Vocal Learning and Songbirds: An Evolutionary Tale of Singing

Heidi Moss Erickson

Heidi Moss Erickson

Journal of Singing, September/October 2022
Volume 79, No. 1, pp. 87–94
https://doi.org/10.53830/XHJI5745

[**Introduction to the "Minding the Gap" series.** Research on the neurobiological underpinnings of vocalization is growing at a rapid pace. Scientists from varied disciplines contribute to this field, elucidating the process from such diverse angles as evolutionary biology, molecular biology, genetics, neuroscience, and psychology.[1] Singing is one of our most complex behaviors, involving coordination of more than 100 muscles, integrating both our musical and linguistic selves. Influences are derived from a looped brain-body-environment continuum and serve as the foundation for vocalization and artistic expression. It is therefore essential that singers, pedagogues, and voice scientists capitalize on this wealth of data to fully understand our instruments. An integrative view of how and why we sing can refine the art of voice pedagogy, demystify long held myths, and yield greater vocal efficiency, making singers better faster.

The goal of this series, "Minding the Gap," is to bridge the divide between traditional voice pedagogy with the most current research on the brain. This is a dynamic process, and the hope is that this snapshot in scientific time will encourage singers and teachers to follow some of this work as it continues to evolve. Given the constant output in neuroscience research, paradigm shifts are the norm. To that end, it is important to follow the latest work from the diverse minds in the field.

The first installment of "Minding the Gap" will explore the field of vocal learning from both an evolutionary and biological perspective. These concepts set the stage for future reviews on the neuroscience of breath, learning, articulation, phonation, emotion, gesture, and more. Each survey will include integrated voice pedagogy and voice science components so that readers can directly apply concepts to their field of study. The hope is that this series will open the door to new lines of inquiry and curiosity, moving our field into a new era of exploration.]

> The singing instinct, which men have in common with the birds is, without a doubt, at the root of the vocal art.
>
> —David Ffrangcon-Davies, *The Singing of the Future* (1905)

Vocalization has both biological and anthropological roots. Humans speak and sing in order to communicate and socially bond. Each individual gains this ability through listening to acoustic models, forming auditory templates, and consolidating sensory information. This results in an identifiable vocal output, as unique as a fingerprint. When we analyze the parallel evolution that gives rise to our vocal abilities, we gain insight into the inner workings of our complex instrument. Traditional voice pedagogy focuses on the question of how to sing; however, by analyzing our evolutionary story, we gain a greater understanding of why we sing. Ultimately, the why serves the how, leading to a deeper appreciation for our instrument by all who care for the voice.

Much discussion around the evolution of complex human behaviors focuses on our primate cousins. We have inherited a great deal from our simian counterparts; chimpanzees, for example, share 99% of our DNA, they laugh when tickled, create strong social bonds, use hand gestures to communicate, mimic,[2] and even demonstrate elements of emotion.[3] However, two critical components for higher order vocalizations—a flexible vocal apparatus and a brain that coordinates complex sounds—are noticeably absent or severely compromised in our close relatives.

Regarding structural elements of the voice, the vocal limitations of nonhuman primates have been partially attributed to the anatomy and configuration of their vocal tract (i.e., a high position of the larynx and a less flexible tongue). This was described in a seminal 1969 paper by Lieberman et al., who used a computer algorithm to extract the phonetic capability of subjects, and thus serve as a model for other nonhuman primates.[4] Their measurements concluded that monkeys do not have a physical capacity for complex speech. Although in recent years scientists proved that there is more variability than initially thought,[5] the fact remains that adult humans possess a more flexible vocal tract capable of more diverse vocalizations than our primate cousins. A more critical difference, however, lies in the correlative brain regions for vocalization. Nonhuman primates cannot voluntarily control pitch and acoustic patterns; the sounds they create are rudimentary and limited.

In contrast, a small class of animals, called vocal learners, have the ability to imitate new sounds, modify inputs, and produce more complex vocalizations.[6] There are only eight vocal learning animal groups: humans, bats, cetaceans (whales), pinnipeds (seals/dolphins), elephants, and songbirds (e.g., parrots, hummingbirds, and finches).[7] Vocal learning requires a great deal of both auditory and motor processing to carry out a vast repertoire of vocalizations. To that end, it takes more than a mechanism; it requires immense brain power as well. Humans have a brain three times the size of a chimpanzee's,[8] and songbirds have higher neuron packing densities than mammalian brains.[9] Given the similar singing skills of humans and songbirds, we can learn a lot about our own vocal processes through study of theirs.

Songbirds have been a model system for studying neural mechanisms of vocal learning since the pioneering work of Nottebohm at Rockefeller University in the 1970s[10] and Marler at UC Davis.[11] Since that time, scientific advancements have opened up the field in remarkable ways, and recent discoveries have revealed extraordinary similarities at both the macro and micro levels in both structure and function between songbirds and humans. Understanding these parallel systems can inform singers and pedagogues to generate biologically tailored methods toward their own vocal learning.

Zebra finches learn their songs through trial and error after repeated listening to a tutor song (Figure 1).[12] A distinct neural pathway enables the birds to make this comparison and to use any discrepancies to improve their subsequent attempts. In other words, songbirds learn vocal repertoire from first hearing their parents and experimenting through trial and error. Researchers also discovered an analogous, stepwise learning process for human infants.[13] Singers can capitalize on this research for optimal vocal learning, including trial and error without judgment; perfection isn't the goal.

Several key steps were essential for both species. Auditory input is the most critical; listening provides the imitative substrate for future sensorimotor learning. Small, repeated patterns imprint in the brain, gradually increasing in complexity and diversity. This creates a repertoire of vocalizations that will be practiced in units before each animal initiates individualized changes. Similarly, to optimize imprinting, vocal music can be extracted in a reductionist fashion; taking a few measures at a time, a singer first can listen and then play with elements of melody, vowels, legato, rhythm, timbre, text, intention, etc. Through repetition of smaller elements, vocal learning is achieved more efficiently. The brain is not deft at attentive multitasking; in order to imprint productive habits, singular directives work best.[14] Diversity within that practice also gives rise to useful options for the brain's predictive coding pro-

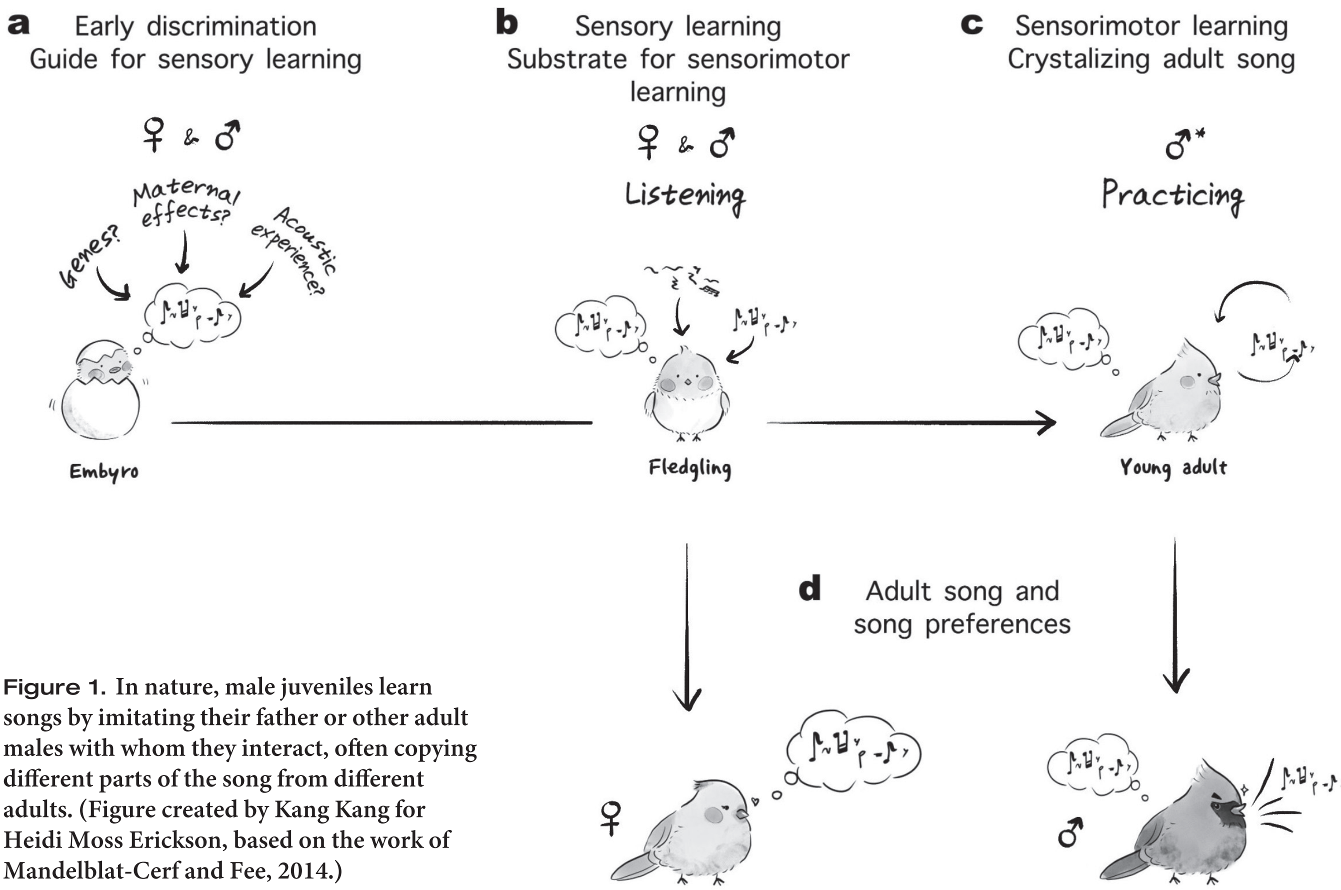

Figure 1. In nature, male juveniles learn songs by imitating their father or other adult males with whom they interact, often copying different parts of the song from different adults. (Figure created by Kang Kang for Heidi Moss Erickson, based on the work of Mandelblat-Cerf and Fee, 2014.)

cess.[15] In other words, the brain "decides" which motor elements to coordinate for a specific sound well before the sound is executed. Therefore, a variety of options serve the system better than "only one right way." For this reason, the auditory step cannot be overlooked, and should be varied and repeated. Audiation (hearing in one's head) primes the motor system, and many vocal muscles are activated through this process without any actual sound.[16] The same holds true for birds; incredibly, if their specific song is played while sleeping, the muscles of their syrinx (bird larynx) is activated.[17] Many singers have experienced the challenge of unlearning habits, but when practice is simplified and repeated in smaller units, a singer has a better chance of success and learning happens more quickly.

Another interesting element is the social requirement; both species require a live adult giving feedback and reward during the vocal learning process.[18] Sequential practice is under the watchful eye (and ear) of a tutor. The process involves dopamine, a neurotransmitter involved in motivation.[19] Each songbird develops independence via their own syntax and expression, creating sounds unique to them; babies babble, becoming increasingly individualized and complex. Both are rewarded by beaming caregivers who respond and guide in kind to the growing utterances. By extension, a voice student is affirmed by their teacher; the psychosocial relationship between teacher and student cannot be overemphasized and learning can be disrupted under excessive stress and criticism.[20] To this end, the dynamic needs to be a healthy one for it to function optimally in vocal learning, an oft overlooked element. The goal of the teacher, like the songbird, is eventually to foster independence and individuality, not codependence and hierarchy.

The functional similarities in vocal learning between human infants and songbirds are quite astounding. Even more revolutionary was the discovery of analogous neural structures, pathways, and genes between these

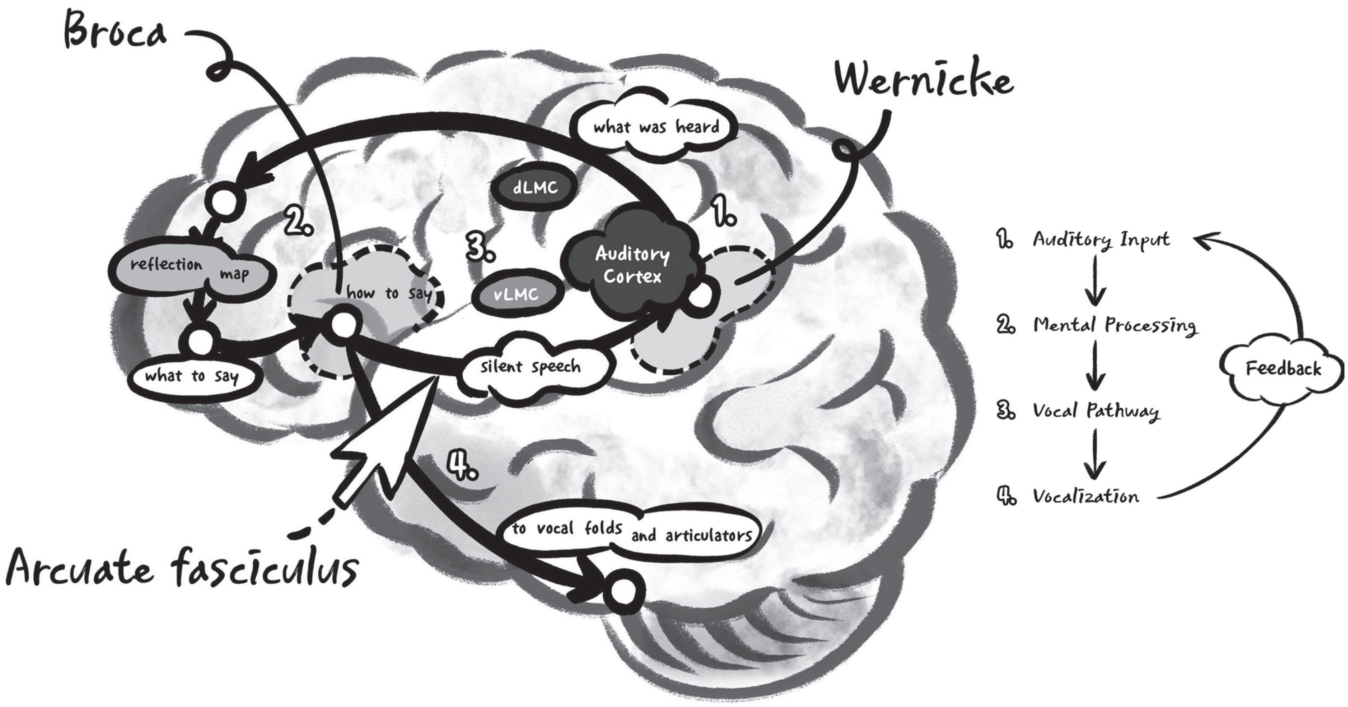

Figure 2. **Human vocal learning pathway. (Image created by Kang Kang and designed by Heidi Moss Erickson.)**

seemingly disparate species. Although it is beyond the scope of this article to delve into complex neuroanatomy, we can get an idea of the parallels by first observing a generalized overview of how a human brain vocalizes. Figure 2 shows some of the pathways involved in creating a vocalization.[21]

Singing involves a complex recruitment of more than 100 respiratory, laryngeal, and vocal tract structures. Given this intricacy, the system evolved in humans to include a robust feedback loop connected by a bundle of axons called the arcuate fasciculus to ensure correct output to target.[22] The arcuate fasciculus, which is involved in audiation and motor planning, has been shown to be more robust in trained singers and can be strengthened with practice.[23] Singers can capitalize on this ability through silent practice, a typically underutilized strategy.

In addition to the functional similarities between human and songbird vocalizations, the process is remarkably similar in the brain as well. In Figure 3, one can see the analogous areas involved in vocal learning and production for birds and humans. In addition to these structural elements, it is important to note that at the molecular level, fifty genes involved in vocal learning show similar patterns of activity in birds and humans, but are either not active or not present in nonhuman primates.[24] These genes allow for learning and plasticity in the brain during vocalization; singing practice will literally change your gene expression, so it is important to practice wisely.[25]

The discovery of these closely analogous pathways was revolutionary. But like many processes, the appearance was quite serendipitous; a simple twist of fate created the shared parallel circuitry. Similarities in disparate species, like sharks and dolphins, are a result of convergent evolution, which is nature's way of stumbling upon useful attributes more than once. Vocal learners, like humans and songbirds, are the result of this kind of convergent evolution. To account for this evolutionary phenomenon, Dr. Erich Jarvis at Rockefeller University proposed the motor theory of vocal learning.[26] In this model, vocal learning is a continuum consisting of both auditory and vocal motor components.[27] Species at the higher end of the continuum, like humans and songbirds, underwent a random duplication of the limb motor pathway adjacent to an innate vocal pathway. This

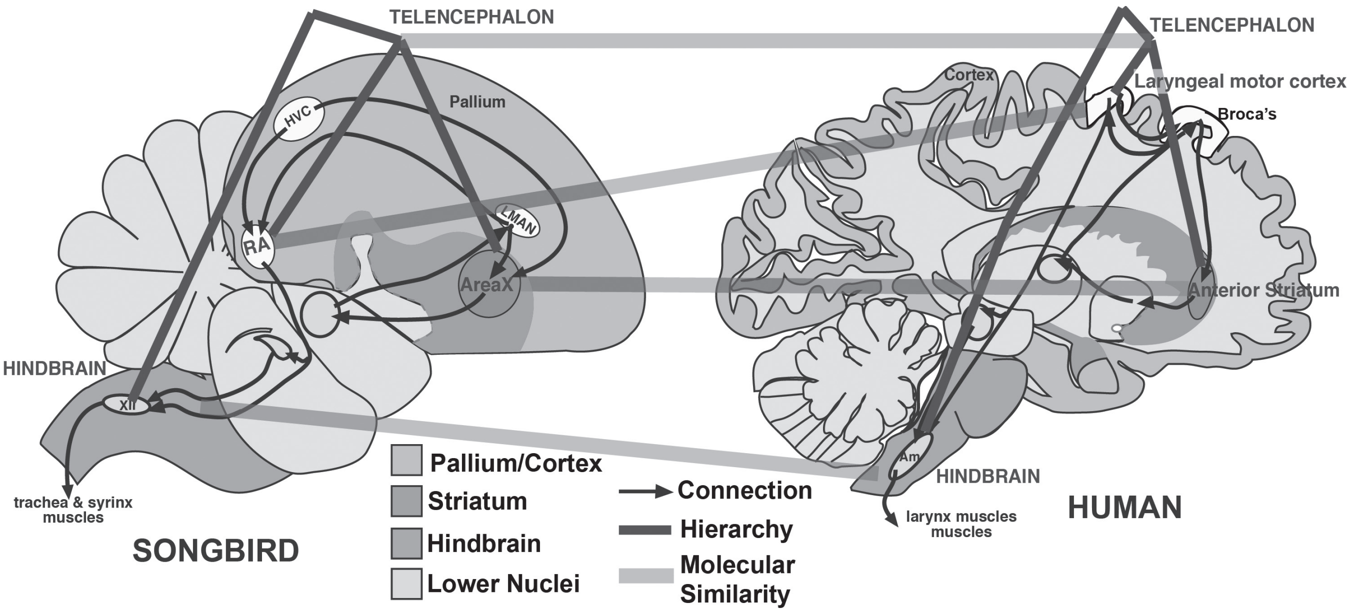

Figure 3. **Analogous schematic diagrams of brain areas involved in vocal learning for birds and humans. Coded for similarities at both the macro and micro levels. (Image from Erich Jarvis modified from Horita and Wada, 2011, and Pfenning et al., 2014.)**

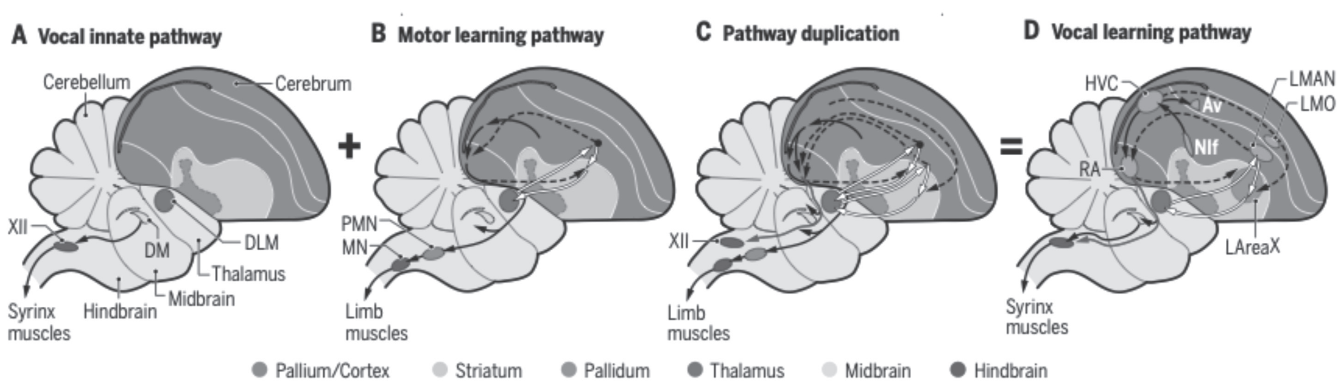

Figure 4. **Schematic of Dr. Erich Jarvis's motor duplication theory of vocal learning from his 2019 review in *Science*.**

resulted in the creation of a novel vocal motor pathway used for vocal learning (Figure 4).

Anecdotally, music teachers have been using movement and hand signals for singing activities, such as solfege and conducting. Gesture can be a useful tool for singers and can change the outcome of a vocalization in surprising ways, for example, tracing a circle for a phrase or wiggling fingers for vibrato. Interestingly, researchers have found that singing, even silently, helps Parkinson's patients improve their gait; strengthening neighboring neuronal pathways in singing influence the ones controlling their limbs.[28]

Evolution gives us a window beyond the mechanics to fully appreciate the depths of our art as singers, speakers, and humans. We can use these ideas to better design target strategies in voice pedagogy (for examples, see Figures 5 and 6). It is not enough to simply observe how things work in a complex mechanism. Singing is more than individual parts; it is a brain-body-environment continuum that requires an integrative understanding

EVOLUTION-BASED *VOCAL LEARNING* STRATEGIES FOR SINGING

Listening:

- Diversify sound inputs (the "tutor"): teachers, coaches, piano tracks, and professional recordings (if available). The more the merrier!
- Create customized "tutors" by making a wide variety of mini recordings. For example:
  - The melody alone, a few measures at a time, repeated ~3 times in a row (using a piano or instrument of choice).
  - The vernacular text, intoned text, and the text in rhythm independent of melody.
  - Combining various elements like speaking the text above the melody.
- Listening should be both active and passive (e.g., while doing the dishes).
- Take time in between each listen to pause and reflect. Space between days as well. This allows the brain to consolidate the input.

Practice:

- Do NOT sing *until you can hear* the excerpt you wish to practice in your head. This goes for exercises as well.
  - Start with humming or lip-trilling the melody as a first step.
- Record each practice step for listening:
  - Take a short section and practice the text, rhythm, vowels, melody, etc. separately.
  - Combine sequentially once each step feels secure. Repeat several times and then put the sections together.
  - Give attention to one specific element at a time: i.e., vibrancy, resonance, diction, breath, intention, etc. Distraction is also a strategy!
  - Play with transpositions, inversions, legato, staccato, etc. This is especially helpful in fretting more challenging phrases.
  - Develop independence of expression and syntax. Play with diverse interpretations, emotions, gestures, tempi, dynamics, and affect. Exercises also deserve artistic treatment! Variety is the key to the brain's ability to adapt and choose motor targets under different sets of circumstances.
- Intersperse silent practice with active practice.
- Remember that positive experiences increase dopamine which strengthens synapses, enhances learning, and inspires motivation: find what is working and apply to other contexts as much as possible.
- Use DO language rather than DON'T language.
- Understand that reductionist practice will make learning go faster, not slower, with less to undo in the end.

Customization:

- Some steps are easier or not as necessary. It is never one size fits all.
- Individuals vary so use your best judgement and instinct.
- Include time off and rest in practice schedules:
  - Rest is where learning is consolidated (stay tuned for more on that process in a future installment).

Figure 5. **Suggested practice strategies based on the science of vocal learning.**

Practice each pattern 1-6:

- Changing rhythms: e.g., dotted, triplets, etc.
- Changing registrations: e.g., a third higher/lower
- Combine in stages: weaving together over lines
- Play with characters, dynamics, tempo, emotion, etc.

Figure 6. **Vocal learning practice example: patterns.**

of both the how and the why. Those who work with the voice need to consider that active processes are not the only arena for optimization; there is a wealth of tools to use in nonsinging spaces where great strides can be made. The voice is an instrument unlike any other, and to that end, we need to educate all musicians about its uniqueness, not just singers. Future articles in this series will parse out these elements and offer novel approaches to singing that will enhance, clarify, and even overturn existing dogmas. We are so fortunate to have our incredible voices and complex brains that command them. Advances in neuroscience offer us new insights, giving the field of voice science a wonderful opportunity to evolve as well.

NOTES

1. Eric R. Kandel, Henry Markram, Paul M. Matthews, Rafael Yuste, and Christof Koch, "Neuroscience Thinks Big (and Collaboratively)," *Nature Reviews Neuroscience* 14, no. 9 (September 2013): 659–664.
2. Marina Davila Ross, Michael J. Owren, and Elke Zimmermann, "The Evolution of Laughter in Great Apes and Humans," *Communicative & Integrative Biology* 3, no. 2 (March 2010): 191–194.
3. Alexandra G. Rosati and Brian Hare, "Chimpanzees and Bonobos Exhibit Emotional Responses to Decision Outcomes," *PLoS ONE* 8, no. 5 (May 29, 2013): e63058.
4. Philip H. Lieberman, Dennis H. Klatt, and William H. Wilson,"Vocal Tract Limitations on the Vowel Repertoires of Rhesus Monkey and Other Nonhuman Primates," *Science* 164, no. 3884 (June 6, 1969): 1185–1187.
5. William Tecumseh Fitch, Bart de Boer, Neil Mathur, and Asif A. Ghazanfar, "Monkey Vocal Tracts Are Speech-Ready," *Science Advances* 2, no. 12 (December 2, 2016): e1600723.
6. Sonja C. Vernes, Buddhamas Pralle Kriengwatana, Veronika C. Beeck, Julia Fischer, Peter L. Tyack, Carel ten Cate, and Vincent M. Janik, "The Multi-Dimensional Nature of Vocal Learning," *Biological Sciences* 376, no. 1836 (October 25, 2021): 20200236.
7. Erich D. Jarvis, "Evolution of Vocal Learning and Spoken Language," *Science* 366, no. 6461 (October 4, 2019): 50–54.
8. Felipe Mora-Bermúdez, Farhath Badsha, Sabina Kanton, J. Gray Camp, Benjamin Vernot, Kathrin Köhler, Birger Voigt, et al., "Differences and Similarities between Human and

Chimpanzee Neural Progenitors during Cerebral Cortex Development," *ELife* 5 (September 26, 2016): e18683.

9. Erich D. Jarvis, "Selection for and against Vocal Learning in Birds and Mammals," *Ornithological Science* 5, no. 1 (July 2006): 5–14.
10. Fernando Nottebohm, "Ontogeny of Bird Song: Different Strategies in Vocal Development Are Reflected in Learning Stages, Critical Periods, and Neural Lateralization," *Science* 167, no. 3920 (February 13, 1970): 950–956.
11. Peter Marler, "Bird Calls: Their Potential for Behavioral Neurobiology," *Annals of the New York Academy of Sciences* 1016, no. 1 (June 2004): 31–44.
12. Yael Mandelblat-Cerf, Liora Las, Natalia Denisenko, and Michale S Fee, "A Role for Descending Auditory Cortical Projections in Songbird Vocal Learning," *ELife* 3 (June 16, 2014): e02152.
13. Dina Lipkind, Gary F. Marcus, Douglas K. Bemis, Kazutoshi Sasahara, Nori Jacoby, Miki Takahasi, Kenta Suzuki, et al., "Stepwise Acquisition of Vocal Combinatorial Capacity in Songbirds and Human Infants," *Nature* 498, no. 7452 (June 2013): 104–108.
14. Kevin Madore and Anthony D. Wagner, "Multicosts of Multitasking," *Cerebrum* (April 2019).
15. Luigi Grisoni, Bettina Mohr, and Friedemann Pulvermüller, "Prediction Mechanisms in Motor and Auditory Areas and Their Role in Sound Perception and Language Understanding," *NeuroImage* 199 (October 2019): 206–216.
16. Tim A. Pruitt, Andrea R. Halpern, and Peter Q. Pfordresher, "Covert Singing in Anticipatory Auditory Imagery: Psychophysiology," *Psychophysiology* 56, no. 3 (March 2019): e13297.
17. Sanne Moorman, Sharon M. H. Gobes, Maaike Kuijpers, Amber Kerkhofs, Matthijs A. Zandbergen, and Johan J. Bolhuis, "Human-like Brain Hemispheric Dominance in Birdsong Learning," *Proceedings of the National Academy of Sciences* 109, no. 31 (July 31, 2012): 12782–12787.
18. Naja Ferjan Ramírez, Sarah Roseberry Lytle, and Patricia K. Kuhl, "Parent Coaching Increases Conversational Turns and Advances Infant Language Development," *Proceedings of the National Academy of Sciences* 117, no. 7 (February 18, 2020): 3484–3491.
19. Kristina Simonyan, Barry Horwitz, and Erich D. Jarvis, "Dopamine Regulation of Human Speech and Bird Song: A Critical Review," *Brain and Language* 122, no. 3 (September 2012): 142–150.
20. Michael A.P. Bloomfield, Robert A. McCutcheon, Matthew Kempton, Tom P. Freeman, and Oliver Howes, "The Effects of Psychosocial Stress on Dopaminergic Function and the Acute Stress Response," *ELife* 8 (November 12, 2019): e46797.
21. Erich D. Jarvis, "Neural Systems for Vocal Learning in Birds and Humans: A Synopsis," *Journal of Ornithology* 148, no. S1 (December 2007): 35–44.
22. Psyche Loui, "A Dual-Stream Neuroanatomy of Singing," *Music Perception* 32, no. 3 (February 1, 2015): 232–241.
23. Gus F. Halwani, Psyche Loui, Theodor Rüber, and Gottfried Schlaug, "Effects of Practice and Experience on the Arcuate Fasciculus: Comparing Singers, Instrumentalists, and Non-Musicians," *Frontiers in Psychology* 2 (2011).
24. Stephanie A. White, "Genes and Vocal Learning," *Brain and Language* 115, no. 1 (October 2010): 21–28.
25. Shin Hayase, Hongdi Wang, Eri Ohgushi, Masahiko Kobayashi, Chihiro Mori, Haruhito Horita, Katsuhiko Mineta, Wan-chun Liu, and Kazuhiro Wada, "Vocal Practice Regulates Singing Activity–Dependent Genes Underlying Age-Independent Vocal Learning in Songbirds," *PLOS Biology* 16, no. 9 (September 12, 2018): e2006537.
26. Mukta Chakraborty and Erich D. Jarvis, "Brain Evolution by Brain Pathway Duplication," *Biological Sciences* 370, no. 1684 (December 19, 2015): 20150056.
27. Erich D. Jarvis, "Evolution of Vocal Learning and Spoken Language," *Science* 366, no. 6461 (4 October 2019): 50–54.
28. Masayuk Satohi and Shigeki Kuzuhara, "Training in Mental Singing While Walking Improves Gait Disturbance in Parkinson's Disease Patients," *European Neurology* 60, no. 5 (2008): 237–243.

Heidi Moss Erickson received a dual biology and voice degree at Oberlin College, where she worked in the voice lab of Richard Miller. Her more than 20 year performing career spans both opera and concert repertoire, with a focus on new music. She received her graduate degree in biochemistry with an emphasis in neuroscience. Her postgraduate research on telomeres at Rockefeller University led to prestigious publications, including a landmark paper in *Cell*. Heidi teaches vocal physiology at the San Francisco Conservatory of Music and applied voice at UC Davis. In 2007 she came down with a rare CNVII nerve injury which resurrected her passion for how the brain controls singing. Her courses and lectures have been featured both nationally and internationally at conferences and universities. www.heidimosserickson.com