A ‘Galilean’ science of language

CHRISTINA BEHME

Dalhousie University

(Received 30 January 2013; revised 20 January 2014)


*The Science of Language*, published in the sixth decade of Noam Chomsky’s linguistic career, defends views that are visibly out of touch with recent research in formal linguistics, developmental child psychology, computational modeling of language acquisition, and language evolution. I argue that the poor quality of this volume is representative of the serious shortcomings of Chomsky’s recent scholarship, especially of his criticism of and contribution to debates about language evolution. Chomsky creates the impression that he is quoting titbits of a massive body of scientific work he has conducted or is intimately familiar with. Yet his speculations reveal a lack of even basic understanding of biology, and an unwillingness to engage seriously with the relevant literature. At the same time, he ridicules the work of virtually all other theorists, without spelling out the views he disagrees with. A critical analysis of the ‘Galilean method’ demonstrates that Chomsky uses appeal to authority to insulate his own proposals against falsification by empirical counter-evidence. This form of discourse bears no serious relation to the way science proceeds.

I. Introduction

*The Science of Language: Interviews with James McGilvray*, henceforth *The Science of Language*, published in the sixth decade of Noam Chomsky’s linguistic career, should have been an impressive summary of the achievements of one of the greatest intellectuals of our time. It is not. Chomsky’s scholarship has

[1] I am greatly indebted to Avery Andrews, Michael Arbib, Derek Bickerton, Paul Bloom, Rudie Botha, Ted Briscoe, Morten Christiansen, Patricia Churchland, Michael Corballis, Peter Culicover, Stanley Dubinsky, Shimon Edelman, Jeff Elman, Dan Everett, Dan Flage, Susan Fred Schmerling, Jim Hurford, Ray Jackendoff, David Johnson, Dan Lassiter, Robert Levine, Philip Lieberman, Brian MacWhinney, Robert Martin, Frederick Newmeyer, David Papineau, Paul Postal, Michael Studdert-Kennedy, Geoffrey Sampson, Pieter Seuren, Maggie Tallerman, Michael Tomasello, and Virginia Valian for very helpful replies to my inquiries and for commenting on earlier drafts. Further gratitude is owed to the editor Bob Borsley and three anonymous referees of *Journal of Linguistics*. All remaining errors are mine.
arguably been slowly deteriorating over decades, and this volume is altogether representative of the problem. I shall argue in this review article that uncritical acceptance of Chomsky’s work despite its steadily declining quality has been doing the field considerable damage.

Chomsky is ‘the leading figure in contemporary linguistics’ (Harman 2001: 265). His technical innovations have set and re-set theoretical agendas for those who share his general perspective, and provided considerable input, if only in the form of novel observations, for those who work in different frameworks. He is regularly invited to deliver keynote or plenary talks at leading universities, and remains the public face of linguistics. Assuming that it is correct that ‘[m]ost educated people have never encountered linguistics, and have no idea what it might even mean to examine a linguistic puzzle scientifically’ (Pesetsky 2013: slide 120), it is to be expected that people who become interested in linguistics will turn their attention first to Chomsky’s publications, and judge the field by their quality. But the field should not be judged by publications like The Science of Language.

One goal of this review article is to engage in what Chomsky advocates as ‘consciousness raising’ (119) and to encourage readers to apply to Chomsky’s work the same standards that are applied throughout science. The Science of Language has been reviewed elsewhere (e.g. Bishop 2012, Pullum 2012, Behme 2013a, Lieberman in press), so readers seeking a more conventional review are already well served. This review article offers only a brief survey of the book as a whole, before focusing on certain key passages in order to discuss serious shortcomings of Chomsky’s recent scholarship. I shall show that his recent work fails to meet serious scientific standards because he rejects scientific procedure, inflates the value of his own work, and distorts the work of others, and that the poor quality of The Science of Language is no isolated ‘misstep’ but characteristic of many of his recent (and some not so recent) publications.²

2. Brief Overview of The Science of Language

The Science of Language contains twenty-five interviews. Part I introduces the reader to Chomsky’s thought on the design and function of human language, language evolution, representationalism, the nature of human concepts, optimality and perfection of Universal Grammar, and Chomsky’s intellectual contributions. Part II includes discussions of human nature, evolutionary psychology, morality, epistemology and biological limits on human understanding. In addition,
McGilvray provides twelve appendices, chapter-by-chapter commentaries, and a glossary. Given that the volume aims (at least in part) at ‘illuminating [Chomsky’s linguistics for] . . . newcomers’ (Paul Pietroski, back cover), the interview format and the frequent focus on philosophical and political issues might seem appropriate. However, as noted by another reviewer, ‘there seemed no coherent structure . . . The style is so discursive that it’s impossible to précis. [Chomsky’s] rhetorical approach seemed the antithesis of a scientific argument’ (Bishop 2012). Indeed, searching for the promised ‘cutting-edge theorizing’ (Robert Stainton, back cover) proves futile, and any hopes for philosophical engagement with long-standing criticism of the Chomskyan paradigm are in vain. Further, it is difficult (especially for newcomers) to follow the presentation because terms are not clearly defined, the conversation meanders through multitudes of obscure, irrelevant digressions, fallacies of reasoning are committed, and far-reaching conclusions are often drawn from meagre premisses. The only attempt in the entire volume at a syntactic analysis (for the sentence *Harry saw a house*) is a sketch outlined by McGilvray in one of the appendices (232).

*The Science of Language* has a variety of shortcomings, which can be summarized as follows. Chomsky dogmatically defends a view that is visibly out of touch with recent research in formal linguistics, developmental child-psychology, computational modeling of language acquisition, and language evolution. He fails to provide novel (or at times, any) arguments supporting his belief in a domain-specific innate biological endowment, saltational language evolution, semantic internalism, or computational optimality. Instead, he proclaims as irrefutable truth what is hotly contested outside the Chomskyan camp. Most of the arguments provided are sketchy repetitions of controversial arguments that Chomsky put forward years or even decades ago. There is virtually no reference to work by Chomskynans that had not been already discussed in recently published volumes, such as *Cartesian Linguistics* (Chomsky 1966a, 3rd edition in 2009), *Of Minds and Language* (Piattelli-Palmarini, Uriagereka & Salaburu 2009), *Noam Chomsky on Language and Cognition* (Özsoy & Nakipoğlu 2009), and *Chomsky Notebook* (Bricmont & Franck 2010). At times Chomsky’s replies to McGilvray’s questions are almost identical to passages from such earlier works. This turns the current volume into a meta-interview: pieces of earlier interviews are recycled and cobbled together in quilt-like fashion.

---

[3] Naturally, one would not expect novel research in a 2009 edition of *Cartesian Linguistics*. Yet it advertises a ‘new and specifically written introduction by James McGilvray, contextualizing the work for the twenty-first century’ (back cover), and contains a letter by Chomsky commenting on computational work.

[4] This volume was originally published as *Chomsky* in French, in 2007.

[5] None of these earlier works reported original research either. There is some discussion of biological and psychological work in *Of Minds and Language*. But discussion of novel linguistic findings has been absent from Chomsky’s publications for a long time.
Apart from the poor quality of argumentation, *The Science of Language* displays numerous violations of basic publication standards. For instance, there are several mismatches in commentary reference pagination, and work criticized by Chomsky is not cited properly in the text (e.g. ‘that guy’ (57); ‘a very good English philosopher’ (30)), and not listed in the references section (Lassiter 2008, Papineau 2008). While ultimately Chomsky is responsible for his own remarks, a share of criticism must also be attributed to the interviewer, James McGilvray, and to the editors at the distinguished academic publisher. Why did they not enforce higher standards?

The problems listed above would be serious flaws in any academic work. But *The Science of Language* is no isolated instance of poor scholarship and sub-standard editing. The public has been exposed in short succession to five books with Chomsky’s name on the cover, all of which fail to live up to enthusiastic back cover endorsements by prominent generativists (Paul Pietroski, Howard Lasnik, Robert Freidin, and Robert Stainton). Despite these failings, the books have elicited the highest praise in reviews published in reputable venues: ‘[Chomsky Notebook] is one of the most comprehensive, sensitive, and imaginative representations of Chomsky’s oeuvre . . . The volume opens with two majestic essays by Chomsky himself’ (Mukherji 2010, *Notre Dame Philosophical Reviews*); ‘[Of Minds and Language] is a book remarkably rich in ideas’ (Kljajevic 2009); ‘the papers [in *Of Minds and Language*] are uniformly thoughtful and provide an excellent guide to some of the best thinking on biolinguistic themes’ (Drummond & Hornstein 2011, *Biolinguistics*). Negative evaluations (Postal 2012, Pullum 2012, Behme 2013b) have elicited open hostility. All this suggests, especially to non-linguists, is that work which fails to meet even minimal standards for scientific publications is held in high esteem by professional linguists.

3. **MEAGER SUPPORT FOR CHOMSKY’S RECENT PROPOSALS**

Chomsky’s early work introduced innovative ideas, earned him recognition beyond linguistics, and inspired generations of linguists. His early proposals in particular have changed the way linguistic research is conducted. Yet anyone unfamiliar with the field who reads one of Chomsky’s early works, e.g. *Syntactic Structures* (1957) or *Aspects of the Theory of Syntax* (1965), back to back with *The Science of Language* cannot fail to be mystified by the difference. Of course, this dramatic change has not occurred overnight, and marked departures from the path laid out in the early works have been visible in Chomsky’s work for a long time. Despite of this, even linguists who do not work in Chomsky’s framework extend the praise that was justified for early publications to Chomsky’s recent work. A representative example is: ‘[Chomsky’s 1950s] aim was to investigate language at the rigorous level that physics is studied. As we will see over the following chapters, this aim is incontrovertibly discernible in his recent work’ (Kinsella 2009: 18).
This situation calls for a demonstration of just how non-rigorous Chomsky’s current ‘theorizing’ is.\textsuperscript{6} For this purpose consider Chomsky’s criticism of and contribution to debates about language evolution. Addressing the problem of language evolution is important for anyone defending the view that language is a biological organ. But Chomsky’s discussion of language evolution reveals the full extent of the double standards that are evident throughout \textit{The Science of Language} (and several other of his 21st century publications). Chomsky’s view is put forward dogmatically as the only rational option. He creates the impression that he is quoting titbits of a massive body of scientific work he has conducted or is intimately familiar with. At the same time, he ridicules the work of virtually all other theorists, without ever spelling out the views he disagrees with.

Chomsky’s language evolution speculations reveal a lack of even basic understanding of biology, and an unwillingness to engage seriously with the relevant literature. In three decades, Chomsky has rarely offered more than ‘off-the-cuff’ remarks about language evolution. Only one proposal sketched, in very broad strokes, a scientifically testable hypothesis (Hauser, Chomsky & Fitch \textsc{2002}).\textsuperscript{7}

In their paper, Hauser et al. (\textsc{2002}) argue that human language requires two components: the faculty of language in the broad sense (FLB) and the faculty of language in the narrow sense (FLN). The former includes a sensorimotor system that produces and receives linguistic signals and a conceptual-intentional system that allows us to categorize/organize information and to understand social cues. The latter is the abstract linguistic computational system alone, independent of the other systems with which it interacts and interfaces . . . The core property of FLN is recursion . . . it takes a finite set of elements and yields a potentially infinite array of discrete expressions’ (Hauser et al. \textsc{2002: 1571}). While details of this proposal have been criticized (Jackendo\textsuperscript{ff} & Pinker \textsc{2005}), it offered the potential for developing a scientific hypothesis. However, when threatening empirical counterevidence was reported (Everett \textsc{2005a, b}), Chomsky refused to follow the rules of science.

\textsuperscript{6} There is an ongoing futile debate about whether Chomsky’s published proposals are theories, hypotheses, programs, or speculations. Much ink has been spilled ridiculing critics who have called a ‘program’ a ‘theory’, etc. Based on such technicalities Chomsky claims that his critics do not know what they are talking about, and evades engagement with their substantial criticism (for discussion of Chomsky’s practices see Behme \textsc{2014}). I do not wish to become involved in this debate and will use scare quotes throughout to indicate that the term used by me refers to whichever term Chomsky deems appropriate.

\textsuperscript{7} Tomalin (\textsc{2011: 308}) argues that ‘the informal discussion in HCF offers no detailed definitions [of recursion] . . . [and that] hazy informal description of this kind are of very little use’. Important definitions given by Hauser et al. (\textsc{2002}) were vague and misleading. Particularly confusing was the ‘terminological distinction between FLN and FLB [that was] intended to help clarify misunderstandings’ (Fitch, Hauser & Chomsky \textsc{2005: 179}), where, as has become evident in recent publications (e.g. Atkinson & Al-Mutairi \textsc{2012}, Legate, Pesetsky & Yang \textsc{2013}, Hornstein \textsc{2014}, Watumull et al. \textsc{2014}), even defenders of the Chomskyan view remain unable to provide precise definitions of the recursive engine allegedly at the core of the human language faculty.
Everett (2005a, b) claimed to have found a human language (Pirahã) that lacked the proposed core property, recursion. Only a naïve falsificationist would have expected Chomsky to give up his theory before Everett’s finding had been independently confirmed. But in a follow-up paper Chomsky and his collaborators not only questioned the veracity of Everett’s report but also changed their hypothesis and essentially abandoned the only empirically testable proposal of the earlier publication. Fitch, Hauser & Chomsky (2005: 203–204) argue: ‘the putative absence of obvious recursion in one of [the human] languages . . . does not affect the argument that recursion is part of the human language faculty [because] . . . our language faculty provides us with a toolkit for building languages, but not all languages use all the tools’, and they suggest that ‘the contents of FLN . . . could possibly be empty, if empirical findings showed that none of the mechanisms involved are uniquely human or unique to language, and that only the way they are integrated is specific to human language’ (ibid.: 181). This new proposal makes UG immune to empirical refutation, because even core properties of UG are now optional. Furthermore, the remaining proposal no longer distinguishes Chomsky’s domain-specific and species-specific UG from proposals like those of Deacon (1997), Arbib (2005), and Tomasello (2008). These authors also argue that none of the mechanisms involved are uniquely human or unique to language, and that only the way they are integrated is specific to human language. In The Science of Language, Chomsky has fully retreated to his unscientific speculations of the period before 2002.

Chomsky’s most controversial proposal in the domain of language evolution is that language could not have evolved gradually through natural selection but must be the result of a single mutation. The main problem for such a proposal is that mutations do not pop into existence in a vacuum (where hypothetically everything is possible). Mutations occur in biological organisms which are constrained by the structures already encoded in their genome. So, to make his speculation minimally plausible, Chomsky would need to identify a structure that could have mutated into what he introduces in The Science of Language as ‘the basic computational principle of all natural languages’ (16), which recursively combines linguistic objects, and which Chomsky calls MERGE. To do that, he would need to specify the biological properties of Merge.8 Chomsky does neither.

In Chomsky’s recent speculations, what he calls ‘third factors’ (non-biological constraints on growth and development) carry a massive explanatory burden. Yet, given the way Chomsky invokes them in The Science of Language, they explain

[8] Other works offer some detailed suggestions about the functions of this abstract operation (Chomsky 1995, 2005, 2007). However, one learns nothing about how this abstract operation is implemented in a human brain. Chomsky claims Merge was the result of a ‘small genetic modification’ (13), some ‘slight rewiring’ (14) of the brain. But it is unclear whether this ‘rewiring’ led to novel arrangements of neurons, additional Merge-neurons, chemical-physiological modification of existing neurons, etc. Therefore, I can only use the term ‘biological’ here.
nothing: ‘The more that’s being learned about evolution and development, the more it looks like most things happen because they have to; there’s no other way’ (53). Third factors apply necessarily to all living organisms, and, hence, offer no explanation for the fact that only humans have language. Third factors constrain what is possible but do not specify which of the possibilities is realized. In order to determine this, neurophysiological research on human brains is needed, but Chomsky does not discuss this kind of research.

Instead, he offers more superficial and irrelevant musings: ‘basically, there’s one organism… the difference between an elephant and a fly is just a rearrangement of the timing of some fixed regulatory mechanism … There is deep conservation; you find the same thing in bacteria that you find in humans’ (53). At one point Chomsky acknowledges that ‘there’s got to be something there [besides third factor contributions, CB]; we’re not all amoebas. Something has got to be there; so what is it?’ (132) Instead of offering a detailed and specific hypothesis (or at least a speculation), Chomsky tells McGilvray ‘I like the edges of puzzles … Think of how boring the world would be if we knew everything we can’ (ibid.). Neither here nor anywhere in The Science of Language does Chomsky bore his readers with details about the genetic factors that he takes as determining the human language faculty.

From a biological perspective, the saltational account is so outlandish that it has been virtually unanimously rejected by researchers who disagree with each other on many other aspects of language evolution (e.g. Pinker & Bloom 1990; Deacon 1997; Studdert-Kennedy 1998; Botha 1999; Jackendoff & Pinker 2005; MacWhinney 2005; Hurford & Dediu 2007; Arbib 2008, 2012; Christiansen & Chater 2008; Tomasello 2008; Jackendoff 2011; Lieberman 2013). Evolutionary biologists call speculations ‘in which a freak mutation just happens to produce a radically different and serendipitously better equipped organism’ (Deacon 1997: 35) ‘hopeful monster’ theories, and emphasize their close resemblance to Intelligent Design or divine creation stories (for discussion see Lieberman 2013: 162–164).

Furthermore, the result of the one-time Merge mutation would have to have been exceptionally stable (allegedly Merge did not change from the moment it appeared in one lucky hominid (called by Chomsky (2005) Prometheus), some 50,000–100,000 years ago), and would have had to result immediately in massive selectional advantages for Prometheus. According to Chomsky (2009b, 2012), it is foolish to speculate about the role of communication when attempting to account for language evolution. This seems to imply that, whatever advantage was conferred on Prometheus, he could not have communicated his novel cognitive powers to other members of the breeding group.

[9] It should be noted that the literature Chomsky relies on here is either not referenced at all (Haldane) or at least 30 years old. Even reference to the recent encyclopedic overview of language evolution research by his coauthor Tecunseh Fitch (Fitch 2010) is missing. No serious scientist would provide such dated account when popularizing his/her ‘cutting edge theorizing’.
It has been argued that Chomsky’s speculations seem to be based on a profound misunderstanding of the developmental biology literature.

[It has been reported that] local genetic changes, for example, on homeobox genes, can influence the expression of other genes, and through a cascade of developmental influences, result in extensive phenotypic consequences ... But the idea that a simple tweak might lead to a complex, highly interdependent, and intricately organized system, such as the putative UG, is highly implausible. Small genetic changes lead to modifications of existing complex systems, and these modifications can be quite far-reaching; however, they do not lead to the construction of new complexity. (Christiansen & Chater 2008: 496)

The reasonable implication is that no single mutation could create from scratch a structure that fulfils the intricate functions which generativists have ascribed to Merge (see Adger 2003, Collins & Stabler 2011). One would expect that, in order to refute such a challenge, Chomsky would try to demonstrate fallacies in the argument of Christiansen & Chater, discuss the relevant biological research, and provide solid evidence in support of his speculation. But in The Science of Language Chomsky dismisses his critics as irrational, and offers only vague speculations in support of his own view. Versions of this view were proposed years ago. Consider:

Within some small group from which we are all descended, a rewiring of the brain took place in some individual, call him Prometheus, yielding the operation of unbounded Merge, applying to concepts with intricate (and little understood) properties. Guided very likely by third factor principles, Prometheus’s language provides him with an infinite array of structured expressions with interpretations of the kind illustrated: duality of semantics, operator-variable constructions, unpronounced elements with substantial consequences for interpretation and thought, etc. Prometheus had many advantages: capacities for complex thought, planning, interpretation, and so on. The capacity would then be transmitted to offspring, coming to predominate (no trivial matter, it appears, but let us put that aside). (Chomsky 2005: 13)

Here Chomsky acknowledges that he is putting aside unspecified non-trivial matters regarding a crucial point of his ‘account’, and that important details of the structural properties of the postulated biolinguistic objects are little understood. One would have expected Chomsky to have focused research efforts on matters that were little understood or had been set aside. Yet seven years later he gives no indication that any such work has been attempted. The first argument, introducing his ‘hypothesis’, in The Science of Language is a stripped down version of the 2005 ‘proposal’:

[S]ome small genetic change led to the rewiring of the brain that made this human capacity available ... Well, mutations take place in a person, not in
a group. We know, incidentally, that this was a very small breeding group – some little group of hominids in some corner of Africa, apparently. Somewhere in that group, some small mutation took place, leading to the great leap forward. It had to have happened in a single person. Something happened in a person that that person transmitted to its offspring. And apparently in a very short time, it [that modification] dominated the group; so it must have had some selectional advantage. But it could have been a very short time in a small [breeding] group. Well, what was it? The simplest assumption – we have no reason to doubt it - is that what happened is that we got Merge. You got an operation that enables you to take mental objects [or concepts of some sort], already constructed, and make bigger mental objects out of them. That’s Merge. As soon as you have that, you have an infinite variety of hierarchically structured expressions [and thoughts] available to you. (13–14)

Chomsky offers no new insights compared to the 2005 ‘account’ and merely presupposes that Merge is the essential computational operation of language and that only Merge is in need of an evolutionary account.

This is unsatisfactory for several reasons. First, Chomsky claims that ‘[internal and external Merge] come for free, unless you stipulate that one of them doesn’t happen’ (25). Yet when responding to Everett’s challenge he suggests that it is possible that not all human languages employ internal Merge: ‘[If Everett is right about Pirahã] this language has limited lexical resources and is not using internal Merge. Well, maybe not . . . No language uses all the options that are available’ (30). It is unclear how ‘limited lexical resources’ could prevent the use of internal Merge. On Chomsky’s view, all human children experience profound poverty of the stimulus and produce a language that goes far beyond the impoverished input they receive. So children receiving input which lacks the unspecified lexical resources should still be able to access their internal lexicon and activate internal Merge. Further, Chomsky claims in his more technical writings that ‘the belief that EM is somehow simpler and preferable to IM [is erroneous]. If anything, IM is simpler, since it requires vastly less search than EM (which must access the workspace of already generated objects and the lexicon) . . . It requires a stipulation that IM is barred’ (Chomsky 2013: 41). So one might expect a language that does not employ the simpler internal Merge also does not employ external Merge and thus lacks Merge altogether. Further, Chomsky also claims that ‘Merge is either available by virtue of UG, or unattainable’ (Chomsky 2012: 15). Chomsky does not clarify how it could be possible that a human language lacks access to internal Merge but insists that Merge must have evolved. But, as in 2005, he remains silent about the biological properties of Merge. Knowing these is of course essential for evaluating the next step in the argument: Merge was the result of a single mutation. This means Christiansen & Chater’s (2008) challenge remains unanswered.

Next, one crucial premiss of Chomsky’s ‘argument’ is that only Merge is in need of an evolutionary explanation. But Merge alone could give nothing resembling
a human language without a rich initial input along with a set of principles controlling generation. For example, in the simplest case, where the only input is X, Merge would yield the simple language $L_X = \{X, XX, XXX, XXXX, \ldots, X_\infty\}$. Unlike any human language, such a language would have virtually no expressive power. In other words, simply claiming only Merge is required for the general structure-building aspect of language (building trees from a pair of input trees) is insufficient to provide an even remotely plausible account of how language might have evolved. To evaluate Chomsky’s proposal, one would need an account of the nature of the linguistic objects providing the input to Merge as well as the various constraints required to achieve the required results, i.e. to generate all and only the structures of any possible human language.

Even granting for argument’s sake that only Merge is required for structure building, a host of essential details are omitted. And when one looks into these details (e.g. Collins & Stabler 2011), one finds a sophisticated and quite abstract system of dedicated linguistic features and principles. For an evolutionary account of the putative language faculty, all key components required to ‘generate’ language need to be included, and clearly these other features and principles are required. So one has to ask: Which of these other aspects have evolved and which have been learned? If they are held to be universal, presumably they must have arisen through biological evolution. Did they also arise in a single mutation? How plausible is that? And if this is not the theory, then what is? Virtually everything determining the resulting structures lies elsewhere in the linguistic system. Therefore, the highly implausible Merge speculation offers at best a grossly incomplete account.

Chomsky never provides a fuller account but talks vaguely about a small rewiring of the brain. There is not even a hint of a suggestion as to which area of the hominid brain might have been affected by the ‘rewiring’, how an ‘operation Merge’ could possibly result from such an unspecified ‘rewiring’, or how one single mutation-event could confer an immediate selective advantage to its bearer, resulting in what he calls the ‘great leap forward’ of the entire species.

An implicit but very important premiss of Chomsky’s argument is that the ‘great leap forward’ provides crucial support for his ‘account’. However, he does not explain what the term refers to, and fails to mention that the ‘great

---

[10] One needs to keep in mind that the ‘rewiring’ must have left all the important functions of the affected brain intact. It could not interfere with the individual’s cognitive or physical abilities (a severely disabled Prometheus would have had a hard time to procreate irrespectively of the miraculous abilities Merge bestowed on him).

[11] In the literature the ‘great leap forward’ refers to (i) an economic and social campaign by the Communist Party of China from 1958 to 1961, (ii) a specific hypothesis named after a term first introduced by Jared Diamond (Diamond 1999), or (iii) it is used informally to refer to marked changes in the archeological record during the Upper Paleolithic. Even though Chomsky never defines his term, we can rule out that he means (i), but it is unclear whether he refers to (ii) or (iii). Therefore I use scare quotes throughout.
leap forward’ hypothesis has been challenged (e.g. McBrearty & Brooks 2000, Lieberman 2013). It is a matter of ongoing scientific debate whether a ‘great leap forward’ occurred in all human groups. That means that another crucial premiss for Chomsky’s argument is much weaker than he makes out.

Furthermore, a sudden change apparent in the archaeological data cannot reliably indicate an increase in overall intelligence and/or signal the arrival of linguistic abilities. A 44th-century scientist, comparing the records of human technology from the 17th and 20th centuries, might conclude that our species dramatically increased in intelligence over that period. But we have little reason to believe in such an increase, and we know that humans had language before the 17th century. Hence, Chomsky would need to establish not only that the ‘great leap forward’ took place but also that it gives a good reason to believe that language evolved at that time. He does neither.

Another ‘argument’ in support of Chomsky’s speculation aims to establish that communication is not an evolutionarily relevant function of human language:

What is [language’s] characteristic use? Well, probably 99.9 percent of its use is internal to the mind. You can’t go a minute without talking to yourself. It takes an incredible act of will not to talk to yourself. We don’t often talk to ourselves in sentences. There’s obviously language going on in our heads, but in patches, in parallel, in fragmentary pieces, and so on. So if you look at language in the way biologists look at other organs of the body and their subsystems—so you take into account all its functions in talking to yourself—what do you get? What are you doing when you talk to yourself? Most of the time you’re torturing yourself [laughter]. So you might think you’re being conned, or asking why does this person treat me that way? Or whatever. So you could say that the function of language is to torture yourself. Now, obviously, that’s not serious. (11−12)

The premiss that 99.9% of language use is internal, if true, supports the conclusion that communication is not the current function of language. But this premiss implausibly implies that someone who is awake 16 hours a day uses language only 58 seconds a day for communication. A much weaker claim (for most people much of their language use is internal) might be true. But such a claim would provide little support for Chomsky’s ‘argument’.

It is not clear how the claim that we do not talk in sentences to ourselves supports the conclusion that communication is not a function of language. Furthermore, Chomsky gives no reason to believe that something, which might be true for him, is true for humanity in general. It would take massive empirical research to rule out the possibility of individual and cultural differences in this area. Finally, Chomsky makes no attempt to specify a function of language ‘in the way biologists [do when they] look at other organs of the body’. Instead, he states that his only explicit ‘proposal’ was ‘not serious’.
Chomsky’s ‘argument’ creates the impression that he subscribes to an outdated essentialist account of biological function that proposes a trait is there only because of the function it is performing. That is:

The function of X is Z iff:

(i) Z is a consequence (result) of X’s being there,
(ii) X is there because it does (results in) Z.

(Wright 1976: 81)

If one sets X to ‘language’ and Z to ‘communication’, and observes that communication can occur in the absence of language and that language does not always result in communication, this principle might support the conclusion that communication is not the function of language. But such an essentialist account is problematic: many traits have more than one function, and a trait that performed one function originally can be ‘exapted’ to perform another function (Gould & Vrba 1982). For language this means that communication could be the biological function for which language was selected even if other traits are used for communication and/or language now has a different function from the one it was originally selected for.

To use a crude analogy: The blinkers of my car are in off mode most of the time, and they can be used for several purposes (signalling slow speed when driving on the highway, signalling change in direction, greeting a friend, entertaining my kids, etc.). At times they are misleading (when I forget to cancel them after a turn or change my mind about turning), and there are other means to signal a change in direction (e.g. the Nova Scotia Driver’s Manual lists a set of hand signals). Yet, uncontroversially, the function of the blinkers is to signal my intention to change direction.

When establishing the function of a biological trait, it is not necessary to demonstrate that the trait fulfils this function most of the time and/or to rule out that other traits can also fulfil that function. It is relevant that the trait under consideration is more effective than other traits available to the organism in contexts that contribute to survival. In the current context one has to ask: Is language more effective than other means of communication to reliably convey information in contexts that matter for survival? If the answer is yes, it is irrelevant whether other traits are available to communicate or whether language is used for other purposes. This point is not even mentioned by Chomsky, let alone addressed. Likewise, numerous other published objections

---

[12] It is impossible to discuss the complexity of modern evolutionary theory adequately here. For an accessible overview of history, recent developments, and philosophical implications see Reiss (2009). Since Chomsky has been citing the same cherry-picked sources cited in The Science of Language for a decade, it is questionable whether he is fully familiar with recent developments, or indeed with older publications (e.g., Pinker & Bloom 1990, Deacon 1997, Studdert-Kennedy 1998, Botha 1999) which challenge his ‘theorizing’. 

12
Having shown how problematic and unsupported Chomsky’s own ‘account’ of language evolution is, I evaluate now his criticism of competing accounts. To illustrate the stark difference between Chomsky’s speculations and a serious proposal I briefly sketch Michael Arbib’s account. Arbib (2005) bases his proposal on comparative studies of human and monkey brains. The premotor area F5 in monkey brains and Broca’s area in human brains contain a mirror neuron system that is active for the execution and observation of manual actions. Arbib believes that meaningful gestures predated the arrival of spoken language and proposes that gradual evolutionary changes within and outside the mirror systems turned the language-less brain of our ancestors into a brain that is capable of processing human language. Arbib proposes that ‘[l]anguage readiness evolved as a multimodal manual/facial/vocal system with protosign (manual-based protolanguage) providing the scaffolding for protospeech (vocal-based protolanguage) to provide ‘neural critical mass’ to allow language to emerge from protolanguage as a result of cultural innovations within the history of Homo sapiens’ (Arbib 2005: 105).

Since both monkeys and humans have mirror systems, these systems alone cannot account for language or even language readiness. Further brain evolution was required for the mirror system for grasping to evolve into a system enabling language. Arbib hypothesizes seven stages in the evolution of language. The first three stages (S1: grasping, S2: mirror system for grasping, S3: simple imitation system for object-directed grasping) are pre-hominid. The next three stages (S4: complex imitation system for grasping, S5: PROTOSIGN, a manual-based communication system, and S6: PROTOSPEECH, requiring the ability to control the vocal apparatus with increasing flexibility) occur only in the hominid line. ‘The final stage . . . result[s] from cultural evolution (historical change) in Homo sapiens: S7: Language, the change from action–object frames to verb–argument structures to syntax and semantics; the co-evolution of cognitive and linguistic complexity’ (ibid.: 107). The proposed stages contain further sub-stages. For example S7 ‘true language’ has the following four sub-stages: LA1 – symbolization and compositionality, LA2 – syntax, semantics and recursion, LA3 – beyond the here-and-now 2: (verb tenses), and LA4 – learnability. Listing and adequately evaluating all proposed sub-stages and the cognitive innovations required for each new stage in detail would lead too far afield.

Linguists may object that this account does not explain what is of most interest to them—the structural complexity of existing human languages. This is a fair point. But Arbib’s account does illustrate three issues that are important for any serious discussion of biological evolution. First, it specifies in detail the

[13] This account is neither universally accepted nor representative for the field of language evolution. I have selected it because it was proposed almost a decade ago (giving Chomsky sufficient time to conduct the research needed to challenge or refute it), provides a scientific hypothesis, and focuses a great deal of attention on brain evolution.
cognitive and neurophysiological changes that are assumed to occur at each stage. That makes it possible to test, modify, or reject the hypothesis. Second, documenting the cognitive and neurological changes that are required to progress just from S₁ (grasping) to S₃ (complex imitation) reveals how extremely unlikely it is that one single mutation could have bridged the neurolinguistic gap between our ape-like ancestors and us. Third, recursion appears late (LA₂) in Arbib’s account. This means that the hominid brain had undergone significant changes from the ape brain before it was ‘ready’ for something like Merge. Hence, several cognitive changes that would have had to precede the arrival of Merge are in need of an evolutionary account.¹⁴

Chomsky regards all evolutionary proposals which challenge his own as not merely wrong but nonsensical:

[T]he overwhelming assumption is that language evolved slowly through natural selection. Yet that doesn’t seem at all consistent with even the most basic facts. If you look at the literature on the evolution of language, it’s all about how language could have evolved from gesture, or from throwing, or something like chewing, or whatever. None of which makes any sense. (49)

Chomsky does not describe or even refer to any of the accounts he sees as problematic, making it impossible for the reader to check the claim that they make no sense. Nor does he say what ‘the most basic facts’ are. He continues his attack: ‘[w]e know almost nothing about the evolution of language, which is why people fill libraries with speculation about it’ (51). Yet he never considers explanations which have been advanced to justify the extensiveness of this literature:

[Language evolution] is truly a multidisciplinary problem that defies analysis from any one perspective alone and where the breadth of the technical topics that must be mastered exceeds even the most erudite scholars’ capabilities. So it is hard to overestimate the task or the risk of superficial analysis, and it is unlikely that any one account can hope to achieve anything close to comprehensive treatment of the problem. (Deacon 1997: 14)

Instead of explaining how one Noam Chomsky can master the multitude of technical topics and offer comprehensive treatment, he attempts to convince his audience that all scientific work on language evolution is to be dismissed:

There are libraries of material on the evolution of human language and some scattered technical papers on the evolution of bee communication, which

---

¹⁴ It may be objected that Chomsky’s accomplishment was to minimize the components of UG that require an evolutionary explanation. But he offers no serious account that could be evaluated. In one publication he speculates that it is conceivable that ‘some random event [maybe some strange cosmic ray shower] causes a language faculty to be installed in [a primate] … In fact it is conceivable … that higher primates, say gorillas or whatever, actually have something like a human language faculty but just have no access to it’ (Chomsky 2000b: 16–17). Yet according to Arbib the differences between gorilla and human brain are too great to allow for such ‘instantaneous installation’.
mostly point out that it’s too hard to study, although it’s vastly easier to study than evolution of human language. This is just irrational. (105)

Saying that the study of language evolution is vastly more difficult than the study of bee communication does not lead to the conclusion that it is irrational to study the former. If the resulting insights are more valuable, it is sensible to spend more resources on that task. By analogy, not knowing how to cure cancer in less complex organisms does not make it irrational to search for a cure of cancer in humans. Chomsky hardly ever identifies the researchers who are producing massive amounts of allegedly worthless literature. Only in one case does he provide a name:

Many of these people, like Dawkins, regard themselves very plausibly as fighting a battle for scientific rationality against creationists and fanatics and so on. And yes, that’s an important social activity to be engaged in, but not by misleading people about the nature of evolution— that’s not a contribution to scientific rationality. (105)

One looks in vain for citation of any non-rational argument that Dawkins has offered, much less any counterevidence to Dawkins (unidentified) non-rational arguments, or any characterization of what it is that makes these (unidentified) arguments irrational. The reader is apparently supposed to come away from this passage convinced that one of the major evolutionary theorists of the past century has provided not merely incorrect arguments, but arguments which are not a contribution to scientific rationality— on the basis of nothing more than Chomsky’s assertion that this is so.

The foregoing has shown that Chomsky’s language evolution speculations in The Science of Language do not meet the criteria of scientific work or its reporting (or even its popularizing). It requires no formal training in linguistics to recognize that Chomsky’s speculations are not supported by credible evidence, are false when taken literally, and are mere rhetorical devices if interpreted as defensible hyperbole in an informal context. Chomsky advances problematic ‘arguments’ that provide virtually no support for his claims, fails to address serious objections, and accuses others of irrationality without providing supporting evidence. This style of arguing from authority, found throughout The Science of Language and many other recent publications (e.g. Chomsky 2000a, b, 2002, 2009a, b, 2010, 2013), contrasts starkly with serious scientific practice. It is also contrary to advice Chomsky has offered to others:

In general, the willingness to rely on ‘experts’ is a frightening aspect of contemporary political and social life. Teachers, in particular, have a responsibility

[15] I object to the attribution of irrationality to researchers who engage in work Chomsky disapproves. Presumably not all work on language evolution will provide valuable results. But this is the case for scientific research in general: if we knew all the answers we would not need to conduct research. Chomsky has filled volumes with proposals that have been retracted later. Does this make him irrational?
to make sure those ideas and proposals are evaluated on their merits, and not passively accepted on grounds of authority, real or presumed (Chomsky 1966c: 46)

Yet in spite of its obvious flaws, *The Science of Language* has been accepted by senior scholars (seemingly based only on Chomsky’s perceived authority), and no generativists have distanced themselves publicly from this work.

4. FROM TRADITIONAL SCIENCE TO ‘GALILEAN STYLE’

Chomsky’s remarkable success has been linked to his innovative proposals, his persuasive debating style, and his emphasis on rigorous scientific methodology for linguistic research. It has been suggested that a cognitive revolution occurred because:

[T]he accepted model or “paradigm” of linguistics was confronted, largely by Chomsky’s work, with increasing numbers of nagging counterexamples and recalcitrant data which the paradigm could not deal with. Eventually the counter-examples led Chomsky to break the old model altogether and to create a completely new one. (Searle 1972: 2–3)

Searle’s remarks referred mainly to the work by Chomsky that was instrumental to replacing the dominant behaviourist framework in psychology with a renewed mentalism. During his early career, Chomsky was perceived to be following accepted scientific procedure, and this work earned him the respect of the scientific community.

Yet over the years Chomsky’s attitude has changed dramatically, and he writes now as if no data can challenge, far less falsify, his own ‘theories’. Decades ago he began advocating a selective approach to data. For example, he proposed that one ought not to be worried by counterevidence to one’s view: '[s]uppose that counterevidence is discovered as we should expect and as we should in fact hope, since precisely this eventuality will offer the possibility of a deeper understanding of the real principles involved’ (Chomsky 1982a: 76). It is, to say the least, odd to claim that scientists should hope to discover counterevidence to their views. Counterevidence may lead to the refinement of a theory or even in some circumstances lead to that theory’s abandonment. Chomsky’s ‘proposal’

---

[16] It is a matter of debate whether Chomsky’s early work was as rigorous as it is often described (for some critical evaluation see Seuren 1998, 2009; Sampson 2001; Pullum 2011a). For my argument it is not necessary to establish how rigorous this work in fact was. I am only claiming that Chomsky’s work was widely perceived as rigorous.

[17] Most scientists are of course aware that even their best theories are preliminary, and that future research may force them to abandon those theories. This philosophical point about progress in science is different from Chomsky’s remark that scientists should hope to discover counter evidence. Chomsky implies incorrectly that falsification is the only way to make intellectual progress.
remains silent about how counterevidence can contribute to theory refinement and seems to have no value beyond insulating his views against rational criticism.

Chomsky’s willingness to accept that his own view could be incorrect and/or that another view could account better for the phenomena under investigation seems to have vanished over the years. For example, when asked what kind of empirical discovery would lead to the rejection of the strong minimalist thesis, Chomsky replied: ‘All the phenomena of language appear to refute it, just as the phenomena of the world appeared to refute the Copernican thesis’ (Chomsky 2002: 124, emphasis added). Instead of offering an explanation for why he continues to defend a thesis which all phenomena of language appear to refute or what makes the refutation only apparent, he provides an irrelevant and misleading reference to an unrelated thesis.18 Copernicus’ proposal of heliocentrism was based on careful observations of the phenomena of the world, which suggested that the accepted view—geocentrism—was incompatible with these phenomena. Further, Copernicus’ incorrect proposal (that the planets move on circular orbits) has been refuted on the basis of observation of the phenomena of the world.

Chomsky has moved further and further towards large-scale dismissal of data that are inconvenient for his view. He even coined a term for this unorthodox methodology: ‘the Galilean style … is the recognition that … the array of phenomena is some distortion of the truth … [and] it often makes good sense to disregard phenomena and search for principles’ (Chomsky 2002: 99). Chomsky calls this attitude the ‘Galilean move towards discarding recalcitrant phenomena’ (Chomsky 2002: 102).

Chomsky’s move from accepted scientific practice to his ‘Galilean style’ of massive data dismissal is based on his idiosyncratic interpretation of Galileo’s work: ‘[Galileo] dismissed a lot of data; he was willing to say: “Look, if the data refute the theory, the data are probably wrong.” And the data that he threw out were not minor’ (Chomsky 2002: 98). Chomsky has provided no evidence that Galileo actually followed this style,19 but he has asserted that it was accepted by other famous scientists (e.g. Copernicus, Newton, Einstein, Monod). He has claimed that it ‘is pretty much the way science often seems to work … You just see that some ideas simply look right, and then you sort of put aside the data that refute them’ (Chomsky 2009a: 36).

---

[18] Chomsky routinely attacks opponents for using the allegedly incorrect term ‘innateness-hypothesis’ when referring to his work. Yet when referring to the work of others, he often uses the kind of convenient shorthand he condemns in others. Here he fails to mention that Copernicus’ main work (De revolutionibus orbium coelestium) contained several theses and to specify which of these he means by ‘the Copernican thesis’. Further, even if there was one ‘Copernican thesis’ that seemingly was refuted by data but turned out to be correct, it does not follow that another, entirely unrelated, Chomskyan ‘thesis’ is also correct.

[19] Fischer (1992) argues convincingly that the Galilean Method has been seriously misrepresented by Paul Feyerabend. It seems that Chomsky’s ‘Galilean style’ is based not on Galileo’s work but on Feyerabend’s misinterpretation of that work.
It is debatable whether in some cases in the past the Galilean style has worked. However, the history of science provides many cases in which ideas that simply looked right turned out to be wrong. Relevantly, many ideas proposed by Chomsky at various stages of his career belong in this second category. In his early work Chomsky proposed that the base of a transformational grammar involves phrase structure (rewriting) rules (e.g. Chomsky 1957: 13–17; 1962: 127; 1975: 80). In later work these rules were abandoned (e.g. Chomsky 1986a: 82–83; 1986b: 3; 1995: 25). Chomsky first proposed that transformational rules are either optional or obligatory (Chomsky 1957: 45; 1962: 136), then abandoned this proposal (Chomsky & Lasnik 1977: 41), and claimed that transformational rules are all optional (Chomsky & Lasnik 1977: 41). This proposal was subsequently abandoned (Chomsky 2000c: 130). In his early work Chomsky proposed that a notion of well formed (sentence) is fundamental to linguistics (Chomsky 1957: 13–14; 1966b: 32; 1972: 64), but, decades later he abandoned that proposal (Chomsky 1995: 194, 213). In early work Chomsky insisted that the notion of deep structure is fundamental to natural language syntax (Chomsky 1966b: 91), later he announced that language has no D-structure (Chomsky 1995: 186–189). And at times only a few years lapsed between a proposal and its subsequent abandonment. For example, in 1982 Chomsky claimed that parasitic gaps are pronominal (and not traces) (Chomsky 1982a: 38, 52), yet four years later this proposal was abandoned (Chomsky 1986b: 56).20

These are only a few of many instances in which Chomsky changed his mind about ideas that just seemed right. It is of course not uncommon that scientists revise their ideas as more evidence becomes available. But given the frequency with which Chomsky has changed his mind about his own ideas, including ideas he claimed to be fundamental to his view (e.g. deep structure), it is irresponsible to insist that his abductive instinct is virtually infallible and allows him to decide which data he can safely set aside.

One might expect that leading generativists have criticized Chomsky’s ongoing violations of accepted scientific practice. This is not the case. It had been proposed: ‘A significant feature of the Generative Revolution in linguistics has been the development of a Galilean style in that field’ (Freidin & Vergnaud 2001: 647) and Chomsky has been complimented for ‘implementing the Galilean style [which] entails a readiness to tolerate unexplained phenomena or even as yet unexplained counterevidence to theoretical constructions’ (ibid.). Another generativist defended Chomsky’s departures from accepted scientific practices: ‘Can [Chomsky’s] goal be pursued by “normal scientific procedure”?’ Let us remember we are talking about someone who tries to reinvent the field every time he sits down to write. Why should we expect Chomsky to follow

---

20 I am indebted to Paul Postal for providing these examples. For further discussion see Postal (2012).
normal scientific practice...?’ (Fiengo 2006: 471). To my knowledge no generativist has responded to Fiengo’s rhetorical question in print. This collective silence could create the impression that generativists consider accepted scientific practices as obsolete for their own discipline. Yet, this is clearly not the case. Perceived violations of scientific practice by others are criticized:

In general, the target article eschews precise analysis and offers instead a number of general descriptive statements about a range of languages that are taken to be refutations of proposed universals. Often these are offered without citation of data or, more importantly, the explicit analysis that would support the claims. Such refutations are based on unstated and unverified assumptions about grammatical analysis. In general, such discussion is not useful. (Freidin 2009: 454)

Freidin makes it clear that he does not tolerate deviations by non-generativists from accepted scientific practice. He is not alone. Other generativists find it unacceptable when a researcher changes his mind about a previous analysis, and (allegedly) does not justify this change adequately:

CA contains statements about the facts (and analyses that predict new facts) that appear to be directly contradicted by data presented and analyzed in DISS and HAL. If CA had reanalyzed the contradictory data successfully, or pointed to factual errors in the earlier work, it would be very reasonable to part company with DISS and HAL and seek some new account of the correct facts. This is not, however, the situation. CA’s new claims are in general presented with only a cover statement that CA ‘supersedes any other published or unpublished statement by me on those aspects of Pirahã grammar here addressed’ (CA: 621, n. 1), without any discussion (or even acknowledgment) of specific relevant published counterevidence. (Nevins, Pesetsky & Rodrigues 2009: 360)

Finally, another leading generativist states unambiguously that doing linguistic research does not warrant abandoning scientific practice:

All other sciences have progressed precisely by taking nothing at face value—diversity included. Evans & Levinson (E&L) claim, in effect, that linguistics is different from all other fields. If they are right, the search for deeper laws behind linguistic structure is a fool’s errand, and languages are just as

[21] It would go beyond the scope of this review article to evaluate whether the researchers criticized below are guilty as charged. The examples are used to demonstrate that generativists are not universally oblivious to violations of accepted scientific practice.

[22] The abbreviations in this passage refer to works by Daniel Everett: CA is an article in Current Anthropology (Everett 2005b), HAL are sections included in the Handbook of Amazonian Languages (Everett 1986), and DISS is Everett’s dissertation from the Universidade Estadual de Campinas (published in Portuguese as Everett 1987); all are referenced in Nevins et al. (2009).
inexplicably diverse as they seem at first glance. It is thus surprising that E&L’s article contains no discussion of the actual research to which they supposedly object. (Pesetsky 2009: 464)

The generativists I have quoted state clearly that the Galilean style is unacceptable when it seems to be adopted by non-generativist researchers. But they offer no answer to Fiengo’s question, at least implicitly suggesting that one should not expect Chomsky to follow the same practices others are required to follow.

Since the linguistic community seems to tolerate Chomsky’s practice, it is unsurprising that The Science of Language continues to promote the Galilean style. He now suggests: ‘if we want a productive theory-constructive [effort], we’re going to have to relax our stringent criteria and accept things that we know don’t make any sense, and hope that some day somebody will make some sense out of them’ (91). Why would anyone defend a position that requires one to accept knowingly things that make no sense? McGilvray seems untroubled by this ‘proposal’. Nor does he object to the wholesale dismissal of empirical data for linguistic theorizing proposed in the ‘argument’ from the Norman Conquest:

Take the Norman Conquest. The Norman Conquest had a huge effect on what became English. But it clearly had nothing to do with the evolution of language - which was finished long before the Norman Conquest. So if you want to study distinctive properties of language - what really makes it different from the digestive system … you’re going to abstract away from the Norman Conquest. But that means abstracting away from the whole mass of data that interests the linguist who wants to work on a particular language. There is no contradiction in this; it’s just a sane approach to trying to answer certain kinds of far-reaching questions about that nature of language. (84)

The vague formulation of this argument makes evaluation difficult. If, when studying distinctive properties of $L_1$, one should abstract away from the whole mass of data of interest to the linguist about $L_2$, the same logic would hold for $L_2 \ldots L_n$. So, if taken literally, the argument abstracts away from everything of linguistic interest about all languages to uncover the nature of language and explain how it differs from digestion.

Idealization and abstraction are of course part of the scientific method. But given how little is currently known about the core properties of language, such wholesale abstraction is hardly responsible. Even on a more charitable reading, the Argument from the Norman Conquest is incompatible with Chomsky’s earlier view that ‘the linguist is always involved in the study of both universal and particular grammar … his formulation of principles of universal grammar must be justified by the study of their consequences when applied in particular grammars’ (Chomsky 1968: 24). Research in the years since 1968 has uncovered counter-examples to many proposed principles that were once considered to
hold universally. This suggests that it remains important to verify that any proposed UG is compatible with the data from particular grammars.

The argument from the Norman Conquest is representative of the problematic nature of the arguments in *The Science of Language*. They are not the sorts of arguments used in science, but suggest that Chomsky is unaware of many current developments, is scornful of the scientific method, and is considering any criticism of his ‘theories’ as irrational dogmatism. The next section highlights some important events that contributed to the change in Chomsky’s approach to scientific work.

5. FROM PROMISSORY NOTES TO REWRITING HISTORY

One of Chomsky’s early works, *Syntactic Structures*, contained many innovative proposals but also a non-trivial promissory note. While certainly innovative by the standards of the time, *Syntactic Structures* was less formal and rigorous than widely believed: ‘The exposition of finite-state systems in *Syntactic Structures* is particularly informal: no definitions are given, grammars are not distinguished from accepting automata, finite-state Markov processes are not distinguished from their transition graphs’ (Pullum 2011a: 279). Furthermore, as an anonymous *JL* referee had pointed out, ‘formulat[ing] rigorously an algebra of transformations’ (Lees 1957: 400) was left as a task for the future. But, *Syntactic Structures* was introduced as a small part of a much larger work (*The Logical Structure of Linguistic Theory*), to which Lees’ review explicitly referred: ‘A good start has already been made in this direction in Chomsky’s *The logical structure of linguistic theory*, Chap. 8 (cf. fn. 1)’ (ibid.: fn. 43). Thus, while it was admitted that a promissory note existed, it was unclear what it contained or how quickly it could be discharged.

It is hard to say whether the success of *Syntactic Structures* encouraged Chomsky to publish other work containing promissory notes and inaccuracies. But he did and, at least initially, was rarely criticized for it. One exception was the largely negative response to *Cartesian Linguistics*. This work was published with the purpose of deepening ‘our understanding of the nature of language and the mental processes and structures that underlies its use and acquisition’ (Chomsky 1966a: ix). Tracing the history of contemporary linguistic theorizing to Cartesian antecedents might be a worthwhile project, but it would require more diligent research than Chomsky undertook. The result was unsatisfactory (especially from a historian’s perspective) and *Cartesian Linguistics* was severely criticized

because the scholarship is poor, because the texts have not been read, because the arguments have not been understood, because the secondary literature that might have been helpful has been left aside or unread even when referred to . . . Universal Grammar is profoundly important in the history of linguistics, but Chomsky’s account fails to grasp both the nature and the history of that importance. (Aarsleff 1971: 584)
Chomsky has never accepted the objections to *Cartesian Linguistics*. He has referred to traditional historical studies as antiquarianism and proposed a better way to do (or rewrite) history. This allows him to ‘refute’ any criticism:

For Descartes the mind is not part of the biological world. Furthermore he appears to regards the mind as uniform and undifferentiated . . . One might then argue that we are not studying Descartes’ problem when we consider the human mind as a specific biological system, one with components and elements of a varied kind, to be explored as we would study any other aspect of the physical world. This conclusion holds, however only if we regard Descartes as an irrational dogmatist. (Chomsky 1980: 30–31)

In other words, those attributing to Descartes the views he actually held consider him to be an irrational dogmatist. For decades the overwhelming majority of the linguistic community were quite tolerant of Chomsky’s refusal to accept justified criticism of his scholarship and his denigration of the standards of a field in which he had claimed expertise. This toleration had far-reaching consequences. I would argue that it allowed Chomsky gradually to relax the standards for his own work and to misrepresent and distort the work of others. These outcomes are illustrated in the next sections.

6. **THE STEADY STATE: ‘ARGUMENTS’ TO SUPPORT ‘PROPOSALS’**

For decades the ‘hypothesis’ that language is a domain-specific innate faculty has been central to Chomsky’s theorizing. Initially poverty of the stimulus arguments seemed to provide strong support for this ‘hypothesis’. However, as philosophical challenges to this ‘hypothesis’ were proposed (e.g. Putnam 1967, Goodman 1968) and empirical counter-evidence surfaced (e.g. MacWhinney & Snow 1985; Pullum & Scholz 2002; Sampson 2002; Tomasello 2003, 2004; MacWhinney 2010), some of Chomsky’s ‘arguments’ in support of his allegedly ‘detailed and specific innateness hypothesis’ became entirely inappropriate for scientific discourse:

The telephone exchange, for example, has ‘heard’ much more English than any of us, but lacking the principles of universal grammar (inter alia) it

---

[23] Apparently Chomsky harboured such deep resentment that he refused to publish in *Language* after another review (Aarsleff 1970) had been published. *Language* editor Bill Bright ‘wrote to [Chomsky] in 1984 urging him to consider submitting to the journal. Chomsky stunned Bill with an angry reply that he would never consider publishing in a journal that has published “outright falsifications so scandalous that they raise serious questions about the integrity of the field . . . that lies . . . couched in a rhetoric of a sort that might be appropriate to some criminal, but that one is surprised to find in a scholarly journal” (Bright 1985 ms. p. 11)’ (Hill 2007: 636).

[24] Chomsky proposes that historians should be focusing only on events that were of interest to them and on the way history should have happened (e.g. Chomsky 1971, 1997). For a detailed discussion see Behne (2009).

[25] Serious Poverty of the Stimulus arguments might well support Chomsky’s position. However, the existence of serious arguments does not excuse Chomsky’s use of ridiculous arguments that denigrate his intellectual opponents. The latter reflects poorly on a field that accepts and/or excuses his practices.
develops no grammar of English as part of its internal structure. (Chomsky 1981: 8)

This ‘argument’ is no isolated example. Chomsky has repeatedly offered arguments of this dubious character in support of the claim that humans have a species and domain-specific innate language faculty:

To say that language is not innate is to say that there is no difference between my granddaughter, a rock and a rabbit. In other words, if you take a rock, a rabbit and my granddaughter and put them in a community where people are talking English, they’ll all learn English. If people believe that, then they believe that language is not innate ... To say ‘language is innate’ is to express the belief that some crucial and relevant internal nature differentiates my granddaughter from rocks, bees and chimpanzees. (Chomsky 2000b: 50–51)

Chomsky creates false dilemmas and trivializes an important scientific dispute. No one denies that there are differences between rabbits, bees, cats, rocks, the telephone exchange, and Chomsky’s granddaughter. Implying otherwise distracts from the fact that Chomsky makes no serious attempt to suggest any research that could establish which of the MANY cognitive and physiological differences between humans and other animals might account for human language acquisition. Some linguists have exposed the silliness of what is known informally as ‘rocks-and-kittens’ argument (Postal 2005: ix; Pullum 2011b). This has not prevented Chomsky from using it in a vindictive attack on Margaret Boden:

[Boden is] questioning the existence of a genetic factor that played a role in my granddaughter’s having reflexively identified some part of the data to which she was exposed as language-related, and then proceeding to acquire knowledge of a language, while her pet kitten (chimp, songbird, etc.), with exactly the same experience, can never even take the first step, let alone the following ones. It is either a miracle, or there is a genetic factor involved. Boden’s suggestion – presumably unwitting – is that it may be a miracle. (Chomsky 2007: 1095)

As can be easily confirmed, the work that Chomsky attacks (Boden 2006) contains no such silly suggestions. Again, to my knowledge, no Chomsky supporter has objected to the use of such illegitimate rhetorical tools to discredit the work of another researcher.

[26] It is of course possible that Boden’s (2006) account of the history of generative grammar and of Chomsky’s work contains factual errors. However, such errors can be identified and corrected without descending into demeaning and incorrect personal attacks. In her reply to Chomsky’s attack, Boden (2008) remarks on the different reactions Chomsky’s review elicited: ‘an American urological surgeon who’d found the review on the web sent me a very funny e-mail, saying “As a urologist, I recognize a “pissing contest” when I see it”’ (Boden 2008: 1953). Apparently the urologist assumed that Chomsky’s review must have been the reaction to an equally malicious publication by Boden. In stark contrast, a linguist colleague remarked: ‘To be savaged by Chomsky is a badge of honor’ (ibid.: 1962).
Furthermore, the rocks-and-kittens argument has been cited in work promoting generative grammar to the general public:

Chomsky professes himself puzzled by such objections [to the innate LAD, CB]: “Is the idea supposed to be that there is no (relevant) difference between my granddaughter, her pet kitten, a rock, a chimpanzee? No one can be that crazy. But if there is a relevant difference, then we accept the ‘innatist scenario.’ So what can possibly be under discussion, except for the nature of the innate endowment?” I find this response convincing. (Smith 1999: 169)

Smith fails to expose the logical fallacy committed by the ‘argument’ or to acknowledge that no serious scientist is that crazy. It is true that some linguists and psychologists have rejected more serious versions of the poverty of the stimulus argument (e.g. Sampson 2002, Tomasello 2003, MacWhinney 2004), but such rejection does not commit anyone to accepting the rocks-and-kittens argument.

The Science of Language does not contain the rocks-and-kittens argument but it offers many equally problematic ‘arguments’. In support of his controversial semantic internalism Chomsky argues:

Take children[’s] stories; they’re based on these [internalist, CB] principles. I read my grandchildren stories. If they like a story, they want it read ten thousand times. One story that they like is about a donkey that somebody has turned into a rock. The rest of the story is about the little donkey trying to tell its parents that it’s a baby donkey, although it’s obviously a rock. Something or another (sic) happens at the end, and— it’s a baby donkey again. But every kid, no matter how young, knows that that rock is a donkey, that it’s not a rock. It’s a donkey because it’s got psychic continuity, and so on. That can’t be just developed from language, or from experience. (27)

This passage is intended as an ‘argument’ to establish that children could not have learned the concept psychic continuity from experience or from instruction. Developmental psychologists study the conditions under which children impute intentionality to objects. But Chomsky does not cite any such work. Instead, he derives his entire evidence from one fairytale, told to an undisclosed number of grandchildren. It is unsurprising that his grandchildren could not have learned from experience with the actual world that donkeys turning into rocks and back into donkeys retain their psychic continuity. Yet this ‘evidence’ hardly establishes that they must have an innate concept of psychic continuity. Like the rocks-and-kittens argument, the donkey tale has been used in other publications (Chomsky 2008, 2009b) as support for Chomsky’s view, and it has found its way into the discourse of fellow biolinguists (e.g. Uriagereka 2009: 405–406).

7. Distortions and Double Standards

Chomsky’s treatment of intellectual opponents warrants special discussion. By the time The Science of Language was published Chomsky had extended his
rewriting of history to current events as well. He now considers it justified to attribute to his opponents views that are easy to refute, instead of engaging with the views they actually hold. And many opponents remain unnamed: ‘It is quite striking how sensible people, even those embedded in the sciences, just take a different approach toward human mental faculties and the mental aspects of the world than they do to the rest... That’s just irrational’ (123). Without knowing who these scientists are, which views they hold, and how their approaches to the study of the development of the visual and linguistic system differ, it is impossible to evaluate whether they are wrong, still less irrational.

When Chomsky supplies enough information to identify his opponent, it becomes clear that his criticism is either incorrect or exaggerated.

Some of the stuff coming out in the literature is just mind-boggling ... The last issue [of Mind and Language] has an article – I never thought I would see this – you know this crazy theory of Michael Dummett’s that people don’t know their language? This guy is defending it. (57)

‘This guy’, at the time a graduate student, had written a philosophical paper (Lassiter 2008) in which he discusses Dummett’s view. Lassiter’s argument is subtle and on a superficial reading it may appear at times as if he agrees with Dummett. However, eventually he rejects Dummett’s view as untenable. Chomsky could have criticized the clarity of expression, but instead he dismisses Lassiter’s entire paper as defending a crazy view. Further, he provides no reference to any work by Dummett that could be consulted to evaluate whether his theory is crazy.27

In addition to many comparable dismissals, The Science of Language contains an explicit attack on one of Chomsky’s favourite targets: Jeff Elman. Chomsky has misconstrued the work of Elman for years. Take for example:

One of the most quoted connectionist papers is Jeffrey Elman’s work on how you can get two nested dependencies. This is true, you can write a program that will do that. But Elman’s program totally breaks down when you get to three, and you have to rewrite the whole program. In Turing machine terms, the control unit has to be totally changed, which means you’re not capturing the rules.

[27] Earlier publications contain several attacks on Dummett and there Chomsky references Dummett (1986, 1991, 1993). However, the evidence offered in support of the claim that Dummett’s proposals have a paradoxical flavor is one citation suggesting that language is a practice that ‘is learned from others and is constituted by rules which it is part of social custom to follow’ (Dummett 1986: 473, cited in Chomsky 2000b: 48) This view hardly qualify as paradoxical or crazy. Furthermore, for his own ‘theories’, Chomsky claims: ‘...paradoxes have been very fruitful in the past, and I imagine this one will be when it is understood and it may indicate some way in which the theories are to be deepened’ (Chomsky 1982b: 76).
And to make things worse, his approach also works for crossing dependencies, so in the case of the example earlier:

(4) *The men who John see is tall.

It works just as well for those. (Chomsky 2009b: 392)

Similar claims appear in *The Science of Language*:

[T]ake Elman’s paper[s] . . . on learning nested dependencies. Two problems:
(1) the method works just as well on crossing dependencies, so doesn’t bear on why language near universally has nested but not crossing dependencies.
(2) His program works up to depth two, but fails totally on depth three.
(Chomsky cited by McGilvray, page 226)28

Chomsky attributes an absurd view to Elman and, on the basis of this view, rejects the work of all connectionists. However, virtually all claims about Elman’s work are incorrect. Elman has never claimed that his program29 works just as well on nesting and crossing dependencies. He has not reported any work on crossing dependencies and, hence, could not have claimed that his method works on them.

Connectionist work on crossing dependencies has been reported by Morten Christiansen and Nick Chater. They state that ‘[t]he fact that cross-dependencies cannot be handled using a context-free phrase structure grammar . . . appears to demonstrate that natural language is not context-free’ (Christiansen & Chater 1999: 162) and report that ‘simulation results indicate that SRNs can embody constraints which limit their abilities to process center-embeddings and cross-dependencies to levels similar to human abilities’ (ibid.: 165). Crossing dependencies are not processed ‘just as well’ as nested dependencies. They are processed worse than right branching dependencies but better than center embeddings. Furthermore, the authors comment explicitly on the surprising difference in performance between center embeddings and crossing dependencies:

[T]he nets appeared to find the cross-dependency language easier to learn than the center-embedding language (at least in terms of their ability to reduce MSE). This is an important result because people appear to be better at dealing with cross-dependency constructions than equivalent center-embedding constructions. This is surprising from the perspective of linguistic theory because,

---

28 Here as elsewhere in *The Science of Language*, Chomsky’s use of terminology is informal and misleading. I do not wish to add to the ‘terminological befuddlement’ (Tomalin 2011: 313) that has plagued linguistic recursion debates for decades and will use mostly direct quotes below. A detailed discussion of the countless misunderstandings resulting from Chomsky’s informal use of the term ‘recursion’ exceeds the scope of this article.

29 Chomsky’s use of ‘program’ indicates that he may not be familiar with what Elman actually did. A simple recurrent network (SRN) is not a program—though it is implemented by way of a program. Elman did not code the SRNs or reprogram them, instead they ‘learn’ from the input they have been exposed to during ‘training’. The use of intentional terminology in reference to SRNs can be criticized. But talking in this context of programs makes no sense.
as we noted above, cross-dependency constructions are typically viewed as more complex than center-embedding constructions because they cannot be captured by phrase-structure rules. (Christiansen & Chater 1999: 177)

Since the performance of humans and SRNs is similar one might suspect that ‘the language processor cannot be primarily based on a stack-like memory store. This is because cross-dependencies, which require a queue, are easier to process than center embeddings, which require a stack’ (ibid.: 186). Rather than discussing (and maybe challenging) the findings reported by Christiansen & Chater (1999), Chomsky does not even mention this important paper, which has been cited by hundreds of researchers. This indicates either a surprising lack of familiarity with research in a field Chomsky so confidently criticizes, or willful omission of discussion of work that contradicts his claims.

Further, Chomsky’s remarks about Elman’s work are incorrect. Elman and his coworkers have repeatedly reported differences in performance for different types of nested dependencies: ‘given the prediction task, the network is more successful at right-branching structures than center-embedded ones’ (Weckerly & Elman 1992: 414, see also below). One important finding of this work was that it showed that ‘the network’s performance parallels that of human listeners (ibid.: 418). In his criticism, Chomsky fails to distinguish between different types of nested dependencies. This is problematic for the claim that ‘[Elman’s] program works up to depth two but fails totally on depth three’ (226). On right branching dependencies, Elman has not reported any failure on depth three and for center embedded dependencies Elman reported:

In the current simulation, the representation degraded after about three levels of embedding. The consequences of this degradation on performance (in the prediction task) are different for different types of sentences. Sentences involving center embedding (e.g., 9c and 9d), in which the level of embedding is crucial for maintaining correct agreement, are more adversely affected than sentences involving so called tail-recursion (e.g., 10d). (Elman 1991: 215)

Elman does not say that his program fails totally at level three but that representations degrade. Further, he specifically states that there exist differences between sentences involving different types of recursive structures at level three.

In three publications, appearing within four years, Chomsky did not once cite any passage from Elman’s work. In The Science of Language he dismissed this

---

[30] A psycholinguistic study of German and Dutch showed that people behave like the SRNs (for details see Bach, Brown & Marslen-Wilson 1986).

[31] For a current number of citations see Google Scholar: http://scholar.google.com/citations?view_op=view_citation&hl=en&user=_qjbd88AAAAJ&citation_for_view=_qjbd88AAAAJ:M3NEmzRMkIC. Presumably this number was slightly lower at the time The Science of Language had been published.

[32] In July 2009 Chomsky had been made aware of Christiansen & Chater (1999). Therefore, intentional deception cannot be ruled out for the remarks in The Science of Language.
work based only on a misinterpretation he had previously been made aware of. In light of such deception, one has to question the sincerity of Chomsky’s remarks about morality in *The Science of Language*: ‘We apply to ourselves the same standards we apply to others – probably more rigorous standards if you’re serious’ (101).

Actually though, the standards one would expect Chomsky to apply to himself are violated countless times in *The Science of Language* and in many of his recent publications. In his attacks on the work of language evolution researchers he writes:

There are a lot of [theories of language evolution, CB] but there’s no justification for any of them. So for example, a common theory is that somehow, some mutation made it possible to construct two- word sentences; and that gave a memory advantage because then you could eliminate this big number of lexical items from memory. So that had selectional advantages. And then something came along and we had three word sentences and then a series of mutations led to five … finally you get Merge, because it goes to infinity. (15)

One example hardly supports the claim that there is no justification for ANY existing theory of language evolution. Further, the ‘common theory’ is absurd indeed but it is not defended by anyone.33 To ridicule work he disagrees with Chomsky has invented an allegedly common account that bears no resemblance to any known theory. Given the variety of existing theories and the willingness of researchers to clarify their views and engage with targeted criticism, detailed discussion and competent critique of existing theories would have been a valuable contribution. Instead, Chomsky is satisfied with creating and dismissing an insultingly bad straw-man ‘argument’.

Examples of severe distortion of the arguments of others are also found in Chomsky’s recent work outside *The Science of Language*. In one publication, Chomsky claimed that proposals by neuroscientist Terrence Deacon seemed to ‘reshape standard problems of science as utter mysteries, placing them beyond any hope of understanding, while barring the procedures of rational inquiry that have been taken for granted for hundreds of years’ (Chomsky 2002: 83).

This amazingly heated accusation is based on one argument by analogy about language evolution, which Chomsky claims to be ‘so vague that it is often hard even to guess what might have been [meant] … Whatever the meaning may be, the conclusion seems to be that it is an error to investigate the brain’ (ibid.).

---

33 The allegedly common theory was not referenced and could not be found in the language evolution literature. I contacted 16 researchers and none of them had heard of it or could imagine anyone would seriously propose such a theory. Several individuals were insulted that anyone would imply they might defend such nonsense. The comments ranged from ‘a caricature’ (Corballis, MacWhinney), ‘willful ignorance’ (Lieberman), ‘very annoying’ (Jackendoff), ‘vaguely incoherent’ (Studdert-Kennedy), ‘ridiculous’ (Bickerton) to terms not suitable for citation.
Yet, in the passage Chomsky criticizes, Deacon proposes that ‘the proper tool for analyzing language structure may not be to discover how best to model them as axiomatic rule systems but rather to study them the way we study organism structure: in evolutionary terms’ (Deacon 1997: 110). Deacon rejects a Chomskyan LAD and proposes that languages are exposed to selection pressure when passed on from one generation to the next (when children learn language from their parents). To illustrate this proposal he uses the virus analogy: ‘In some ways it is helpful to imagine language as an independent life form that colonizes and parasitizes human brains, using them to reproduce’ (Deacon 1997: 111). One may, of course, question the utility of this analogy. But the proposal does not commit Deacon to the belief that languages are viruses any more than Chomsky’s perennial invocation of Martian scientists commits him to the belief that Martian scientists exist.

One might be willing to grant that Chomsky truly did not understand that Deacon was speculating. But even the most charitable reader cannot assume Chomsky had missed a section called Brain, which is more than 100 pages in length. In that section Deacon describes in detail how brain research can shed light on our understanding of language. The aim of this section is clearly stated: ‘[we can now attempt] to determine what happened to human brains to make it possible for our ancestors to break through [the language-learning, CB] barrier. If the symbol-learning problem is the threshold that separates us from other species, then there must be something unusual about human brains that helps us to surmount it’ (Deacon 1997: 145). Chomsky can of course challenge any specific proposal Deacon has made. But claiming that Deacon believes it is an error to investigate the brain is preposterous.

Chomsky’s irresponsible remarks about scientists are not limited to his linguistic publications. For example, in The Philosophy of Generative Linguistics one reads:

If you read the scientific literature, there’s a lot of expressed contempt for what philosophers have to say. So you’ll read someone working on the neurophysiology of consciousness – a hot topic – and there will be an obligatory first couple of paragraphs saying ‘well, the philosophers have mucked this up for centuries, but now we’ll show how it’s done’ but with no engagement in the arguments that have been given or the thinking behind them. (Chomsky 2011: 174–175)

Here Chomsky gives the impression that he paraphrases something written by a scientist without providing the name of the author or a publication in which the alleged statement was made. This rhetorical trick turns what one individual might have said into a blanket accusation that scientists hold philosophers in contempt. Ignoring the irresponsible nature of this remark, interviewer Peter Ludlow replies: ‘Does it trouble you that, say, most linguists don’t engage philosophers on these issues’ (Ludlow 2011: 75). Apparently neither of the two is troubled by the message this passage might send to the scientific community.
8. Conclusions

Pesetsky (2013) claims that the scientific community is largely ignorant of work by generative linguists. Given the problems I have discussed in this review article, one almost hopes he is partly right, and that not too many scientists have read Chomsky’s recent work. But since Chomsky is the best-known linguist, one must assume that non-linguists in particular will read his books before turning to other linguistic publications. Few scientists who have read any of the works I have quoted will fail to notice their dubious quality, and they are bound to wonder why many generative linguists apparently still hold such work in the highest esteem. And even if The Science of Language does not discourage readers from going on to works by other linguists, they would find barely any reference to recent linguistic work in the bibliography. The only purpose I can see that The Science of Language could serve is as a reminder for linguists that theorizing and philosophizing about a science like biology cannot be based on what ‘sounds reasonable’, but must be grounded in thorough knowledge of biology. ‘Do not try to bring science down to ordinary knowledge but rather learn some science before philosophizing about it’ (Bunge 2002: 4).

REFERENCES


30


JOURNAL OF LINGUISTICS


Author’s address: Department of Philosophy, Dalhousie University, 6135 University Avenue, Halifax, Nova Scotia, B3H 4R2, Canada
Christina.Behme@dal.ca