are also new discoveries in the study of the development of musical skills in themselves. Several research teams report different kinds of responses to music in babies.

The presence of sophisticated responses to music in infants too young to have been culturally conditioned to respond in certain ways leads to the question of how and why musicality evolved. The issue, which already left Darwin scratching his head, still attracts a large number of hypotheses and lively debates.

Thus, when a journal published two “target articles” on the issue for debate, the pair attracted 60 responses, adding up to 149 pages of debate. Samuel Mehr and colleagues described the origins of music in terms of evolutionary psychology as a case of credible signalling, with rhythmic signals used in coalition forming and melodic ones more likely in infant care (Behav. Brain Sci. (2021) *44*, 23–29).

Patrick Savage from Keio University at Fujisawa, Japan, and colleagues aimed to synthesise many of the existing hypotheses under the umbrella concept of social bonding, which covers all kinds of bonding processes between humans and the resulting ties that bind them to form families, groups or societies (Behav. Brain Sci. (2021) *44*, 1–22). The authors argue that “the evolution of musicality involves gene–culture coevolution, through which proto-musical behaviors that initially arose and spread as cultural inventions had feedback effects on biological evolution because of their impact on social bonding.”

They suggest that music started from cultural innovations that became useful for bonding and thus influenced reproduction fitness and genetic selection. It can be compared to the way innovations like dairy making (Curr. Biol. (2018) *28*, R1171–R1173) or, much earlier, the use of fire for cooking food, have influenced genetic traits controlling our metabolism. Savage and colleagues also offer a range of predictions for cross-domain, cross-cultural, and neurobiological effects to be tested. Clearly, more research is needed.

Michael Gross is a science writer based at Oxford. He can be contacted via his web page at www.michaelgross.co.uk

---

Essay

AI and the Doctor Dolittle challenge

Yossi Yovel[1] and Oded Rechavi[2]

Talking to animals is a fundamental human desire. The emergence of powerful AI algorithms, and specifically Large Language Models, has driven many to suggest that we are on the verge of fulfilling this wish. A few large scientific consortia have been formed around this topic and several commercial entities even offer such services. We frame the task of communicating with animals as ‘The Doctor Dolittle challenge’ and identify three main obstacles on the route to doing so. First, although generative AI models can create novel animal communication samples, it is very difficult to determine their context, and we will forever be biased by our human *umwelt* when doing so. Second, using AI to extract context in an unsupervised manner must be validated through controlled experiments aiming to measure the animals’ response. This is difficult, and moreover, AI algorithms tend to cling on to any available information and are thus prone to finding spurious correlations. And third, animal communication focuses on a restricted set of contexts, such as alarm and courtship, highly limiting our ability to communicate regarding other contexts. Nevertheless, using the tremendous power of novel AI methods to decipher and mimic animal communication is both fascinating and important. We thus define the criteria for passing the Doctor Dolittle challenge and call upon scientists to take on the mission.

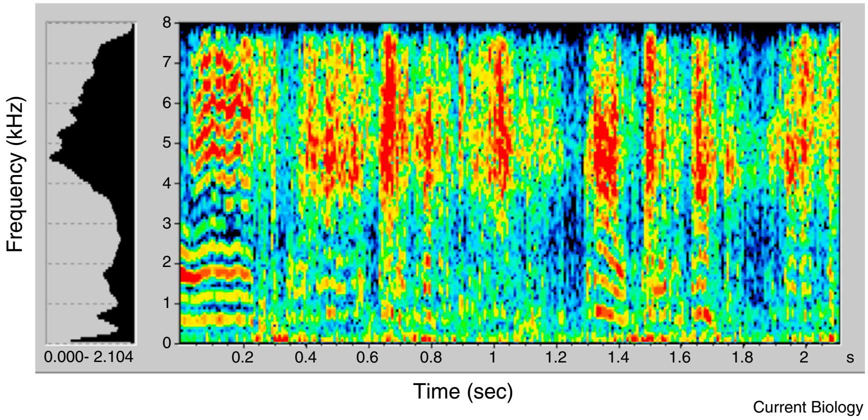
> *“...animals don’t always speak with their mouths, said the parrot in a high voice, raising her eyebrows. They talk with their ears, with their feet, with their tails, with everything”*
> — Polynesia the Parrot, The Story of Doctor Dolittle

Humans have always wanted to talk with animals. Whether it was King Salomon, Francis of Assisi, or Mowgli, almost every human culture evolved myths about people who can talk to animals. All around the globe, children strive to understand their pets and parents struggle to communicate with pre-lingual newborns. Talking with other organisms would not only enable empathizing with animals, learning about other levels of consciousness, and enriching our understanding of life, it would also have immense practical value. Farmers might benefit from a machine that allows them to ask a cow about its health, though it might turn them vegetarian. Such a machine would also come in handy once we land on a planet which is inhabited by some simple alien life-form, while if advanced aliens ever make it to earth, they will probably already have such a machine.

In this essay, we examine whether we can harness the power of the AI revolution and large language models (LLMs) in order to talk to animals. LLMs have shocked the world with their ability to talk to humans: these algorithms receive complicated sentences, interpret what the user wants, and compose new text, pretending efficiently to be human, although it is debated whether they really communicate with humans as humans do. Here, we ask whether similar algorithms can be developed to communicate with animals. Framed more popularly, we challenge the scientific community to create ‘CatGPT’ — a machine that allows us to chat with our pet cats.

Specifically, we will ask if we can develop a machine that: (1) communicates with an animal using *its own endogenous signals*; (2) does so in *various behavioral contexts*; and (3) works such that the animal *produces a measurable response* to it as if it were communicating with a conspecific and not a machine. Loosely paying homage to the Turing test, we term this the Doctor Dolittle challenge.

We expanded specifically on acoustic communication, but our discussion can be easily generalized to other communication modalities including vision and olfaction (which we touch on briefly). Importantly,



Current Biology

Figure 1. The spectrogram of a sequence of Egyptian fruit bat vocalizations generated using a generative neural network.
To generate the sequence, we first trained a generative adversarial network (GAN) model with several thousand calls of adult fruit bats and then used the model to generate a new sequence, never uttered before by a bat (based on our unpublished data).

we frame the challenge in terms of *communication*, not *language*. Animal communication falls short of human language in several key aspects. First, reduced semantics — there are only a few examples of animal signals which have an abstract meaning[1–3]. Second, poor syntax — although some studies suggest that animal vocal sequences are ordered, the importance of this order for generating meaning (compositional syntax) has been rarely demonstrated[1,4–6]. And third — limited vocal learning — with just a few exceptions[7–10], most animals exhibit a very limited ability to incorporate new signals into their repertoire based on exposure to new signals, as humans easily do.

What would it take for AI to talk with animals?

Generating animal *signals* is not difficult. We can record and mimic them and even generate new ones. Generative AI models, for example generative neural networks, can already mimic animal communication pretty well. These models have the ability to learn the statistics of the examples they are fed with, and to generate new examples. Just like they can generate a novel Van-Gogh like picture, they can also generate a novel vocalization (Figure 1 presents a sequence of Egyptian fruit bat social vocalizations that we generated in this way).

Current transformer networks — the type that LLMs are based on — and future AI algorithms will become even more efficient in learning data statistics and generating new data, but mimicking animal communication is not enough for passing the Doctor Dolittle challenge. To really communicate with animals, we must first decode their signals and then mimic their communication in the right context, to elicit the correct response. We identify several fundamental obstacles *en route* to achieving this.

First, *text* versus *context.* As demonstrated above, we can already generate new ‘text’, that is, new examples of communication, but we have no clue about the meaning of these examples or at least about their context. In order to decipher the context of communication, scientists typically annotate signals emitted in different contexts, and then examine if they can distinguish between them statistically. We have done this with fruit-bat vocalizations and have shown that we can distinguish the context of a vocalization based on its spectrum[11]. AI has definitely taken us forward in this respect, allowing us to work with high dimensional data and saving us the need to predefine the signals’ informative features. But when assigning context to signals, human scientists are limited by their human *umwelt*. We use the term umwelt here to express something that is beyond species-specific sensory perception and includes the entire human way of life. In the case of fruit bat communication, we could only name contexts such as feeding, sleeping or mating, based on our human-biased perception. This obstacle is strongly related to the current debate regarding ChatGPT’s ability to understand what it is generating: is it creating *text* or *context*?

Second, eliciting a correct natural response. One common approach to overcoming this human bias is to use an un-supervised approach where all data are divided into clusters based on some similarity criterion without any human annotation. To validate that this approach managed to reveal actual clusters of context, we must perform an experiment — we must measure a response. The animal must respond differently and appropriately to signals of different contexts. It should, for example, escape when we broadcast an alarm signal and approach or vocalize when we generate a contact signal. As we will demonstrate below, measuring a response is often not trivial because, in many cases, there is no clear external measurable response.

Importantly, because of its tremendous statistical power and ability to learn any statistical correlation available in the data, AI-based clustering should be treated with much caution. The literature is full of examples where AI attended background noise instead of actual signals (for example see[12]). Such errors are of very high probability when using real-world animal communication data (unlike when using cropped images or text as if often done). Moreover, a failure to generate a response could be the result of completely missing additional signals that are broadcast in channels that we humans have no direct access to[13]. We might, for example, be focusing on the acoustic cues salient to us, while missing the volatile cues that the animals secrete.

Notably, to pass the Doctor Dolittle challenge the animal should respond without learning, even when exposed to our signal for the first time. Humans have been teaching animals to communicate using associative learning for a very long time, and their abilities to learn in this way can be remarkable. One can train a dog to associate hundreds of actions with human vocal commands[14], but this essay focuses on communicating with animals using their own natural communication *signals*. Ecologists distinguish between *signals*, which evolved to convey a message delivered by a sender to a receiver, and *cues*, which might carry information although they have not evolved to do so[15]. The smell of dung attracts dung beetles from afar, although it is not a signal evolved to do so.

Third and last, the Wittgenstein obstacle: even if we manage to reveal *context* in animal communication and generate signals that elicit a response, we will (probably) never be able to communicate with animals about contexts that are not part of their communication repertoire. Because of the poor semantics and syntax of their communication, if cats do not talk about their feelings with each other or find puns funny, we will never be able to ask them 'how they feel' or explain that ChatGPT already means CatGPT in French (and that it might be funny).

This third obstacle might dramatically limit our capacity for inter-species communication as many studies suggest that much of animal communication concerns no more than the animal's arousal level[16,17], and if this is the case, there is not much to 'talk about' with these species. In other words, we might be able to develop an algorithm that encodes the extent to which our cat is angry with us for pushing it off the keyboard, but we can already do that today without sophisticated AI algorithms. Wittgenstein famously argued that: *"even if a lion could speak, we could not understand it*". His argument completes ours in suggesting that animal communication might not only be context poor, but it probably includes contexts that are completely alien to us. Human language is unique in this sense, as it allows describing phenomena that are beyond our perception, for example what vision is to a blind person. This will come in handy if we ever have to talk to intelligent aliens, who might have a completely different *umwelt* from ours.

Some examples

The honeybee dance is often raised as the best example of semantics in the animal kingdom. Using a specific series of movements, known as the 'waggle dance', bees transmit vectorial information about the location of available food. In this case, measuring the response of the animal to the signal (obstacle number two) is straightforward, because the receiver of the information flies to the location that was pointed to by the dance. The waggle dance is a unique case in which it seems that humans have (to some extent) hacked animal communication and thus might be able to communicate with them in one specific context. Indeed, a recent study[18,19] found that a robotic bee was able to recruit bees to follow the instructions of a dancing robot and to fly to a specific location in the field, thus fulfilling the Doctor Dolittle challenge criteria in one specific context.

As explained above, however, we will likely never be able to ask a bee 'how do you feel today' using the waggle dance (obstacle number three). Moreover, the dance probably entails much more information than meets the human eyes, including subtle tactile and acoustic signals about the quality of the resource[20]. These data would also need to be collected and fed into the AI algorithm if it were asked to crack the code, but we are not even sure which other types of data would need to be recorded, or what is the communication channel (obstacle number one). Will olfaction and acoustic[21] recordings suffice, or should we also record electric fields[22]? Humans have not yet deciphered the meaning of these signals, and we argue that the three above-mentioned obstacles will make it very difficult to fully do so.

There are a few additional examples where scientists managed to trick animals to respond to endogenous communication signals using robots[23]. Two such examples include a female-like robo-frog that was able to attract real male frogs to attempt mating with it[24] and a fish-robot that interacted with live fish during schooling behavior and affected their movement[25].

Notably, all these machines communicated with animals in a single context (foraging, mating or movement). Moreover, these are all cases that did not require any AI, as the relevant signals which elicit a response directly strike the eye. First coined by Niko Tinbergen, ecologists refer to such signals that stimulate a clear response as sign-stimuli. Similarly, the essentials of the waggle dance were deciphered nearly 100 years ago, before computers existed, using careful experiments, and all additional understanding gained since then has been secondary. We argue that it is not coincidental that only machines using very salient communication signals have so-far been devised. Much of animal communication is not about the salient signal, but rather about the fine details of how this signal is presented[26]. Like the mechanical nightingale of the Chinese emperor in Andersen's fairytale, none of these robots attempted to mimic the nuances of animal communication and would thus probably fail when competing with real conspecifics.

Another highly studied communication system is the courtship (and territorial) singing of songbirds, which is considered a prime example of communicative virtuosity[7]. Some songbirds exhibit advanced vocal learning abilities including an ability to learn syntactic rules defining the order or syllables in a sequence[27]. These abilities have led to the development of local song dialects[28], somewhat resembling human dialects. Despite these features, songbirds seem to have very poor semantics, limiting their role in intra-specific communication. A female conspecific can probably extract valuable information about the value of mating with a specific male based on its singing, but this does not leverage our ability to communicate with birds.

Studies on song-bird communicative signals such as alarm calls, not considered to be part of the song repertoire, have yielded some understanding of their communication which to some extent even allows us to hack it[29]. Researchers have identified various types of bird (and mammalian) alarm calls and they were even able to elicit an alarm response in these organisms using acoustic playback[30]. In fact, some animals mimic heterospecific alarm calls to trick them[31]. Just as in the case of the bee dance, however, scientists remain limited to communication in very specific contexts (such as alarm) and it is not clear how AI will take us farther.

On the other extreme of the communication spectrum, *Caenorhabditis elegans* nematodes primarily rely on chemical communication[32]. This model organism has a relatively simple, stereotypic, and fully mapped nervous system, composed of just 302 neurons. The worm genome encodes more than a thousand receptors that detect different chemicals[33]. *C. elegans* secretes many metabolites, for example pheromones such as ascarosides, to communicate with conspecifics. Advances in mass spectrometry-based metabolomics now make it possible to trace amounts of thousands of such chemicals and reveal that the bouquet of odors that the worms secrete is dynamic and changes depending on the context (for example it differs dramatically between sexes[34]). The worm's behavior is assessed by quantifying its movement, which, because the worm is approximately a curved line, can be fully represented by a vector of angles. To the human observer it appears that the worm has a limited set of behavioral repertoires[35] without a quantifiable syntax (order[32]). Importantly, it is already possible to simultaneously record the function of the entire *C. elegans* nervous system (using calcium imaging[36]). Could this type of data be used by an AI to talk with the worms? In other words, if we were to record the movement of millions of worms and in parallel measured their chemical secretion and brain activity at single-neuron resolution and fed all this information into a powerful AI algorithm, would it be able to pass the Doctor Dolittle criterion?

It is sometimes suggested that AI algorithms can simply analyze a mass of data and reveal patterns, for example find which worm bouquets are secreted under which posture and brain activity condition. Because of the limited number of postures that worms assume[35], movement is probably highly ambiguous and indistinguishable under many different contexts, so we will focus on their neuronal activity. Such a brute statistical force approach might yield new insight about worm communication, revealing for example that a certain bouquet is secreted in correlation with a certain neural state. Moreover, feeding unsupervised AI algorithms with continuous neural recordings will probably allow identification of distinguishable internal states[37], but will this allow us to understand worms, let alone communicate with them? Finding that a certain neural state correlates with a communicative signal simply pushes the problem to a different level — the neural level — but all three obstacles mentioned above will still hold.

For example, we will need to decide which contexts are relevant for the animal and what they mean to the animal, and we will also need to measure a response. While worms definitely secret chemicals to communicate, we do not know if they 'talk back' by secreting chemicals, and if not, what would be considered a response? Can a body turn of 30 degrees be considered a response? There are also many technical challenges. For instance, we will have to determine what are the relevant temporal timescales; would we need to measure secretions every 5 seconds or every hour? These are difficult questions that will remain difficult even if AI Doctor Dolittle's computational power would improve substantially.

Finally, as our closest relatives, primates and among them apes arguably make the most appealing case for inter-species communication, especially when searching for the origins of our own communication. Primates and apes communicate using a complex combination of vocalization[38], gestures and displays[39]. Starting with Richard Garner[40], who pioneered primate vocal recordings in the late eighteen-hundreds, aiming to decipher monkey language, much effort has been dedicated to decoding primate vocal communication. Garner himself was the first to use a play-back paradigm, recording primate vocalizations and then playing them back to other individuals. His reports were heavily doubted, and the play-back method was abandoned for nearly 100 years in primate research until Seyfarth *et al.*[41] used it to reveal predator-specific alarm calls in vervet monkeys. Attempts to teach apes human vocal language by rearing them with humans mostly failed[42], although some success has been achieved with sign language, suggesting that they might have a symbolic capacity while lacking the vocal abilities. In fact, the dogma in the field was that primates have no capacity for vocal learning, although more and more recent studies bring evidence that disproves this claim[43].

Considering all the above, a similar approach to the one proposed for *C. elegans* would be to record primate interactions continuously using both audio and video. This would be very difficult to do in the wild where the natural scenery would often occlude the behavior. Several recent machine learning algorithms can automatically track all the animal's joints[44], allowing posture to be represented as a three-dimensional vector, but unlike the situation in worms, this might not be enough for describing posture completely. A baboon might have the same posture in two different conditions: curious and vigilant/ready-to-escape, but while the muscles of the first will be relaxed, those of the latter will be contracted and ready for action. Still, this partial-posture data could be fed into a powerful learning algorithm which would probably find correlations between certain sequences and postures and certain vocalizations. However, the above-mentioned obstacles, such as how we will identify relevant contexts and what will be considered a response, will create high barriers here too, and thus it is not clear whether this will bring us closer to communicating with them.

In conclusion, the giant leap in AI has dramatically enhanced our ability to encode data and perform predictions on unseen data or even generate new data. These advances can also be applied to animal communication. They will likely yield a better understanding of animal communication, but they will not necessarily enable communicating with animals. We point to several obstacles that substantially hinder our ability to 'talk' with animals and argue that these obstacles are inherent and will not vanish even if the power of AI increases a million fold. Our view may seem pessimistic but we do not intend it to be discouraging; on the contrary, our goal is to highlight the challenges, but we brought several examples for partial success in communicating with animals.

Even if we will never be able to talk to animals in the human way, understanding how complex animal communication is and attempting to tap into it and mimic it is a fascinating scientific endeavor and it has multiple practical and commercial advantages. Moreover, there is still much that is unknown. Some species might use communication systems that are more complex than those demonstrated so far, perhaps, but not necessarily, using a sensory modality that is not easily accessible to us. The advances in neural recordings might also be used in ways that we currently cannot imagine, and these are just a few examples. We thus call on scientists to apply AI to decipher animal communication according to the Doctor Dolittle challenge criteria, that is, to develop machines that communicate with animals, using their own signals, in a variety of contexts, such that they respond to the machines as if they were responding to a conspecific.

ACKNOWLEDGMENTS

We would like to thank Andre Brown, Arnon Lotem, Ilya Ruvinsky and Roni Katzir for their critical and insightful comments, and Yoni Amit for helping with audio analysis. No AI tool was consulted with regarding the ideas presented in this text.

REFERENCES

1. Suzuki, T.N., Wheatcroft, D., and Griesser, M. (2020). The syntax–semantics interface in animal vocal communication. Philos. Trans. R. Soc. Lond. B *375*, 20180405.
2. Arnold, K., and Zuberbühler, K. (2006). Semantic combinations in primate calls. Nature *441*, 303–303.
3. Soltis, J., King, L.E., Douglas-Hamilton, I., Vollrath, F., and Savage, A. (2014). African elephant alarm calls distinguish between threats from humans and bees. PLoS One *9*, e89403. https://doi.org/10.1371/journal.pone.008940.
4. Cate, C., ten, and Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. Philos. Trans. R. Soc. Lond. B *367*, 1984–1994.
5. Zuberbühler, K. (2020). Syntax and compositionality in animal communication. Philos. Trans. R. Soc. Lond. B *375*, 20190062.
6. Amit, Y., and Yovel, Y. (2023). Bat vocal sequences enhance contextual information independently of syllable order. iScience *26*, 106466.
7. Hauser, M.D., and Konishi, Mark. (1999). The Design of Animal Communication (Cambridge, MA: MIT Press).
8. Prat, Y., Azoulay, L., Dor, R., and Yovel, Y. (2017). Crowd vocal learning induces vocal dialects in bats: Playback of conspecifics shapes fundamental frequency usage by pups. PLoS Biol. *15*, e2002556.
9. Janik, V.M. (2014). Cetacean vocal learning and communication. Curr. Opin. Neurobiol. *28*, 60–65.
10. Fernandez, A.A., Burchardt, L.S., Nagy, M., and Knörnschild, M. (2021). Babbling in a vocal learning bat resembles human infant babbling. Science *373*, 923–926.
11. Prat, Y., Taub, M., and Yovel, Y. (2016). Everyday bat vocalizations contain information about emitter, addressee, context, and behavior. Sci. Rep. *6*, 39419.
12. Artificial ignorance: The 10 biggest AI failures of 2017. TechRepublic https://www.techrepublic.com/article/the-10-biggest-ai-failures-of-2017/.
13. Hebets, E.A., Barron, A.B., Balakrishnan, C.N., Hauber, M.E., Mason, P.H., and Hoke, K.L. (2016). A systems approach to animal communication. Proc. R. Soc. Lond. B *283*, 20152889.
14. Pilley, J.W., and Reid, A.K. (2011). Border collie comprehends object names as verbal referents. Behav. Process. *86*, 184–195.
15. Smith, M.J., and Harper, D.G.C. (1995). Animal signals: models and terminology. J. Theor. Biol. *177*, 305–311.
16. Probert, R., Bastian, A., Elwen, S.H., James, B.S., and Gridley, T. (2021). Vocal correlates of arousal in bottlenose dolphins (Tursiops spp.) in human care. PLoS One *16*, e0250913.
17. Filippi, P. (2016). Emotional and interactional prosody across animal communication systems: A comparative approach to the emergence of language. Front. Psychol. *7*, 1393.
18. Landgraf, T., Bierbach, D., Kirbach, A., Cusing, R., Oertel, M., Lehmann, K., Greggers, U., Menzel, R., and Rojas, R. (2018). Dancing honey bee robot elicits dance-following and recruits foragers. Preprint at arXiv, arXiv:1803.07126.
19. Landgraf, T., Moballegh, H., and Rojas, R. (2008). Design and development of a robotic bee for the analysis of honeybee dance communication. Appl. Bionics Biomech. *5*, 157–164.
20. Łopuch, S., and Tofilski, A. (2020). Impact of the quality of food sources on the wing beating of honey bee dancers. Apidologie *51*, 631–641.
21. Michelsen, A., Towne, W.F., Kirchner, W.H., and Kryger, P. (1987). The acoustic near field of a dancing honeybee. J. Comp. Physiol. A *161*, 633–643.
22. Greggers, U., Koch, G., Schmidt, V., Dürr, A., Floriou-Servou, A., Piepenbrock, D., Göpfert, M.C., and Menzel, R. (2013). Reception and learning of electric fields in bees. Proc. R. Soc. Lond. B *280*, 20130528.
23. Klein, B.A., Stein, J., and Taylor, R.C. (2012). Robots in the service of animal behavior. Commun. Integr. Biol. *5*, 466–472.
24. Narins, P.M., Grabul, D.S., Soma, K.K., Gaucher, P., and Hödl, W. (2005). Cross-modal integration in a dart-poison frog. Proc. Natl. Acad. Sci. USA *102*, 2425–2429.
25. Faria, J.J., Dyer, J.R.G., Clément, R.O., Couzin, I.D., Holt, N., Ward, A.J.W., Waters, D., and Krause, J. (2010). A novel method for investigating the collective behaviour of fish: introducing “robofish”. Behav. Ecol. Sociobiol. *64*, 1211–1218.
26. Zahavi, A., and Zahavi, A. (1999). The Handicap Principle: A Missing Piece of Darwin’s Puzzle (Oxford: Oxford University Press).
27. Bloomfield, T.C., Gentner, T.Q., and Margoliash, D. (2011). What birds have to say about language. Nat. Neurosci. *14*, 947–948.
28. Kershenbaum, A., Ilany, A., Blaustein, L., and Geffen, E. (2012). Syntactic structure and geographical dialects in the songs of male rock hyraxes. Proc. R. Soc. Lond. B *279*, 2974–2981.
29. Suzuki, T.N. (2020). Other species’ alarm calls evoke a predator-specific search image in birds. Curr. Biol. *30*, 2616–2620.
30. Suzuki, T.N. (2018). Alarm calls evoke a visual search image of a predator in birds. Proc. Natl. Acad. Sci. USA *115*, 1541–1545.
31. Igic, B., and Magrath, R.D. (2014). A songbird mimics different heterospecific alarm calls in response to different types of threat. Behav. Ecol. *25*, 538–548.
32. Gomez-Marin, A., Stephens, G.J., and Brown, A.E. (2016). Hierarchical compression of Caenorhabditis elegans locomotion reveals phenotypic differences in the organization of behaviour. J. R. Soc. Interface *13*, 20160466.
33. Robertson, H.M., and Thomas, J.H. (2006). The putative chemoreceptor families of C. elegans. WormBook, 1–12.
34. Bobkov, Y.V., Walker III, W.B., and Cattaneo, A.M. (2021). Altered functional properties of the codling moth Orco mutagenized in the intracellular loop-3. Sci. Rep. *11*, 3893.
35. Stephens, G.J., Johnson-Kerner, B., Bialek, W., and Ryu, W.S. (2008). Dimensionality and dynamics in the behavior of C. elegans. PLoS Comput. Biol. *4*, e1000028.
36. Kato, S., Kaplan, H.S., Schrödel, T., Skora, S., Lindsay, T.H., Yemini, E., Lockery, S., and Zimmer, M. (2015). Global brain dynamics embed the motor command sequence of *Caenorhabditis elegans*. Cell *163*, 656–669.
37. Calhoun, A.J., Pillow, J.W., and Murthy, M. (2019). Unsupervised identification of the internal states that shape natural behavior. Nat. Neurosci. *22*, 2040.
38. Zuberbühler, K. (2018). Combinatorial capacities in primates. Curr. Opin. Behav. Sci. *21*, 161–169.
39. Arbib, M.A., Liebal, K., and Pika, S. (2008). Primate vocalization, gesture, and the evolution of human language. Curr. Anthropol. *49*, 1053–1076.
40. Burling, R. (2010). The simian tongue: The long debate about animal language by Gregory Radick. Language *86*, 244–247.
41. Seyfarth, R.M., Cheney, D.L., and Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. Anim. Behav. *28*, 1070–1094.
42. Suddendorf, T. (2013). The Gap: The Science of What Separates Us from Other Animals (New York: Basic Books Press).
43. Fischer, J., and Hammerschmidt, K. (2020). Towards a new taxonomy of primate vocal production learning. Philos. Trans. R. Soc. Lond. B *375*, 20190045.
44. Lauer, J., Zhou, M., Ye, S., Menegas, W., Schneider, S., Nath, T., Rahman, M.M., Di Santo, V., Soberanes, D., Feng, G., *et al*. (2022). Multi-animal pose estimation, identification and tracking with DeepLabCut. Nat. Methods *19*, 496–504.

[1]School of Zoology, Wise Faculty of Life Sciences & Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel.
[2]Department of Neurobiology, Wise Faculty of Life Sciences & Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel.
E-mail: yossiyovel@gmail.com (Y.Y.); odedrechavi@gmail.com (O.R.)