Á. Gallego & D. Ott (eds.). 2015. *50 Years Later: Reflections on Chomsky's* Aspects. Cambridge, MA: MITWPL.
Available for purchase at http://mitwpl.mit.edu
© The Authors

LEIBNIZIAN LINGUISTICS

IAN ROBERTS
University of Cambridge
JEFFREY WATUMULL
Oceanit Laboratories

"Nothing is more important than to see the sources of invention which are, in my opinion, more interesting than the inventions themselves." (Leibniz, in Pólya 1945: 123)

1 Introduction

The formulation of generative grammar in *Aspects of the Theory of Syntax* (1965) was foundational to the Second Cognitive Revolution—the revival of rationalist philosophy first expounded in the Enlightenment. (That the First Cognitive Revolution of the 17th Century contained "the sources of invention" for generative grammar was appreciated only in retrospect.) Chomsky discussed this historical connection at some length in *Aspects* (see especially Chomsky's *Cartesian Linguistics* (1966/2009)), identifying Leibniz as one of his precursors in adopting the rationalist doctrine that "innate ideas and principles of various kinds [...] determine the form of the acquired knowledge in what may be a rather restricted and highly organized way" (Chomsky 1965: 48). There are in fact many aspects of Leibniz's prodigious thought which anticipate core aspects of generative grammar, including some rather specific features of current theory. Below we sketch out what these are, and suggest that we should take Leibniz's lead in situating the ideas about human nature that generative grammar as a theory of language embodies, and which Chomsky has expounded many times, in a wider philosophical context.

First a very brief introduction to Leibniz's achievements. It is completely impossible to do justice to the breadth and depth of Leibniz's work in a short article like this (or even in a long book; Russell 1900/1992). Suffice to say that he was one of the greatest polymaths of all time—a "universalgenie" (a universal genius). Among his achievements are that he discovered calculus (independently of Newton) and other important aspects of mathematics, invented mechanical calculators, propounded a system of rationalist philosophy, formulated his own metaphysics ("the best of all possible worlds"), and made significant innovations in physics, probability theory, biology, medicine, geology, psychology, and what would centuries later become computer science. He also made a major contribution to the nascent subject of comparative linguistics in that he was one of the first to argue that the Finno-Ugric and Samoyedic families may be connected, thus taking "an early step towards the recognition of Uralic" (Campbell and Poser 2008: 91; this was almost a century before comparative Indo-European linguistics started). His achievements in logic were extraordinary; Kneale and Kneale (1962: 320) say that he "deserves to be ranked among the greatest of all logicians." In *The History of Western Philosophy*, Russell observes that Leibniz anticipated much of modern formal logic, as developed in the late 19th and early 20th centuries by Frege and others (including Russell himself). Indeed Russell (1946: 541) states that "Leibniz was a firm believer in the importance of logic, not only in its own sphere, but as the basis of metaphysics. He did work on mathematical logic which would have been enormously important if he had published it; he would, in that case, have been the founder of mathematical logic, which would have become known a century and a half sooner than it did in fact."[1] (Indeed it was not until the 1950s that mathematical logic enlightened linguistics: "Mathematical logic, in particular recursive function theory and metamathematics, were becoming more generally accessible [in the 1950s], and developments in these areas seemed to provide tools for a more precise study of natural languages as well. All of this I personally found most stimulating" (Chomsky 1955/1975: 39).[2])

Although this sketch is extremely perfunctory, it conveys something of Leibniz's staggering achievements. But our aim here is to focus on just a few aspects of this huge legacy, namely: Leibniz's formal innovations, his metaphysics, and his rationalist epistemology. The first and last of these show direct connections to modern linguistics (as Chomsky pointed out in relation to rationalist epistemology in *Aspects*, as we have already mentioned):

"Leibniz took Hobbes seriously when [Hobbes] said that reason is nothing but reckoning [computation]. [Leibniz] devoted much of his life to inventing a scheme that would perfect the computations underlying thought, turning arguments into calculations and making fallacies as obvious as errors in arithmetic[...]. The idea that intelligence arises from manipulation of symbols by rules is a major doctrine of the school of thought called rationalism, generally associated with Leibniz and Descartes. When the symbols stand for words and the rules arrange them into phrases and sentences, we have grammar, the subject of Cartesian linguistics, which later inspired Humboldt and then Chomsky. When the symbols stand for concepts and the rules string them into chains of inference, we have logic, which became the basis for digital computers, the artificial intelligence systems that run on them, and many models of human cognition" (Pinker 1999: 88-89).

Finally, Leibnizian metaphysics resonates with certain ideas in modern physics, and we will speculate on a connection to cognitive science in a way that, we believe, is very much in Leibniz's spirit.

---

[1] It is telling that Leibniz was the subject of the only book the prolific Russell ever wrote about another philosopher.

[2] Leibnizian philosophy was an obsession of one the greatest mathematical logicians, Kurt Gödel (see Goldstein 2005). Thus it is unsurprising that when Gödel and Chomsky met, generative grammar was overshadowed by Leibniz's ghost: "In 1958-9, I spent a year at the IAS. I naturally paid a visit to Gödel. He must have had every book on Leibniz in the Princeton library in his study. As soon as we started talking he instructed me that I was on the wrong track and that I should immerse myself in Leibniz's work, because that's where all the answers were about language and meaning" (Noam Chomsky, personal communication).

2 Leibniz's Formal Innovations

The eminent mathematician and computer scientist Martin Davis (2012) describes the history of computer science as "the road from Leibniz to Turing." Leibniz devised a *calculus ratiocinator* (an abstract computer) operating on a *characteristica universalis* (a symbolic language). This system is recognizably a universal computational system in the modern (Turing) sense: it is a system which manipulates symbols in a precisely defined (i.e., *computable*), *recursive* fashion: "A good century and a half ahead of his time, Leibniz proposed an algebra of logic, an algebra that would specify the rules for manipulating logical concepts in the manner that ordinary algebra specifies the rules for manipulating numbers. He 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); and Merge itself is but a generalization and simplification of the phrase structure rules assumed in *Aspects*.[3] Leibniz defined some of the properties of ⊕—call it *Lerge*—thus:

(1) $X \oplus Y$ is equivalent to $Y \oplus X$.

(2) $X \oplus Y = Z$ 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 $A \oplus B$." 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 indiscernible, they are identical: Merge *is* Lerge.[4] These functions embody the modern inductive/combinatorial conception of sets: "naturally prior involves the more simple[...]. Prior by nature is a term which consists of terms less derived. A term less derived is equivalent to one [which includes] a smallest number of primitive simple terms" (Leibniz, in Nachtomy 2008: 78). Here the intuition is clearly close to *definition by induction*, a central aspect of recursive computational systems such as generative grammar (see Watumull, et al. 2014).

In addition to definition by induction, recursive Merge/Lerge realizes the *mathematical induction* from finite to infinite (e.g., drawing from a finite lexicon, Merge generates an infinite

[3] "[Leibniz] grasped the concept of having formal, symbolic, representations for a wide range of different kinds of things. And he suspected that there might be universal elements (maybe even just 0 and 1) from which these representations could be built. And he understood that from a formal symbolic representation of knowledge, it should be possible to compute its consequences in mechanical ways—and perhaps create new knowledge by an enumeration of possibilities" (Wolfram 2013: http://blog.stephenwolfram.com/2013/05/dropping-in-on-gottfried-leibniz/). The connections to Chomsky's ideas are self-evident.

[4] Indeed, both Merge and Lerge are purely mechanical (i.e., explicit, generative, deterministic): "all these operations are so easy that there would never be any need to guess or try out anything" (Leibniz), just as a generative grammar "does not rely on the intelligence of the understanding reader" (Chomsky 1965: 4). And the operations Leibniz posited allow us to "proceed to infinity by rule" (Leibniz), just as a generative grammar computes a discrete infinity of syntactic structures; see http://blog.stephenwolfram.com/2013/05/dropping-in-on-gottfried-leibniz/.

set of syntactic objects; see footnote 4). From simple primitives, Nature recursively generates infinite complexity. A complex is generated by iterative "reflection" on simples; "God" reflects on his thinking. "Leibniz [...] presupposes logical simples[...], which he identifies with God's simple attributes (or God's simple forms). At the same time, for Leibniz, God is an active mind whose primary activity is thinking and self-reflection. God's reflections on his simple attributes are mental combinations of his simple forms that produce complex forms. Likewise, God's reflective operations are iterative, so that he reflects upon his reflections. Thus God thinks the combinations among his simple forms, and more complex concepts arise in his mind. In this view, God combines the simple forms in a natural order—from the simple to the complex—and, in this sense, Leibniz's system of possibility is both recursive and yields the production of infinite concepts" (Nachtomy 2007: 42-43). Evidently God is equipped with Merge/Lerge. So too are we. Leibniz observes: "The following operation of the mind seems to me to be most wonderful: namely, when I think that I am thinking, and in the middle of my thinking I note that I am thinking about my thinking, and a little later wonder at this tripling of reflection. Next I also notice that I am wondering and in some way I wonder at this wonder" (Leibniz, in Nachtomy 2008: 77). This recursive operation of reflection instances the more general notion of "Reducing reasoning to formal rules", which "was the underlying basis for Leibniz's dream of a universal computational language. And it underlay Turing's achievement in showing that all computation could be carried out on his universal machines" (Davis 2012: 177). And of course, in a decidedly rationalist move, Chomsky has demonstrated that linguistic cognition reduces to computation: "there is no difficulty in principle in programming a computer with [...] a universal grammar of the sort that has been proposed in recent years[...]. I believe these proposals can be properly regarded as a further development of classical rationalist doctrine, as an elaboration of some of its main ideas regarding language and mind" (Chomsky 1968/2006: 73-74). Here we see a profound connection between Leibniz's thought and modern linguistic theory. But the recursivity of Lerge is but one instance of recursion in Leibnizian philosophy.

Leibniz demonstrated the quadrature of the circle by generating an infinite series from a recursive function: $\pi/4 = 1 - 1/3 + 1/5 - 1/7 + 1/9 - 1/11 + 1/13 - ....$ Leibniz rejected as too weak the Aristotelean (constructivist) notion of *potential* infinities and accepted the Platonic existence of *actual* infinities. We can therefore assume that he would not have balked at the formulation of discrete infinity in generative grammar: "A grammar of [a language *L*] can be regarded as a function whose range is exactly *L*[;] We can consider a grammar of *L* to be a function mapping the integers onto *L*" (Chomsky 1959: 137, 138). And of course the integers form an actual infinity in Leibniz's sense—the Platonic sense. More generally, as Chomsky put it in *Aspects*: "The syntactic component specifies an infinite set of abstract formal objects, each of which incorporates all information relevant to a single interpretation of a particular sentence" (Chomsky 1965: 16). Such unboundedness would have enthralled Leibniz: "I am so in favor of the actual infinite that instead of admitting that Nature abhors it, as is commonly said, I hold that Nature makes frequent use of it everywhere, in order to show more effectively the perfections of its Author" (Leibniz, in Davis 2012: 51). And that Author generates infinite complexity from finite simplicity: "the simplicity of the means counterbalances the richness of the effects" so that in nature "the maximum effect [is] produced by the simplest means" (Leibniz, in Chaitin 2005: 63). This is an informal statement of a fundamental theorem in game theory, finally formalized in the 20th Century.

Indeed Leibniz anticipated important aspects of game theory, particularly von Neumann's *minimax theorem*. Game theory is the mathematical theory explicative of goals, logics, strategies, and related concepts, for which "the fundamental theorem on the existence of good strategies" (von Neumann 1953: 125) is the minimax theorem: "minimize maximum negatives and maximize minimum positives." This is a good strategy—a principle of optimality—independent of the domain of application because, as Leibniz observed, "the simplicity of the means counterbalances the richness of the effects" so that in nature "the maximum effect [is] produced by the simplest means." And "given that things exist," and given the assumption that "nothing comes about without a sufficient reason[...], we must be able to give a reason why [things] have to exist as they are and not otherwise" (Leibniz, in Chatin 2005: 63); this is the powerful *Principle of Sufficient Reason*. Observing this principle in the domain of recursive function (computability) theory as applied to generative grammar, the optimal Merge function—and here we are interested in the inductive generation of sets (i.e., the recursive generation of structured expressions)—is minimax (see Watumull and Roberts, Under Review): an effectively unary binary function. Unarity implements minimality (e.g., minimal search, noncounting, etc.): one input suffices to generate an output; but such output is set theoretically trivial (i.e., the potentially infinite subsumption of a single substantive element: α, {α}, {{α}}, ..., {...{{α}}...}...}—an embedded lexical item). Binarity implements maximality: the capacity to combine two inputs recursively suffices to generate infinitely structured (information carrying) outputs (i.e., {α, β}, {γ, {α, β}}, {δ*i*, {γ, {α, β}}}, ..., {δ*n*, ..., {δ*i*, {γ, {α, β}}}...}—a complex sentence or discourse).

The reason to minimize is self-evident (i.e., anything superfluous violates the principle of sufficient reason); the reason to maximize, Leibniz argues, is that it "leave[s] nothing [...] which does not display that beauty of which it is capable" such that a mechanism with the potential to generate infinite complexity ought to realize that potential. Failure to maximize results in a reserve—an unnecessary and hence inexplicable resource—inconsistent with the principle of sufficient reason and perfect design. "When the simplicity of God's way is spoken of, reference is specially made to the means which he employs, and on the other hand when the variety, richness and abundance are referred to, the ends or effects are had in mind. Thus one ought to be proportioned to the other, just as the cost of a building should balance the beauty and grandeur which is expected" (Leibniz, in Rescher 1991: 196). Nature has "paid" for the unbounded generative capacity of Lerge/Merge, and for Platonic balance thus "expects" and "deserves" unbounded generation. Crudely put, getting "the biggest bang for the buck" is a principle of nature.

Thus in the Leibnizian "best of all possible worlds," Merge would be propertied with the minimality of unarity and the maximality of binarity. Let us postulate such a minimax Merge as a formally explicit—hence comprehensible and testable—hypothesis. The Minimalist Program (Chomsky 1995) can be seen, in this light, as investigating the hypothesis that the computational system of language is that we would expect to find in a Leibnizian "best of all possible worlds".

Such a world is a reflection of perfection in that everything ought to be derivable from—explicable in terms of—the simple forms.[5] These are the primitives, the axioms, of nature.

[5] "The assertion that nature is governed by strict laws is devoid of all content if we do not add the statement that it is governed by mathematically simple laws[...]. That the notion of law becomes empty when an arbitrary complication is permitted was already pointed out by Leibniz in his *Metaphysical Treatise*[...]. The astonishing thing is not that there exist natural laws, but that the further the analysis proceeds[...], the finer the elements to which the phenomena are reduced, the simpler—and not the more complicated, as one would originally expect—the fundamental relations become and the more exactly do they describe the actual occurrences" (Weyl 1932: 40, 41, 42). Leibniz and Weyl's comments apply perfectly to the Minimalist Program: "Some basic properties of language are unusual among biological systems, notably the property of discrete infinity. A working hypothesis in generative grammar has been that languages are based on simple principles that interact to form often intricate structures, and that the language faculty is nonredundant, in that particular phenomena are not 'overdetermined' by principles of language. These too

Nothing complex could exist but for irreducible simples. "For this reason, complex thoughts or concepts presuppose constituents. Leibniz indeed presupposes absolutely simple constituents or forms. He writes that 'nothing can be said of forms on account of their simplicity' [...]" (Leibniz, in Nachtomy 2008: 76).

The notion that complex thoughts presuppose constituents is axiomatic to Leibniz's universal computational language. Moreover, the notion has been adduced to argue that not only are natural languages universally constituency-based, but that *any language anywhere in the universe* would be: we should predict convergence on elementary recursive functions operating on simple constituents; and thus we should predict that "intelligent aliens will be intelligible" (Minsky 1984). However, at the time of *Aspects*, Chomsky would probably have rejected this hypothesis: "according to the theory of transformational grammar, only certain kinds of formal operations on strings can appear in grammars—operations that, furthermore, have no a priori justification. For example, the permitted operations cannot be shown in any sense to be the most 'simple' or 'elementary' ones that might be invented" (Chomsky 1965: 55). Leibniz would not have accepted this theory, for in his philosophy nothing is without a priori justification (i.e., everything has a sufficient reason) and such justifications are ultimately based on a metaphysics in which the most complex objects are generated by the most elementary combinatorics (i.e., $\oplus$) operating over the most simple representations (*monads*, on which more below).

Interestingly, since *Aspects*, generative grammar has become increasingly Leibnizian. Already in *Lectures on Government and Binding* (1981), there were hints that language may be "a system that goes well beyond empirical generalization and that satisfies intellectual or even esthetic standards" (Chomsky 1981: 14). And now, in the Minimalist Program, we can seriously entertain the thesis that the language faculty conforms to "virtual conceptual necessity, so that the computational system [is] in some sense optimal" (Chomsky 1995: 9). If this Leibnizian progression continues, future research will prove language to be governed by *absolute* conceptual necessity—truly the best of all possible worlds.[6]

At this point, we should turn to Leibniz's metaphysics.

3 Metaphysics

It is ironic that the codiscoverer of the calculus should have based his metaphysics not on its *continuum*—a "labyrinth" to Leibniz's mind—but rather on *discrete* substances: *monads*. For Leibniz, reality is composed of these irreducible immaterial objects—their existence an inescapable logical truth: everything complex is ipso facto composed of simpler things; if those simpler things are themselves extended (i.e., material), then ipso facto they are further reducible; thus it follows that at bottom the simplest of simple substances must be un-extended (i.e., immaterial); "these simple forms are unanalyzable and indefinable [...]. Interestingly, Leibniz identifies these simple forms with the attributes of God. He writes that 'God is the subject of all

are unexpected features of complex biological systems, more like what one expects to find (for unexplained reasons) in the study of the inorganic world. The approach has, nevertheless, proven to be a successful one, suggesting that the hypotheses are more than just an artifact reflecting a mode of inquiry" (Chomsky 1995: 168).

[6] Just as Chomsky's simplification of phrase structure rules in *Aspects* revealed greater systematicity, so did Leibniz reveal elegant structure in mathematics by formulating binary arithmetic: "as numbers are reduced to the simplest principles, like 0 and 1, a wonderful order is apparent throughout" (Leibniz); see http://blog.stephenwolfram.com/2013/05/dropping-in-on-gottfried-leibniz/.

absolute simple forms' [...] and that '[a]n attribute of God is any simple form'" (Nachtomy 2008: 76). Substitute "nature" for "God," for a naturalistic philosophy (i.e., scientific—indeed rational—legitimacy). What are these simple forms? What simply *is*? What is in need of no explanation? *The truths of mathematics*. Like Leibniz, "Descartes had raised the question of whether God had created the truth of mathematics. His follower Nicolas Malebranche [...] firmly expressed the view that they needed no inception, being as eternal as anything could be" (Dennett 1995: 184). For such reasons, mathematics was foundational to Leibnizian philosophy (and generative grammar).[7] Indeed monads are inherently mathematical (even Platonic).

Monadology implies recursivity: "a theory that explains the status of physical phenomena (together with their relations) in terms of collections of monads mutually representing the objective contents of one another's representations in a fashion that is infinitely recursive" (McGuire 1985: 215). The simplest forms are binary. Here again we see a link to modern linguistic theory (Merge) and to cognitive science (the computational theory of mind). But Leibniz's metaphysics, the recursive monadology, led him further: "Leibniz was very proud to note how easy it is to perform calculations with binary numbers, in line with what you might expect if you have reached the logical bedrock of reality. Of course, that observation was also made by computer engineers three centuries later" (Chaitin 2005: 61). Binary enables minimax: "It is the mystic elegance of the binary system that made Leibniz exclaim *Omnibus ex nihil ducendis sufficit unum.* (One suffices to derive all out of nothing.)" (Dantzig 1930/2007: 15). To derive everything from (virtually) nothing allows for a complete explanation.

Laplace commented as follows on how Leibniz linked his binary arithmetic to metaphysics: "Leibniz saw in his binary arithmetic the image of Creation[...]. He imagined that Unity represented God, and Zero the void; that the Supreme Being drew all beings from the void, just as unity and zero express all numbers in his system of numeration" (Laplace, in Dantzig 1930/2007: 15). Here there is a connection with some ideas in contemporary physics: "Leibniz's vision of creating the world from 0's and 1's refuses to go away. In fact, it has begun to inspire some contemporary physicists, who probably have never even heard of Leibniz[...]. What is the Universe made of? A growing number of scientists suspect that information plays a fundamental role in answering this question. Some even go so far as to suggest that information-based concepts may eventually fuse with or replace traditional notions such as particles, fields and forces. The Universe may literally be made of information, they say, an idea neatly encapsulated in the physicist John Wheeler's slogan: 'It from Bit' [Matter from Information]. So perhaps Leibniz was right after all" (Chaitin 2005: 61, 62). If information derives matter, if the mathematics of 0's and 1's constitutes reality, if that which requires no explanation—that which is self-explanatory—derives all that does require explanation, then a *theory of everything* is possible. As Leibniz observes: "given that things exist," and given the assumption that "nothing comes about without a sufficient reason[...], we must be able to give a reason why [things] have to exist as they are and not otherwise." Thus Leibniz anticipated Chomsky's (2004) enterprise to go "beyond explanatory adequacy." Stated generally, to go beyond explanatory adequacy is to give a sufficient reason. Here we speculate, but for generative grammar in particular, the reason

[7] "[Leibniz's] typical approach seems to have been to start by trying to create a formal structure to clarify things— with formal notation if possible. And after that his goal was to create some kind of 'calculus' from which conclusions could systematically be drawn" (http://blog.stephenwolfram.com/2013/05/dropping-in-on-gottfried-leibniz/). This relates to that idea that "The most interesting contribution a generative grammar can make to the search for universals of language is specify formal systems that have putative universals as *consequences* as opposed to merely providing a technical vocabulary in terms of which autonomously stipulated universals can be expressed" (Gazdar, et al. 1985: 2).

the language faculty assumes the form it does could be that *it too has come from bit*: the language faculty *is* a mathematical object (see Watumull (Under Review) for more on this Leibnizian—indeed Platonist or even Pythagorean—idea). This particular mathematical object (computational system) is obviously encoded neurobiologically, which leads us to rationalism.

4 Rationalism

In *Cartesian Linguistics*, Chomsky (1966/2009) writes that in the early modern era, rationalism "focus[ed] attention on the innate interpretative principles that are a precondition for experience and knowledge and [emphasized] that these are implicit and may require external stimulation in order to become active[...]. The psychology that develops in this way is a kind of Platonism without preexistence. Leibniz makes this explicit in many places. Thus he holds that 'nothing can be taught us of which we have not already in our minds the idea,' and he recalls Plato's 'experiment' with the slave boy in the *Meno* as proving that 'the soul virtually knows those things [i.e., truths of geometry, in this case], and needs only to be reminded (animadverted) to recognize the truths. Consequently, it possesses at least the idea upon which these truths depend. We may say even that it already possesses those truths, if we consider them as the relations of the ideas'" (Chomsky 1966/2009: 62-63).[8] This rationalist current flowed into romanticism, as Chomsky observed in *Aspects*: "Like Leibniz, [Humboldt] reiterates the Platonistic view that, for the individual, learning is largely a matter of Wiedererzeugung, that is, of drawing out what is innate in the mind" (Chomsky 1965: 51). "In the traditional view a condition for [...] innate mechanisms to become activated is that appropriate stimulation must be presented[...]. For Leibniz, what is innate is certain principles (in general unconscious), that 'enter into our thoughts, of which they form the soul and the connection'. 'Ideas and truths are for us innate as inclinations, dispositions, habits, or natural potentialities.' Experience serves to elicit, not to form, these innate structures[...]. It seems to me that the conclusion regarding the nature of language acquisition [as reached in generative grammar] are fully in accord with the doctrine of innate ideas, so understood, and can be regarded as providing a kind of substantiation and further development of this doctrine" (Chomsky 1967: 10).

Chomsky quotes Leibniz at some length in *Aspects* for the proposition explicitly formulated in generative grammar that we are innately endowed with a competence to generate unbounded knowledge in specific cognitive domains. This knowledge develops if we are presented with the appropriate stimulation. Here Leibniz discusses our related faculty for mathematics: "The truths of numbers are in us, yet nonetheless one learns them [...] by drawing them from their source when we learn them through demonstrative proof (which shows that they are innate),"[9] and thus "all arithmetic and geometry are in us virtually, so that we can find them there if we consider attentively and set in order what we already have in the mind[...]. [In general,] we have an

[8] This "Platonism without preexistence" was carried over into modern biology: "Plato says [...] that our 'necessary ideas' arise from the preexistence of the soul, are not derivable from experience—read monkeys for preexistence" (Darwin, in Desmond and Moore 1991: 263). Indeed evolutionary psychology—and evolutionary linguistics—does look to the minds of related species for evidence of homologues/analgoues to our psychological (e.g., linguistic) capacities.

[9] Leibniz is rehearsing Plato's (Socrates') idea: "Socrates employs his maieutic methods [...]. Maieutics is the pedagogical method that tries to draw a conclusion out of a mind where it latently resides. As Leibniz remarks of the method, it consists in drawing valid inferences" (Goldstein 2014: 304). It is interesting to consider the similarities and differences when what is being drawn out of the mind is language.

infinite amount of knowledge of which we are not always conscious, not even when we need it[...]. The senses, although necessary for all our actual knowledge, are not sufficient to give it all to us, since the senses never give us anything but examples, i.e., particular or individual truths. Now all the examples which confirm a general truth, whatever their number, do not suffice to establish the universal necessity of the same truth[...]. Necessary truths [...] must have principles whose proof does not depend on examples, nor consequently upon the testimony of the senses'" (Chomsky 1965: 50).

The rationalist theory of language acquisition as running a form of minimum description length algorithm was implicit in Leibniz's discussion of laws of nature. In connection with data analysis (a central concept in any theory of language acquisition and learnability), Leibniz runs a Gedankenexperiment in which points are randomly distributed on a sheet of a paper and observes that for any such random distribution it would be possible to draw some "geometrical line whose concept shall be uniform and constant, that is, in accordance with a certain formula, and which line at the same time shall pass through all of those points" (Leibniz, in Chaitin 2005: 63), as with a Lagrangian interpolation. For any set of data, a general law can be constructed. "How can we decide if the universe is capricious or if science actually works? And here is Leibniz's answer: If the law has to be extremely complicated ('*fort composée*') then the points are placed at random, they're '*irrégulier*,' not in accordance with a scientific law. But if the law is simple, then it's a genuine law of nature" (Chaitin 2005: 63). In Leibniz's words, the true theory is "the one which at the same time [is] the simplest in hypotheses and the richest in phenomena, as might be the case with a geometric line, whose construction was easy, but whose properties and effects were extremely remarkable and of great significance" (Leibniz, in Chaitin 2005: 63). This Leibnizian logic is the logic of program size complexity or, equivalently, minimum description length, as determined by some evaluation/simplicity metric (see Chomsky and Halle 1968: Chapter 9). "Leibniz had all the pieces, he had only to put them together. For he worshipped 0 and 1, and appreciated the importance of calculating machines. [Leibniz could have anticipated Chaitin by understanding] a scientific theory as a binary computer program for calculating the observations, written down in binary. [There exists] a law of nature if there is compression, if the experimental data is compressed into a computer program that has a smaller number of bits than are in the data that it explains. The greater the degree of compression, the better the law" (Chaitin 2005: 64). This logic applies straightforwardly to language acquisition.

Language acquisition can be represented as navigating a hierarchy of parameters computing a series of decision points (see Roberts 2012, Biberauer and Roberts 2013 and references given there). In these hierarchies, Leibniz's notion of "the simplest in hypotheses and the richest in phenomena" is represented in the minimax principle of feature-economy (minimize formal features) combined with input generalization (maximize available features). The parameterized language—the steady state—is attained by traversing a path in the hierarchies: the parameterized language is by definition the shortest possible (if described in bits), consistent with Leibnizian program size complexity. As Berwick (1982: 6, 7, 8) puts it, a "general, formal model for the complexity analysis of competing acquisition [...] demands [can be] based on the notion of program size complexity[—]the amount of information required to 'fix' a grammar on the basis of external evidence is identified with the size of the shortest program needed to 'write down' a grammar. [Formally, this] model identifies a notational system with some partial recursive function Φi (a Turing machine program) and a rule system as a program p for generating an observed surface set of [primary linguistic data]. Like any computer program, a program for a rule system will have a definite control flow, corresponding roughly to an augmented flowchart that describes the implicational structure of the program. The flow diagram specifies [...] a series of 'decision points' that actually carry out the job of building the rule system to output. [The] implicational structure in a developmental model corresponds rather directly to the existence of implicational clusters in the theory of grammar, regularities that admit short descriptions. [T]his same property holds more generally, in that all linguistic generalizations can be interpreted as implying specific developmental 'programs'". Thus the complexities of language acquisition and linguistic generalizations are determined by simple programs—a Leibnizian conclusion evident in the theory of parameter hierarchies. More generally, "Leibniz, like Hobbes (who had influenced him), was ahead of his time in recognizing that intelligence is a form of information processing that needs complex machinery to carry it out. As we now know, computers don't understand speech or recognize text as they roll off the assembly line; someone has to install the right software first" (Pinker 2002: 35).

In the rationalist view of language acquisition, a generative grammar is genetically preprogrammed and activated by primary linguistic data. In *Aspects*, Chomsky "follow[s] Leibniz's enlightening analogy, [with] 'the comparison of a block of marble which has veins, rather than a block of marble wholly even, or of blank tablets, i.e., of what is called among philosophers a tabula rasa. For if the soul resembled these blank tablets, truths would be in us as the figure of Hercules is in the marble, when the marble is wholly indifferent to the reception of this figure or some other. But if there were veins in the block which should indicate the figure of Hercules rather than other figures, this block would be more determined thereto, and Hercules would be in it as in some sense innate, although it would be needful to labor to discover these veins, to clear them by polishing, and by cutting away what prevents them from appearing. Thus it is that ideas and truths are for us innate, as inclinations, dispositions, habits, or natural potentialities, and not as actions; although these potentialities are always accompanied by some actions, often insensible, which correspond to them'" (Chomsky 1965: 52). So, like a figure revealed in marble (the Leibnizian analogy), or a geometrical theorem elicited from an untutored mind (the Socratic example), the form of linguistic knowledge—a generative grammar—is predetermined, but its expression is contingent. (Hence did Leibniz distinguish *necessary necessity* from *contingent necessity*.)

5 Conclusion

Contra Leibniz, it is conceivable that some things exist for no reason—things for which there is no explanation—but to conclude so is in general logically impossible (modulo some cases in mathematics): it cannot be proved that an "unsuccessful" search of the infinite space of possible explanations has been exhaustive; it cannot be proved that our search will never succeed (i.e., converge on an explanation); this is analogous to the incomputability of the halting problem and Kolmogorov- Chaitin complexity. But we should try, as in Hilbert's dictum: *Wir müssen wissen. Wir werden wissen*. This, it seems to us, is an apt description of the program for linguistic theory set out in full for the first time in *Aspects*, a program which has set linguistics firmly in a venerable rationalist position, both epistemologically and metaphysically. We seek theories that are not merely observationally and descriptively adequate. Nor are we satisfied with explanatory adequacy. The desideratum, implicit in *Aspects* and explicit in the Minimalist Program, is to go beyond explanatory adequacy and give a sufficient reason—a reason that would satisfy Leibniz—for why language assumes its unique form—perhaps the best of all possible forms.

References

Berwick, R.C. 1982. Locality principles and the acquisition of syntactic knowledge. Cambridge, MA: MIT dissertation.

Biberauer, T. and I. Roberts. 2013. The clausal hierarchy, features and parameters. In U. Shlonsky (ed.), *Cartography: The Way Forward*. Oxford: Oxford University Press.

Campbell, L. and W.J. Poser. 2008. *Language Classification: History and Method*. Cambridge: Cambridge University Press.

Chaitin, G. 2005. *Meta Math: The Quest for Omega*. New York: Pantheon Books.

Chomsky, N. 1955/1975. *The Logical Structure of Linguistic Theory*. New York: Plenum Press.

Chomsky, N. 1959. On certain formal properties of grammars. *Information and Control* 2: 137-167.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.

Chomsky, N. 1966/2009. *Cartesian Linguistics*. Cambridge: Cambridge University Press.

Chomsky, N. 1967. Recent contributions to the theory of innate ideas. *Synthese* 17: 2-11.

Chomsky, N. 1968/2006. *Language and Mind*. Cambridge: Cambridge University Press.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2004. Beyond explanatory adequacy. In A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures*, vol. 3, 104-131. Oxford: Oxford University Press.

Chomsky, N. and M. Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.

Dantzig, T. 1930/2007. *Number: The Language of Science*. New York: Plume.

Davis, M. 2012. *The Universal Computer: The Road from Leibniz to Turing*. New York: Taylor & Francis Group.

Dennett, D.C. 1995. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. London: Penguin Books.

Desmond, A. and J. Moore. 1991. *Darwin*. London: Michael Joseph.

Gazdar, G., E. Klein, G.K. Pullum, and I.A. Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge: Harvard University Press.

Goldstein, R.N. 2005. *Incompleteness: The Proof and Paradox of Kurt Gödel*. New York: W.W. Norton & Company.

Goldstein, R.N. 2014. *Plato at the Googleplex: Why Philosophy Won't Go Away*. New York: Pantheon Books.

Kneale, W. and M. Kneale. 1962. *The Development of Logic*. Oxford: Oxford University Press.

McGuire, J.E. 1985. Phenomenalism, relations, and monadic representation: Leibniz on predicate levels. In J. Bogen and J.E. McGuire (eds.), *How Things Are*, 205-233. Dordrecht: Reidel.

Minsky, M. 1984. Why intelligent aliens will be intelligible. In E. Regis (ed.), *Extraterrestrials: Science and Alien Intelligence*, 117-128. Cambridge, MA: MIT Press.

Nachtomy, O. 2007. *Possibility, Agency, and Individuality in Leibniz's Metaphysics*. Dordrecht: Springer.

Nachtomy, O. 2008. Leibniz's rationality: Divine intelligibility and human intelligibility. In M. Dascal (ed.), *Leibniz: What Kind of Rationalist?*, 73-82. Dordrecht: Springer.

Pinker, S. 1999. *Words and Rules: The Ingredients of Language*. NewYork: Perennial.

Pinker, S. 2002. *The Blank Slate: The Modern Denial of Human Nature*. New York: Penguin Books.

Pólya, G. 1945. *How to Solve It*. Princeton: Princeton University Press.

Rescher, N. 1991. *G.W. Leibniz's Monadology*. New York: Routledge.
Roberts, I. 2012. Macroparameters and minimalism: A programme for comparative research. In C. Galves, S. Cyrino, R. Lopes, F. Sândalo and J. Avelar (eds.), *Parameter Theory and Linguistic Change*, 320-335. Oxford: Oxford University Press.
Russell, B. 1900/1992. *A Critical Exposition of the Philosophy of Leibniz*. New York: Routledge.
Russell, B. 1946/2004. *The History of Western Philosophy*. London: Routledge.
von Neumann, J. 1953. Communication on the Borel notes. *Econometrica* 21: 124-125.
Watumull, J. Under Review. Language and the Computable Universe Hypothesis.
Watumull, J., M.D. Hauser, I.G. Roberts, and N. Hornstein. 2014. On recursion. *Frontiers in Psychology* 4: 1-7.
Watumull, J. and I. Roberts. Under Review. Syntax all the way up.
Weyl, H. 1932. *The Open World: Three Lectures on the Metaphysical Implications of Science*. New Haven: Yale University Press.