Bio-Linguistics: Monkeys Break Through the Syntax Barrier

W. Tecumseh Fitch
Department of Cognitive Biology, Faculty of Life Sciences, University of Vienna, Vienna, A-1090 Austria
Correspondence: tecumseh.fitch@univie.ac.at
https://doi.org/10.1016/j.cub.2018.04.087

Macaque monkeys can be trained to produce complex spatial sequences beyond the simplest levels of grammar previously known from animal studies. This indicates cognitive capabilities in the spatial-motor domain that approach the computational complexity level of human syntax.

The human capacity for language, allowing us to express any thought we can think, appears to be unique on the planet: although most animals communicate, none but humans show this unbounded expressive power. But our capacity to use and acquire language consists of multiple interlocking subcomponents, many of which are shared with other species [1]. For example, humans are the only living primate species known to be capable of learning and reproducing novel vocalizations, including words or melodies: but we share this capability with other, more distantly related species including birds, bats, seals, or elephants [2–4]. Shared traits are a boon to biologists interested in language, as animal models allow us to deploy a panoply of neuroscientific tools to understand their inner workings, and to test evolutionary hypotheses about adaptive function.

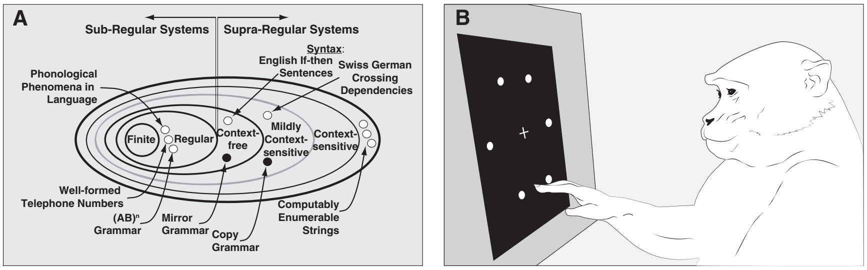
One component of language has until now resisted the search for parallels in our animal brethren: syntax. All human languages have at their core complex sets of rules which enable us to combine phonemes into syllables into words into sentences with precise, specific meanings. The last, most complex stage (words into sentences) requires rule systems — 'grammars' — previously thought to be beyond the capabilities of nonhuman animals [5,6]. In this issue of *Current Biology*, Jiang *et al.* [7] show that, with adequate training, monkeys can break beyond this barrier.

Working in Liping Wang's laboratory, together with Stanislas Dehaene, Jiang and colleagues [7] trained two rhesus macaques to produce structured sequences by pressing a touchscreen at specific locations arrayed around a circle. After intensive training, the monkeys could learn rules more complex than any previously demonstrated in nonhuman species. Most tellingly, they learned to produce mirror sequences following the pattern ABC|CBA, and generalized this ability to new sequence lengths. This is important because such 'mirror grammars' require computational capabilities beyond the simplest type, the so-called 'regular' or 'finite state' grammars. Like the grammars of all human languages, mirror grammars require a learner to possess 'supra-regular' computational abilities, which requires specific computational machinery not needed at the lower sub-regular level (Figure 1). The new results thus suggest that the monkey's brain possesses the kind of cognitive mechanisms required for human linguistic syntax, at least in this specific cognitive domain, and after intensive training.

To fully unpack the significance of this result requires a bit of computational theory. Formal language theory is a branch of mathematics, originating with the work of Alan Turing [8], that plays a central role in theoretical computer science [9]. It specifies the types of computational mechanisms required to cope with potentially infinite sets of strings — termed 'languages' — that obey certain constraints or follow certain patterns [10]. Note that, despite using such words as 'grammar' and 'language', this body of theory is not limited to human languages: it applies across diverse domains including mathematical expressions, music, visual patterns, or even well-formed telephone numbers. Any system where some strings are valid ('grammatical' or well-formed) and others invalid can be analysed using formal language theory, which tells us what type of computational system is needed to identify or generate that particular set of well-formed strings.

The most limited class consists of the regular languages, which require computational machinery termed 'finite state automata'. Such systems have a limited 'rote' memory, but limited ability to keep track of past occurrences or context. Abundant previous work shows that many nonhuman species possess, at least, this level of computational power. But human language has many examples where such limited systems are inadequate: for example a sentence with an 'if' will typically have, some arbitrary number of words later, the word 'then', and such if/then pairs can be nested within one another. This is one of many examples in linguistic syntax where complex tree structures are required to capture the empirical facts about human language, and crucially, supra-regular systems are needed to capture such patterns [9].

Previous animal work on supra-regularity started with the finding of [11] that cotton-top tamarins, while able to learn a simple regular rule ('repeat AB indefinitely') were unable to learn a closely matched supra-regular grammar termed A^nB^n (meaning 'some number of As followed by the same number of Bs'). This suggested that supra-regularity may represent a threshold between humans and other species. However, it used a habituation paradigm involving very little training. Later work using intensive training in songbirds appeared to show success on this same grammar [12] but, along with several subsequent studies [7], has been faulted on methodological grounds. The central empirical challenge



Current Biology

Figure 1. Macaques can master supra-regular grammars.

(A) The formal language hierarchy categorizes computational systems at different levels of complexity. Each small circle is a special case of the outer, enclosing circles. The outermost circle of all represents anything that is Turing-machine computable. Until now, high performance on supra-regular systems was known only for humans. (B) A rhesus macaque working on a touchscreen illustrates the basic paradigm developed by Jiang *et al.* [7], where a circular array of screen locations is used to encode grammars at varying levels of complexity. The monkeys in this study succeeded in mastering two complex grammars, at the supra-regular level, a first for nonhuman animals.

is to exclude the possibility that animals 'succeed' on a supra-regular task by inferring 'shortcuts' or heuristics at the finite-state level; animals have indeed been demonstrated in several studies to adopt such alternative strategies [13,14]. Although supra-regular, the A^nB^n grammar underlying most of this previous work has also been faulted for allowing strategies such as counting which have little relevance for the structural analyses at the heart of human languages.

The mirror grammar used by Jiang *et al.* [7] neatly avoids this problem, as simple counting or various finite-state shortcuts will not yield success. This grammar, combined with the novel experimental paradigm, makes the new work a paragon of how to perform this type of animal research. In addition to the mirror grammar, the monkeys in this study learned two additional grammars, including a 'repeat' or 'copy' grammar (requiring similar supra-regular resources to the mirror grammar). One monkey was able, impressively, to combine a spatial rule ('progress around the circle') with the mirror repetition, so that given 'A' it could produce the subsequent 'BC', followed by the entire 'CBA' mirrored sequence.

Finally, to put these achievements into perspective, the researchers tested pre-schoolers aged 5–6 years. Human children learned the tasks easily and almost instantly (requiring about five demonstrations), and vastly out-performed the monkeys.

This monkey/pre-schooler comparison suggests two rather different interpretations of the overall results. On the one hand, the new results are bad news for those who want to draw a strict line separating humans from other animals. On the other hand, the fact that monkeys require intensive training (tens of thousands of trials) to successfully learn the task, while children master it nearly perfectly with almost no training, suggests a major quantitative distinction between humans and macaques in this cognitive domain. While humans may not be the only species *capable* of mastering supra-regular systems, we may have an unusually strong *propensity* to do so [15]. I have previously dubbed this human propensity to infer tree structures 'dendrophilia' [16]; the new work suggests that macaques may be dendro-competent with training, but not dendrophilic by nature.

Future research can proceed on two major pathways. The first concerns the generality of macaque's capabilities: are they limited to this particular visuospatial domain and manual output modality, or can macaques generalize to other systems (such as auditory sequences)? If supra-regularity turns out to be domain specific, it may be that spatial cognition and/or motor control provided the original evolutionary function of supra-regularity, later exploited and strengthened during human evolution. Second, now that a task exists where monkeys successfully exhibit supra-regular abilities, neuroscientists can begin a full exploration of the neural mechanisms involved (for example, with intracranial recording, functional magnetic resonance imaging, and so on). Such research should greatly illuminate the brain mechanisms underlying human language. My money personally is on involvement of the region of the prefrontal cortex known as Broca's area, which exists in monkeys but is greatly expanded in humans [17,18]. It is possible that the massive expansion of this region in humans, combined with greatly increased connectivity, underlies the domain-general dendrophilia typifying our species [19].

In conclusion, after more than a decade of searching, a paradigm allowing a nonhuman species to break through the 'syntax barrier' has finally been found. While many other mechanisms are needed for full human language (including both complex semantics and the neural control required for speech), one central component of syntax can now be explored at the neural level. Be sure to watch this space!

REFERENCES

1. Fitch, W.T. (2010). The Evolution of Language (Cambridge: Cambridge University Press).
2. Stoeger, A.S., Mietchen, D., Oh, S., de Silva, S., Herbst, C.T., Kwon, S., and Fitch, W.T. (2012). An Asian elephant imitates human speech. Curr. Biol. *22*, 2144–2148.
3. Knörnschild, M. (2014). Vocal production learning in bats. Curr. Opin. Neurobiol. *28*, 80–85.
4. Nowicki, S., and Searcy, W.A. (2014). The evolution of vocal learning. Curr. Opin. Neurobiol. *28*, 48–53.
5. Fitch, W.T., and Friederici, A.D. (2012). Artificial grammar learning meets formal language theory: An overview. Philos. Trans. R. Soc. Lond. B *367*, 1933–1955.
6. ten Cate, C. (2016). Assessing the uniqueness of language: Animal grammatical abilities take center stage. Psychon. Bull. Rev. *24*, 91–96.
7. Jiang, X., Long, T., Cao, W., Li, J., Dehaene, S., and Wang, L. (2018). Production of supra-regular spatial sequences by macaque monkeys. Curr. Biol. *28*, 1851–1859.
8. Turing, A.M. (1937). On computable numbers, with an application to the Entscheidungsproblem. Proc. Lond. Math. Soc. *42*, 230–265.
9. Jäger, G., and Rogers, J. (2012). Formal language theory: refining the Chomsky hierarchy. Philos. Trans. R. Soc. Lond. B *267*, 1956–1970.
10. Hopcroft, J.E., Motwani, R., and Ullman, J.D. (2000). Introduction to Automata Theory, Languages and Computation, 2nd Edition (Reading, Massachusetts: Addison-Wesley).

11. Fitch, W.T., and Hauser, M.D. (2004). Computational constraints on syntactic processing in a nonhuman primate. Science *303*, 377–380.

12. Gentner, T.Q., Fenn, K.M., Margoliash, D., and Nusbaum, H.C. (2006). Recursive syntactic pattern learning by songbirds. Nature *440*, 1204–1207.

13. Ravignani, A., Westphal-Fitch, G., Aust, U., Schlumpp, M., and Fitch, W.T. (2015). More than one way to see it: Individual heuristics in avian visual cognition. Cognition *143*, 13–24.

14. van Heijningen, C.A.A., de Vissera, J., Zuidema, W., and ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. Proc. Natl. Acad. Sci USA *106*, 20538–20543.

15. Miller, G.A. (1967). Project Grammarama. In Psychology of Communication, G.A. Miller, ed. (New York: Basic Books).

16. Fitch, W.T. (2014). Toward a computational framework for cognitive biology: unifying approaches from cognitive neuroscience and comparative cognition. Physics Life Rev. *11*, 329–364.

17. Rilling, J.K., Glasser, M.F., Preuss, T.M., Ma, X., Zhao, T., Hu, X., and Behrens, T.E.J. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. Nat. Neurosci. *11*, 426–428.

18. Schenker, N.M., Hopkins, W.D., Spocter, M.A., Garrison, A.R., Stimpson, C.D., Erwin, J.M., Hof, P.R., and Sherwood, C.C. (2010). Broca's area homologue in chimpanzees (*Pan troglodytes*): probabilistic mapping, asymmetry and comparison to humans. Cerebr. Cort. *20*, 730–742.

19. Fitch, W.T. (2018). What animals can teach us about human language: The phonological continuity hypothesis. Curr. Opin. Behav. Sci. *21*, 68–75.

Division of Labor: How Microbes Split Their Responsibility

Babak Momeni
Department of Biology, Boston College, Chestnut Hill, MA 02467, USA
Correspondence: momeni@bc.edu
https://doi.org/10.1016/j.cub.2018.05.024

Within a biofilm, individual cells might perform only a subset of activities required for overall success of the biofilm. A new study examining matrix production, a task necessary for biofilm formation, shows possible mechanisms of genetic or phenotypic division of labor.

Division of labor — subpopulations performing different tasks simultaneously within an assembly — is pervasive in biology. In multicellular organisms, differentiation of cells into different tissues and organs is a clear example. Division of labor can happen among a group of individuals as well. In termite colonies, for instance, queens and kings are specialized for reproduction, workers forage and collect food, and soldiers primarily defend workers [1]. Division of labor is an important concept, allowing a better understanding of how complexity arises and is maintained in biological systems. A new study by Dragoš *et al.* [2], published in this issue of *Current Biology,* aims to reveal the molecular mechanism used by a common soil bacterium and model system, *Bacillus subtilis*, to explain the division of labor amongst cells engaged in biofilm matrix production.

Microbial assemblies, especially within a biofilm, are already known to demonstrate division of labor [3–11]. In *B. subtilis* biofilms, for example, even a genetically clonal population gets divided into subpopulations with drastically different activities [12]. Genetically similar cells in these biofilms differentiate into cells specialized for motility, matrix production, and sporulation, all of which are important for overall success of the biofilm. But how do the costs and benefits of division of labor for each cell favor the maintenance of such a scheme amongst individuals?

Revealing the mechanisms of how division of labor is implemented in a biological system is an important question. Take persistent microbial infections as an example: synergy among microbes through division of labor can make them more harmful to us. In chronic wounds, *Pseudomonas aeruginosa* and *Staphylococcus aureus* divide virulence tasks, making it harder for the immune system or antibiotics to suppress them [13]. Similarly, in the lungs of humans with cystic fibrosis, the differentiation of *P. aeruginosa* is thought to make it harder to treat the infection [14,15].

It is challenging to tease apart how division of labor takes place in natural settings in which underlying mechanisms are often muddied by the multitude of unknowns and uncertainties about individuals, interactions, and the environment. Therefore, in the new study, Dragoš and colleagues [2] focus on *B. subtilis* as a tractable system to explain the origination and maintenance of division of labor. The capability to monitor, control, and manipulate subpopulations in this system offers a direct way to mechanistically explore the maintenance of division of labor. Two techniques are primarily used: fluorescent markers are used to monitor the expression of relevant genes, and specific mutants are constructed to assess the impact of corresponding genes. This combination allows them to examine both genetic and phenotypic aspects of division of labor.

In particular, Dragoš *et al.* [2] focus on the production of an extracellular matrix, which is comprised of two major