


Mukherji, N. (2022). *The human mind through the lens of language: Generative explorations*. Bloomsbury. 280 pp. ISBN 9781350062689.

A Future Without a Past: Philosophical Consequences of Merge

Elliot Murphy¹ 

[1] Vivian L. Smith Department of Neurosurgery, University of Texas, Houston, TX, USA.

Biolinguistics, 2023, Vol. 17, Article e13067, <https://doi.org/10.5964/bioling.13067>

Received: 2023-10-20 • Published (VoR): 2023-10-31

Handling Editor: Kleantes K. Grohmann, University of Cyprus, Nicosia, Cyprus

Corresponding Author: Elliot Murphy, Vivian L. Smith Department of Neurosurgery, UTHealth, Houston, TX 77030, USA. E-mail: elliott.murphy@uth.tmc.edu

In his book *The Primacy of Grammar*, Nirmalangshu Mukherji (2010) claims that the development of the theory of Merge in theoretical syntax unveiled, for perhaps the first time in cognitive science, a “real joint of nature” (p. xviii) or “a (new) aspect of nature” (p. 235). I was lucky enough to stumble across this book during my undergraduate studies in English at the University of Nottingham, and a few months later Mukherji and I had become friends over social media; he a distinguished professor of philosophy in Delhi, India, and me in my over-priced halls of residence in Nottingham, England. Mukherji was even generous enough to provide feedback on a draft of my second book, which explored the philosophical foundation of the biolinguistics enterprise (Murphy, 2012). It is therefore with considerable joy that I now turn to reviewing his new book, *The Human Mind Through the Lens of Language: Generative Explorations*, an impressive foray into the Cartesian basis of modern generative grammar.

Mukherji’s central goal in the book is to answer the following question: What notion of the human mind follows from the study of language? That is to say, given the framework provided by generative grammar, what can be said of mentality more generally? Mukherji presents a brief overview of Cartesian linguistics; linguistics as part of the cognitive sciences; the relation between language and thought; economy conditions in minimalism; Merge as a factor in human-uniqueness; and the topic of whether Merge is “(perhaps) all we need”. This review will briefly interrogate some of the book’s core proposals.



The human mind “puts ideas together in different domains” (p. x), termed *Principle G* (‘G’ for ‘Generative’) in the book. Though never clearly formalized precisely, *Principle G* seems to be Mukherji’s way of noting the widespread use of limitless combinatoriality across cognitive systems, without being tied down to one particular domain of application. This is effectively the basic compositional machinery needed for various Languages of Thought (LoT), symbolic systems like language and music. The human mind exhibits a distinguishing “ability to entertain thoughts entrenched in a variety of sign or symbol systems” (p. 5). Mukherji, following Hauser and Watumull (2017), invokes a universal governing generative system to subsume language, “arithmetical thought, musical thought, artistic thought” under *Principle G* (see also McCarty et al., 2023). It remained unclear to me throughout the text how precisely Mukherji’s proposal differed in substantive ways from Hauser and Watumull’s proposal, although Mukherji’s general focus on the long history of these (domain-general) generative explorations certainly forms a major portion of the book’s early chapters. The emergence of *Principle G* may explain “both the unique origin of symbolic forms in humans and the subsequent production of structured thoughts in language that created human history and altered the character of the planet” (p. 231).

Mukherji claims that the human mind is “a set of structuring principles, probably a unit set”. This is *Principle G*, which he equates directly with the human mind (“Principle G is the human mind”; p. 13). Instead of seeing the mind as a combination of separate functions, like language, consciousness, attention, reasoning, and so on, Mukherji sees it as a single generative principle. Hence while Merge “satisf[ies] some of the major conditions” (p. 14) for *Principle G*, Mukherji wishes to re-formulate Merge in such a way as to make it accommodate the generativity of mind in its entirety, including extra-syntactic functions. He reviews, in later chapters, how one might go about achieving this.

Mukherji rightly sees the generative enterprise, though focused on language, as bearing possible implications for our understanding of compositional creativity more broadly: “knitting, cooking, yoga and gardening” (p. x) exhibit some abstract level of design, some structured relations between simple elements. Mukherji expounds on the Galilean style of science, through which the basic sciences have managed to achieve extraordinary explanatory depth through focusing on simple systems (“complex systems defy principled inquiry”; p. 1).

Early on, the book explores some historical background to the second cognitive revolution in the middle of the 20th century, touching on major topics like Turing computation, Gestalt psychology, the origins of generative grammar, and classical symbolic artificial intelligence. “For the first time in the history of thought”, he notes, the cognitive revolution, in particular theoretical syntax, gave us “a glimpse of what it means for language to be a ‘mirror of mind’” (p. 66). One of Mukherji’s main complaints is that “there is very little connection between the formal and empirical studies of human language and related aspects of human cognition and the rest of the topics pursued

in cognitive science” (p. 54). There have been few attempts to migrate the concerns of theoretical linguistics to other domains of the cognitive sciences (see [Murphy, Holmes, & Friston, 2022](#)). Mukherji’s charge is legitimate, but given his philosophical interests there is limited discussion of some of the more recent developments in cognitive neuroscience and experimental psychology that bear on these topics.

Mukherji then turns to Chomsky’s discussions concerning the formulation of mentality and physicality (e.g., [Chomsky, 2000](#)). Mukherji seeks to find contradictions within Chomsky’s writings about philosophy of mind and the mind-body problem, but he mostly poses questions about his positions that have the rhetorical force of a complaint but do not concretely state the source of the putative category error or contradiction. But for the interested reader, there are some more nuanced sections of text that review Chomsky’s writings on the boundaries of the mental.

Mukherji claims that Chomsky does not centralize in his Cartesian philosophy the human-unique aspects of language. Some readers may be already convinced by this stance, although I do not see how this is the case; much recent work by Chomsky and collaborators has been dedicated to highlighting unique components of (capital) MERGE, its restrictions, and the complexities of lexical representations that appear to belie any unification with non-human conceptual systems.

Where the differences become clearer is perhaps found in how Mukherji deviates from [Chomsky’s \(1966\)](#) traditional Cartesian linguistics. For Mukherji, Chomsky’s conception “falls short of the distinction between mind and cognition” (p. 6). Mukherji claims that Chomsky’s modularity of mind thesis should be replaced in favor of his *Principle G*, which seeks to unify structured symbolic representations across language, music and arithmetic. For Mukherji, the principles of language are *not* specific to the domain of language. Mukherji’s focus is placed mostly on this Cartesian theme of philosophical framing, and a little less on the ‘Galilean style’ of methodological naturalism (for this, see [Collins, 2023](#)).

Mukherji provides some interesting critiques of recent minimalist proposals (p. 154). Consider how Merge is strictly binary-branching, making syntax a minimal kind of algebra, termed a magma. Mukherji objects to theories based on the idea that because minimalism uses binary branching in much of its formalism, these operations can be rewritten in the binary notation vector of physics (e.g., [Piattelli-Palmarini & Vitiello, 2015](#)). Steps of syntactic derivation could then be called vector states, but nothing in terms of explanation is gained by comparing these remote formal systems. Mukherji is quick to call out these and other similar proposals that purport some force towards explanatory adequacy yet ultimately appear an awful lot like re-description.

The book also reviews a range of fascinating history into Enlightenment thought. He assesses Descartes’s *Compendium on musick*, published posthumously, in which Descartes claims music to be a “passion of the soul” and examines some of its mathematical foundations and how formal structure determines aesthetics. Mukherji relates various

strands of Enlightenment thinking to motivate a more general conception of how *Principle G* provides structured mental representations over variable domains (mathematics, music, etc.). For Descartes, animals did not have ‘thought’ because they cannot use signs. Much the same for Mukherji’s worldview.

The following chapters in the book then advocate for a saltationist account of cognitive evolution (a punctuated, sudden and large mutational change across generations) yielding *Principle G*, followed by a gradualist phase interfacing *Principle G* with other representational systems. A comprehensive evolutionary timeline is presented concerning when Merge, the lexicon, and the interfaces may have emerged (see [Murphy, 2019](#)).

Mukherji turns to cognitive neuroscience mid-way through the book. He writes that “[s]plit-brain research and imagery of brain function [...] say interesting [...] things on the localization of cognitive abilities” (p. 69). But, he notes, such research is not “*explanatory*”. This is a fairly common critique. Mukherji, nevertheless, is justified in critiquing much current work in cognitive neuroscience of language that does little to formalize the initial object of inquiry (“What do we mean by language?”), and little to ground features of this object directly within endogenous properties of neural systems (for some recent attempts, see [Murphy, 2015, 2020, 2023](#)). As Mukherji puts it (p. 69), “the limitations of split-brain research is essentially linked to the limitations of the neuroscientist’s answer to the question ‘what is [language]?’” Effectively moving against some contemporary ‘inside-out’ accounts of cognitive neuroscience (e.g., [Buzsáki, 2019](#)), Mukherji notes that “no notion of competence will emerge from simply looking at the brain” (p. 69).

Mukherji cites some recent work exploring “how the hierarchical structure of syntactic processing is represented in the brain” (p. 70), discussing [Nelson et al. \(2017\)](#) and others. Nelson et al. is discussed by Mukherji in connection with reported effects of ‘node closure’ of syntactic constituents and high gamma activity increases at points of structure composition. However, [Woolnough et al. \(2023\)](#) note that Nelson et al. did not consider word frequency in this analysis, and show that once this is taken into consideration, most of these effects of syntactic structure-building in intracranial high gamma activity cannot be replicated. [Woolnough et al. \(2023\)](#), [Murphy, Woolnough, et al. \(2022\)](#) and [Murphy \(2023\)](#) provide some alternative avenues to seek neural signatures of syntax.

Mukherji also mentions recent work arguing that “periodic oscillations in certain areas of the brain [...] seem to relate to what linguists call *recursive hierarchical phrase structure*”. Mukherji then pushes back against Friederici’s notion of ‘grounding’ syntax in the brain through her neuroimaging work (e.g., [Friederici et al., 2017](#)), and notes that it is “totally unclear what ‘grounding in the brain’ means beyond the trivial observation that certain events of grammatical construction are associated with certain neural events” (p. 70). He even goes so far as to say that “there is nothing in the *formulation* of the computational system of human language that requires that the system cannot be located in the knee joints” (p. 70). Given that we know enough now about feedforward/feedback,

bottom-up/top-down interactions, cyclic connectivity trajectories, and cortical spirals and travelling wave architectures amongst still more abstract neural architectures, I would disagree here, and note instead that since Merge requires some organic system that is capable of recursive structure-building, the human brain seems to be a much better candidate than knee joints. Nothing in the knees can ever *in principle* satisfy the needs of cyclic self-call, whereas a host of dynamical neural processes certainly can.

Still, it remains surely true that “[t]here is nothing in the description of the brain that tells us anything about linguistic categories” (p. 71). Mukherji goes further and claims that “[n]eural research throws no additional light on the properties of language at all. For all its fancy technological show, the entire discipline of neurolinguistics is basically an afterthought”. (p. 71). These observations, radical though they may appear, effectively amount to highlighting how challenging conducting interdisciplinary research is (it is an ongoing task to try and bridge computational with algorithmic and implementational accounts), and indeed they also indirectly reflect some of the sociological boundaries, grounded in some potentially hidebound attitudes, that block the way towards multi-disciplinary action, and which serve to stymie and discourage inquiry outside purely theoretical domains. Mukherji repeatedly stresses, for example, that “we do not know *how* [Principle G] was made available to early modern humans, not to mention *when*” (pp. 110–111). This is undoubtedly correct—but this difficulty is something that both the theoretician and experimentalist have to grapple with.

Reading Mukherji more charitably, he is indeed correct to note that, regardless of advances in the brain sciences, and quite independent of the biological explanation for the language faculty, “biolinguistics has already uncovered an aspect of nature in its own terms” (p. 71). This, for Mukherji, is essentially Merge cast as domain-general. As mentioned, this amounts to a universal generative faculty that, when applied to words, generates thoughts in the language system; “it works on tones to generate the musical system, on social relations to generate the kinship system” (p. 73). The mind cannot be reduced to linguistic thought; it is “a primaeval system that gave rise to a range of more complex closely related systems” (p. 101). With respect to simplex lexical items, “the origin of these atomic concepts continues to be a mystery” (p. 230).

For mental domains, Mukherji contends that “the computational system uses the general ‘law of generativity’ to construct structures in accordance with the properties of domain-specific lexicon” (p. 223). He believes that the task of the cognitive scientist is to provide an optimal, parsimonious account of these structures. When reviewing recent formulations of (capital) MERGE, Mukherji concludes that the workspace-based re-definition looks like “a general operation of the mind implementing Principle G in mental systems” (p. 226), but skips over the various syntax-specific restrictions on MERGE, like Resource Restriction, Minimal Yield, amongst some others discussed in current inquiry (Chomsky, 2023; Mizuguchi, 2023; Murphy & Shim, 2020). MERGE is inherently Markovian; it cannot “see” and has no memory of what came before it, it

only applies itself to a given workspace with no extensive lookback memory into past applications. The history of the derivation is not preserved in the current stage, and it is Minimal Yield that makes derivations strictly Markovian in that the next step has no access to derivational history. This, critically, renders linguistic recursion distinct from other types of recursion. It is unclear how Mukherji's system would deal with these domain-specific restrictions on application, but they provide exciting avenues to direct future work in philosophy of mind.

Perhaps the most important message of Mukherji's book is that the principles of linguistic theory do not simply form a part of mental life; they in fact help characterize the central notion of the human mind. By centralizing the generative structure-building capacity in his philosophy of mind, Mukherji successfully revives a more classical theme that linguistics should be seen as the 'queen' of the cognitive sciences.

Funding: The author has no funding to report.

Acknowledgments: The author has no additional (i.e., non-financial) support to report.

Competing Interests: The author has declared that no competing interests exist.

References

- Buzsáki, G. (2019). *The brain from inside out*. Oxford University Press.
<https://doi.org/10.1093/oso/9780190905385.001.0001>
- Chomsky, N. (1966). *Cartesian linguistics*. Harper and Row.
- Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511811937>
- Chomsky, N. (2023). *The miracle creed and SMT*. University of Arizona & MIT.
- Collins, J. (2023). Generative linguistics: 'Galilean style'. *Language Sciences*, 100, Article 101585.
<https://doi.org/10.1016/j.langsci.2023.101585>
- Friederici, A. D., Chomsky, N., Berwick, R., Moro, A., & Bolhuis, J. (2017). Language, mind and brain. *Nature Human Behaviour*, 1, 713–722. <https://doi.org/10.1038/s41562-017-0184-4>
- Hauser, M. D., & Watumull, J. (2017). The universal generative faculty: The source of our expressive power in language, mathematics, morality, and music. *Journal of Neurolinguistics*, 43(B), 78–94.
<https://doi.org/10.1016/j.jneuroling.2016.10.005>
- McCarty, M. J., Murphy, E., Scherschligt, X., Woolnough, O., Morse, C. W., Snyder, K., Mahon, B. Z., & Tandon, N. (2023). Intraoperative cortical localization of music and language reveals signatures of structural complexity in posterior temporal cortex. *iScience*, 26(7), Article 107223.
<https://doi.org/10.1016/j.isci.2023.107223>
- Mizuguchi, M. (2023). MERGE, minimal yield, and workspace accessibility. *Linguistic Research*, 40(1), 27–65. <https://doi.org/10.17250/khisli.40.1.202303.002>

- Mukherji, N. (2010). *The primacy of grammar*. MIT Press.
<https://doi.org/10.7551/mitpress/9780262014052.001.0001>
- Murphy, E. (2012). *Biolinguistics and philosophy: Insights and obstacles*. Lulu.
- Murphy, E. (2015). The brain dynamics of linguistic computation. *Frontiers in Psychology*, 6, Article 1515. <https://doi.org/10.3389/fpsyg.2015.01515>
- Murphy, E. (2019). No country for Oldowan men: Emerging factors in language evolution. *Frontiers in Psychology*, 10, Article 1448. <https://doi.org/10.3389/fpsyg.2019.01448>
- Murphy, E. (2020). *The oscillatory nature of language*. Cambridge University Press.
<https://doi.org/10.1017/9781108864466>
- Murphy, E. (2023). *ROSE: A neurocomputational architecture for syntax*. arXiv.
<https://doi.org/10.48550/arXiv.2303.08877>
- Murphy, E., Holmes, E., & Friston, K. (2022). *Natural language syntax complies with the free-energy principle*. arXiv. <https://doi.org/10.48550/arXiv.2210.15098>
- Murphy, E., & Shim, J.-Y. (2020). Copy invisibility and (non-)categorical labeling. *Linguistic Research*, 37(2), 187–215. <https://doi.org/10.17250/khisli.37.2.202006.002>
- Murphy, E., Woolnough, O., Rollo, P. S., Roccaforte, Z. J., Segaert, K., Hagoort, P., & Tandon, N. (2022). Minimal phrase composition revealed by intracranial recordings. *The Journal of Neuroscience*, 42(15), 3216–3227. <https://doi.org/10.1523/JNEUROSCI.1575-21.2022>
- Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., Cash, S. S., Naccache, L., Hale, J. T., Pallier, C., & Dehaene, S. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences of the United States of America*, 114(18), E3669–E3678. <https://doi.org/10.1073/pnas.1701590114>
- Piattelli-Palmarini, M., & Vitiello, G. (2015). Linguistics and some aspects of its underlying dynamics. *Biolinguistics*, 9, 96–115. <https://doi.org/10.5964/bioling.9033>
- Woolnough, O., Donos, C., Murphy, E., Rollo, P. S., Roccaforte, Z. J., Dehaene, S., & Tandon, N. (2023). Spatiotemporally distributed frontotemporal networks for sentence reading. *Proceedings of the National Academy of Sciences of the United States of America*, 120(17), Article e2300252120. <https://doi.org/10.1073/pnas.2300252120>