PNAS COMMENTARY

Evolving alternative neural pathways for vocal dexterity

Yisi S. Zhang[a] and Asif A. Ghazanfar[a,1]

We humans are remarkably adept at using our voices. We can use them for comic effect when we imitate the peculiar speech of other humans or to time the delivery of a punchline. We can control our pitch and loudness for singing musical notes. Even the basic ability to produce consonant and vowel sounds is only expressed in its fullest capacity in us (but see refs. 1 and 2). Our vocal flexibility is not something we really share with our close primate relatives like chimpanzees and macaque monkeys, which produce seemingly simpler vocalizations. Why do they lack our vocal skills? It does not seem to be because the nonhuman primate vocal anatomy (larynx, lungs, and vocal tract) is radically different. After all, computer modeling shows that the vocal anatomy of a macaque monkey is capable of human speech (3). The behavioral differences must stem from neural differences. A popular idea is that our vocal skills are accounted for by the existence of a direct, human-specific neural pathway for controlling the larynx. It goes from primary motor cortex to a brainstem area called the nucleus ambiguus (in one step; this structure then connects to the muscles of the larynx). In PNAS, Cerkevich et al. directly test this hypothesis (4).

There were two key factors to their study: choice of model species and choice of methodology. Marmoset monkeys were the model species. Marmosets are small primates native to Brazil and distantly related to us; they exhibit remarkable vocal skills relative to other nonhuman primates. Like us, they evolved the ability to adjust the timing and amplitude of their vocalizations according to social context (5). They can also change their "dialect," gradually modifying their voices to sound similar to those of nearby marmosets (6). Thus, the specific hypothesis to test was, Are the vocal skills of marmosets due to the existence of direct connections from the primary motor cortex to the nucleus ambiguus (as it seems to be in humans)? The methodology was to inject rabies virus into a laryngeal muscle known as the cricothyroid which can modify laryngeal tension; this was done in both marmoset monkeys (vocally skilled) and macaque monkeys (not so vocally skilled). The rabies virus is transported "backward," so to speak, and across synapses (7). This transsynaptic movement of the rabies virus is time-dependent. By controlling how long the rabies virus is in the brain, infected neurons of different synaptic distances from the injection site (in this case, the laryngeal muscle) can be identified and muscle-specific neural pathways are revealed.

The results were unambiguous. In neither marmosets nor macaques are there direct connections from primary motor cortex to the nucleus ambiguus (4). Furthermore, there were not any direct connections from any other motor-related ("premotor") cortical areas to that brainstem structure. The data showed that in both species, despite their vocal behavioral differences, all vocal–motor cortical signals must first be relayed to another brainstem region—the reticular formation—and from there to the nucleus ambiguus (a two-step pathway). Since marmosets have considerable vocal skills, the hypothesis that such skills can only be conferred by such connections was falsified. So, then, what accounts for the difference in vocal skills between marmosets and macaques, and the similarities between marmosets and humans?

If we assume that more neurons mean better ability—an assumption that is aligned with a number of observations (see ref. 8 for a thoughtful discussion of this idea)—then one possibility is that there are more primary motor cortical neurons projecting to the reticular formation in marmosets than in macaques (relative to brain size). This could lead to increased voluntary vocal control despite the multistep pathway to the nucleus ambiguus. This also turned out not to be true; however, a variant of the idea was substantiated. Among the four premotor cortical areas investigated in this study, a subset of them in marmosets had significantly more connections with the reticular formation when compared to macaque brains: the lateral ventral area 6 and the medial supplementary motor area (Fig. 1) (4). It is these neural differences that account for the vocal behavioral differences between species. Moreover, the increased robustness of these premotor cortical pathways points to an alternative means for generating vocal skills—some such skills do not require a direct connection from primary motor cortex to the nucleus ambiguus (Fig. 1). To exclude the possibility that maybe motor control in marmosets is generally different from macaque monkeys [perhaps due to developmental constraints (9)], Cerkevich et al. traced the pathway for manual skill, i.e., from the hand muscle (macaques are more dexterous with their hands than marmosets). The patterns of connections in these neural pathways deviated strikingly from the laryngeal one in marmosets (so there was no general motor control pattern specific to this species) and differed between marmosets and macaques. Thus, the vocal–motor circuit in marmosets appears to be a specialization for vocal skills.

While marmosets and humans exhibit convergent evolution of certain (but not all) vocal skills, the neural implementation of these skills is different between species. This is reminiscent of the convergent evolution of flight. Birds lost the separate digits that their reptilian ancestors had

Author affiliations: [a]Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544

Author contributions: Y.S.Z. made the figure; and Y.S.Z. and A.A.G. wrote the paper.

The authors declare no competing interest.

Copyright © 2022 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

See companion article, "Cortical basis for skilled vocalization," 10.1073/pnas.2122345119.

[1]To whom correspondence may be addressed. Email: asifg@princeton.edu.

Published June 10, 2022.

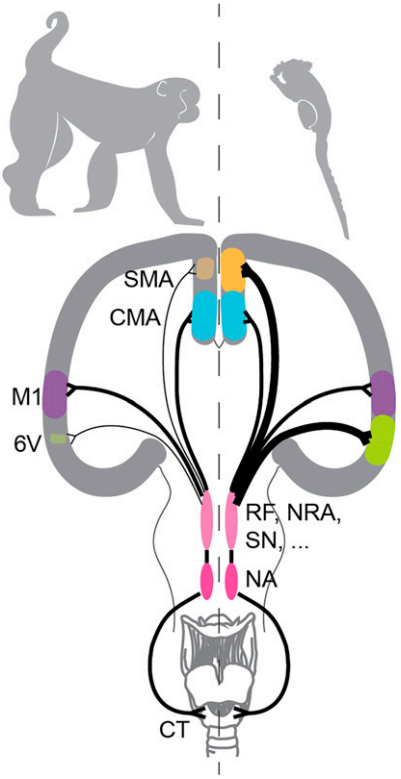

Fig. 1. Neural pathways to a laryngeal muscle. The left and right halves illustrate the difference between macaque (*Left*) and marmoset (*Right*) brains. Size represents the proportion of labeled neurons relative to M1. SMA: supplementary motor area; CMA: cingulate motor area; M1: primary motor cortex; 6V: ventral area 6; RF: reticular formation; NRA: nucleus retroambiguus; SN: solitary nucleus; NA: nucleus ambiguus; CT: cricothyroid.

by fusing them together to undergird the structure of the wing. Bats still have five separate digits that elongated to accommodate the structure of their wings. Both can fly, but bats can manipulate their wings more, which allows them to be much more agile in the air compared to birds. Similarly, marmoset monkey vocal skills are not completely on par with human skills. For example, marmosets cannot imitate others in real time, nor do they produce the variety of sounds that we can. Whether these differences are solely due to the direct pathway from motor cortex to the nucleus ambiguus present in humans is unlikely, but the idea does provide a foundation upon which to launch further laudable investigations like the work of Cerkevich et al. described above.

Why did marmosets and humans converge on increased vocal skills relative to other primates? The life-history strategy and early postnatal vocal learning of marmoset monkeys parallels the life-history strategy and prelinguistic vocal learning of humans (reviewed in ref. 10). Both species incur massive energetic costs during pregnancy relative to other primates. In humans, this is due to the large and rapidly growing fetal brain; in marmosets, it is due to the gestation of twins. These energetic constraints lead to earlier births in both species, resulting in altricial offspring (again, relative to other primates) that cannot move by themselves and lack good control of their vocalizations. As a result, mothers in both species rely on help from fathers, older siblings, and other group members. Such a cooperative breeding strategy is rare among primates. Relatedly, to reliably elicit care from others, infants of both species use learning via social reinforcement to produce better, more-mature-sounding vocalizations. Thus, one possibility is that the vocal dexterity in marmoset monkeys and humans arose as a by-product of the convergent evolution of a developmental system that includes altriciality and a cooperative breeding system.

ACKNOWLEDGMENTS. Y.S.Z. and A.A.G.'s research is supported by National Institute of Neurological Disorders and Stroke, NIH Grant R01NS054898.

1. L.-J. Boë *et al.*, Evidence of a vocalic proto-system in the baboon (Papio papio) suggests pre-hominin speech precursors. *PLoS One* **12**, e0169321 (2017).
2. A. R. Lameira, I. Maddieson, K. Zuberbühler, Primate feedstock for the evolution of consonants. *Trends Cogn. Sci.* **18**, 60–62 (2014).
3. W. T. Fitch, B. de Boer, N. Mathur, A. A. Ghazanfar, Monkey vocal tracts are speech-ready. *Sci. Adv.* **2**, e1600723 (2016).
4. C. M. Cerkevich, J.-A. Rathelot, P. L. Strick, Cortical basis for skilled vocalization. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2122345119 (2022).
5. A. A. Ghazanfar, D. A. Liao, D. Y. Takahashi, Volition and learning in primate vocal behaviour. *Anim. Behav.* **151**, 239–247 (2019).
6. Y. Zürcher, E. P. Willems, J. M. Burkart, Are dialects socially learned in marmoset monkeys? Evidence from translocation experiments. *PLoS One* **14**, e0222486 (2019).
7. R. P. Dum, P. L. Strick, Transneuronal tracing with neurotropic viruses reveals network macroarchitecture. *Curr. Opin. Neurobiol.* **23**, 245–249 (2013).
8. G. F. Striedter, *Principles of Brain Evolution* (Sinauer Associates, Sunderland, MA, 2005).
9. B. L. Finlay, R. B. Darlington, Linked regularities in the development and evolution of mammalian brains. *Science* **268**, 1578–1584 (1995).
10. T. T. Varella, A. A. Ghazanfar, Cooperative care and the evolution of the prelinguistic vocal learning. *Dev. Psychobiol.* **63**, 1583–1588 (2021).