? A study on black swans. *Ethology* **116**, 554-563. (doi:10.1111/j.1439-0310.2010.01769.x).

[11] Faurby, S. & Svenning, J.C. 2015 A species-level phylogeny of all extant and late Quaternary extinct mammals using a novel heuristic-hierarchical Bayesian approach. *Mol. Phylogenet. Evol.* **84**, 14-26. (doi:10.1016/j.ympev.2014.11.001).

[12] RCoreTeam. 2017 R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL http://www.r-project.org/.

[13] Leys, C., Ley, C., Klein, O., Bernard, P. & Licata, L. 2013 Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *J. Exp. Soc. Psychol.* **49**, 764-766. (doi:10.1016/j.jesp.2013.03.013).

[14] Barnard, G.A. 1947 Significance tests for 2 × 2 tables. *Biometrika* **34**, 123-138. (doi:10.2307/2332517).

[15] Garcia, M., Herbst, C.T., Bowling, D.L., Dunn, J.C. & Fitch, W.T. 2017 Acoustic allometry revisited: morphological determinants of fundamental frequency in primate vocal production. *Sci. Rep.* **7**, 10450. (doi:10.1038/s41598-017-11000-x).

[16] Garcia, M. & Dunn, J.C. 2019 No evidence that maximum fundamental frequency reflects selection for signal diminution in bonobos. *Curr. Biol.* **29**, R732-R733. (doi:10.1016/j.cub.2019.06.022).

[17] Lameira, A.R., Hardus, M.E., Mielke, A., Wich, S.A. & Shumaker, R.W. 2016 Vocal fold control beyond the species-specific repertoire in an orang-utan. *Sci. Rep.* **6**, 30315. (doi:10.1038/srep30315).

[18] Arriaga, G., Zhou, E.P. & Jarvis, E.D. 2012 Of mice, birds, and men: The mouse ultrasonic song system has some features similar to humans and song-learning birds. *PLoS ONE* **7**, e46610. (doi:10.1371/journal.pone.0046610).

[19] Portfors, C.V. & Perkel, D.J. 2014 The role of ultrasonic vocalizations in mouse communication. *Curr. Opin. Neurobiol.* **28**, 115-120. (doi:10.1016/j.conb.2014.07.002).

[20] Assaf, Y., Bouznach, A., Zomet, O., Marom, A. & Yovel, Y. 2020 Conservation of brain connectivity and wiring across the mammalian class. *Nat. Neurosci.* (doi:10.1038/s41593-020-0641-7).

[21] Fitch, W.T. 1999 Acoustic exaggeration of size in birds via tracheal elongation: comparative and theoretical analyses. *J. Zool.* **248**, 31-48. (doi:10.1111/j.1469-7998.1999.tb01020.x).

[22] Fitch, W.T. & Reby, D. 2001 The descended larynx is not uniquely human. *Proc. Biol. Sci.* **268**, 1669-1675. (doi:10.1098/rspb.2001.1704).

[23] Weibull, J.W. 1997 *Evolutionary Game Theory*, MIT Press.

[24] Nowak, M.A. 2006 *Evolutionary dynamics: exploring the equations of life*, Harvard University Press.

[25] Dawkins, M.S. & Guilford, T. 1991 The corruption of honest signalling. *Anim. Behav.* **41**, 865-873. (doi:10.1016/S0003-3472(05)80353-7).

[26] Charlton, B.D., Frey, R., McKinnon, A.J., Fritsch, G., Fitch, W.T. & Reby, D. 2013 Koalas use a novel vocal organ to produce unusually low-pitched mating calls. *Curr. Biol.* **23**, R1035-1036. (doi:10.1016/j.cub.2013.10.069).

[27] Hunt, B.G., Ometto, L., Wurm, Y., Shoemaker, D., Yi, S.V., Keller, L. & Goodisman, M.A.D. 2011 Relaxed selection is a precursor to the evolution of phenotypic plasticity. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15936. (doi:10.1073/pnas.1104825108).

[28] Awan, S.N. 1991 Phonetographic profiles and F0-SPL characteristics of untrained versus trained vocal groups. *J. Voice* **5**, 41-50. (doi:10.1016/S0892-1997(05)80162-6).

[29] Mendes, A.P., Rothman, H.B., Sapienza, C. & Brown, W.S. 2003 Effects of vocal training on the acoustic parameters of the singing voice. *J. Voice* **17**, 529-543. (doi:10.1067/S0892-1997(03)00083-3).