Forum

Rebuttal to "Merge Is Not 'Lerge'"

Jeffrey Watumull[1], Ian Roberts[2]

[1] *Oceanit Laboratories, Honolulu, HI, USA.* **[2]** *Department of Linguistics, University of Cambridge, Cambridge, United Kingdom.*

Abstract

This paper is a rebuttal to Gärtner (2023, https://doi.org/10.5964/bioling.11715).

Keywords

merge, lerge, Leibniz

It was our contention in "Leibnizian Linguistics" (Roberts & Watumull, 2015) that one could read Leibniz to have formulated a precursor to Merge—the basic set-formation operation of I-language posited within the minimalist program (see Seely et al., 2022). We wrote:

> "[Leibniz] introduced a special new symbol ⊕ to represent the combining of quite arbitrary pluralities of terms. The idea was something like the combining of two collections of things into a single collection containing all of the items in either one" (Davis 2012: 14-15). This operation is in essence formally equivalent to the *Merge* function in modern syntactic theory (Chomsky 1995); [...]. Leibniz defined some of the properties of ⊕—call it *Lerge*—thus:
>
> (1) X ⊕ Y is equivalent to Y ⊕ X.
>
> (2) X ⊕ Y = Z signifies that X and Y "compose" or "constitute" Z; this holds for any number of terms.

> "Any plurality of terms, as A and B, can be added to compose a
> single term A ⊕ B." Restricting the plurality to two, this describes
> Merge exactly: it is a function that takes two arguments, α and
> β (e.g., lexical items), and from them constructs the set {α, β} (a
> phrase). (We can also see that ⊕ shares with Merge an elegant
> symmetry, as (1) states.) And according to Leibniz's principle of the
> *Identity of Indiscernibles*, if Merge and Lerge are formally indiscerni-
> ble, they are identical: Merge *is* Lerge.

Lo, these many years later, Gärtner (2023) asserts to the contrary that "Merge is Not 'Lerge'" because Lerge, unlike Merge, is associative. According to Swoyer (1994), ⊕ is governed by a number of axioms:

(A1) Commutativity: A ⊕ B = B ⊕ A

(A2) Idempotentce: A ⊕ A = A

(A3) Associativity: (A ⊕ B) ⊕ C = A ⊕ (B ⊕ C)

The first two axioms are unproblematic for the Merge = Lerge conjecture. Merge generates unordered 2-sets, and therefore conforms to A1. A2 is simply the case of Internal Merge applied to a single lexical item, yielding the successor function (see Chomsky, 2008). The allegation is that A3 is "where the Leibnizian perspective and minimalist Merge definitely part ways" because, for instance, expressions like {X, {Y, Z}} and {{X, Y}, Z}} can "differ [...] in terms of thematic structure and/or grammatical functions"; and "given these discernible differences regarding associativity, we are forced to conclude that Merge is not 'Lerge.'" (Gärtner, 2023).

The conclusion is simply a *non sequitur*.

First, a metamathematics/metalinguistics point. Merge is binary set-formation, *tout court*. Associativity is a problem only if one takes a myopic picture of Merge: that is, only if one sees it as a *constructive* operation—analogous to a constructive proof—where the order of operations is all that matters. However, as Watumull and Chomsky (Forthcoming) argue, there is the other side of the coin: the *classical* side, analogous to a classical proof, where *all possible applications of Merge apply*, such that {X, {Y,Z}} and {{X,Y}, Z}—amongst other structures—*are* generated. (They are enumerated in the range of the function.) Analogously, the axioms of arithmetic generate—the *extension* of the *intension* includes—(A + B) + C and A + (B + C), and every other possible combination. The order of operations only matters when we seek to understand what parts of our knowledge we can *use*, factoring in third factors, etc. (see Chomsky, 2023).

Secondly, and less esoterically, associativity is an *additional* operation that can be applied to the outputs of Merge and Lerge. It just so happens that the application to Merge has effects that end up mattering for the "thematic structure and/or grammatical functions" of language. But those effects are external to the internal properties of Merge

and the structures it generates. (Alternatively, associative is a relation that can—but need not—be defined over—read onto—the structure Merge generates.) Consider: Given three objects (X, Y, Z), External Merge can combine any two into a set and then merge that set with the third object. (We can set aside Internal Merge here, but the point is the same.) Ditto for Lerge. *QED*. This is all that needs to be said. The fact that with Merge {{X,Y}, Z} and {X, {Y, Z}} constitute different thoughts is irrelevant. The computational/combinatorial properties of Merge and Lerge are identical in the syntactic (set theoretic) sense. The fact that {{X,Y}, Z} and {X, {Y, Z}} mean different things is due to extra-set-theoretic factors. Merge generates all possible structures, only some of which converge (where convergence is dependent upon satisfying factors extrinsic to Merge).

Third, Merge and Lerge do not apply to distinct domains because, if Watumull and Chomsky (Forthcoming) are correct, the syntactic and semantic domains are not distinct. Language (the calculus) = Thought (the semantics). All structures/thoughts exist (i.e., Merge generates the infinite set), but only a subset (nonetheless infinite) can be used (and an even smaller subset, finite for now (i.e., so long as we are finite beings) are in fact used). Those which can be used are those that can be constructed consistent with third factors (e.g., principles of computational efficiency, etc.). All of this relates to distinguishing knowledge and use, competence and performance, classical and constructive proofs, etc.

Fourth, to be scholarly scrupulous, we should note that Leibniz did *not* define ⊕ as associative, contrary to what Gärtner asserts. His authority, Swoyer (1994), himself admits: "Leibniz does not include this axiom in his calculus". The reason Swoyer stipulates it is because Frege and others thought it necessary, and that "several of Leibniz's proofs fail without it". We simply note that perhaps Leibniz knew what he was doing.

Finally, let us imagine, *arguendo*, that we are wrong that Leibniz's formulation of Lerge is equivalent to Merge. Imagine we are engaged in historical pareidolia, seeing Merge where it does not exist. Such "misreadings" of history would not undermine our analysis one jot, for our "Leibnizian Linguistics" is modelled on Chomsky's "Cartesian Linguistics", and as Watumull and Chomsky (Forthcoming) explain:

> "The history discussed is neither exhaustive nor conventional, but decidedly selective and idiosyncratic". "[I]t is meant to excavate and preserve foundational ideas—in their original, explicit, and often most cogent formulations—that structure our approach for expository coherence, rhetorical power, and logical force. 'A knowledge of the historic and philosophical background gives that kind of independence from prejudices of his generation from which most scientists are suffering. This independence created by philosophical insight is—in my opinion—the mark of distinction between a mere artisan or specialist and a real seeker after truth' (Einstein 1944). This approach is 'not to be confused with efforts […] to reconstruct

exactly how the issues appeared and how ideas were constructed at an earlier time' (Chomsky 1979), as in traditional historical scholarship. Our equally legitimate approach, successfully adopted by others in other domains (e.g., Popper 1945/2013, Chomsky 1966/2009, Feyerabend 1975/2010, Lakatos 1976/2015), is 'to recover insights that have long been neglected, approaching earlier work [...] from the standpoint of current interests and trying to see how questions discussed in an earlier period can be understood, sometimes reinterpreted, in the light of more recent understanding, knowledge, and technique'. For instance, were we studying seventeenth century art, we would 'not [be] proceeding in the manner of an art historian so much as that of an art lover, a person who looks for what has value to him in the seventeenth century, for example, that value deriving in large measure from the contemporary perspective with which he approaches these objects. Both types of approach are legitimate' (Chomsky 1979)".

As we explained at length in our original work, there is considerable value to be gained by understanding Merge as a descendant of Lerge, just as it is insightful to understand the Turing machine as a descendent of Leibniz's *calculus ratiocinator* and *characteristica universalis*. This exercise does not distort history, but illuminates the consilience and universality of mathematical and computational notions discovered to typify human intelligence.

References

Chomsky, N. (2008). On phases. In R. Freiden, C. P. Otero, & M. L. Zubizarreta (Eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud* (pp. 133–166). MIT Press.

Chomsky, N. (2023). *Working toward the strong interpretation of SMT.* Theoretical Linguistics at Keio-EMU Linguistics as Scientific Inquiry Lecture Series 3.

Gärtner, H.-M. (2023). Merge vs. "Lerge:" Problems of association. *Biolinguistics, 17*, Article e11715. https://doi.org/10.5964/bioling.11715

Roberts, I. G., & Watumull, J. (2015). Leibnizian linguistics. In Á. J. Gallego & D. Ott (Eds.), *50 years later: Reflections on Chomsky's aspects* (pp. 211–222). MIT Press.

Seely, T. D., Epstein, S. D., & Kitahara, H. (2022). *A minimalist theory of simplest merge.* Routledge.

Swoyer, C. (1994). Leibniz's calculus of real addition. *Studia Leibnitiana, 26*, 1–30.

Watumull, J., & Chomsky, N. (Forthcoming). *Life, information, language, intelligence.*