Chomsky's "Galilean" Explanatory Style

Nicholas Allott, Terje Lohndal & Georges Rey

1. Introduction[1]

Chomsky pursues a methodology in linguistics that abstracts from substantial amounts of data about actual language use in a way that has met considerable resistance from many other linguists. He thinks of this method as like that employed by Galileo and later physicists who proposed laws of motion in considerable abstraction from many of the motions we observe in daily life, focusing, for example, not on leaves in the wind, but on frictionless environments that virtually never occur on earth. Thus, Chomsky's theoretical proposals are supported not by studies of the corpora of actual language use, but often by the intuitions of native speakers; and the relevant intuitions are not about what they think is *often* or is *likely* to be said, but rather about what "*can't"* be said (so called "negative data"), and about what types of interpretation a sentence can or cannot have. But doesn't this fly in the face of good, commonsensical scientific methodology? Aren't theories confirmed by greater data, and refuted by data that seem to conflict with them? With regard to this issue, Chomsky (1980) writes:

> Substantial coverage of data is not a particularly significant result, it can be attained in many ways, and the result is not very informative as to the correctness of the principles employed. It will be more significant if we show that certain far-reaching principles interact to provide an explanation for crucial facts – the crucial nature of these facts deriving from their relation to proposed explanatory theories. (Chomsky 1980, 2)

We'll argue below that Chomsky's observation here in fact accords with good explanatory practice elsewhere in science, but it does conflict with a traditional methodology in linguistics. In the spirit of the positivism/empiricism of the 1930s, the 'structuralist' linguist Leonard Bloomfield (1933, 20) insisted that "the only useful generalizations about language are inductive generalizations", and linguists for the next several decades tried to specify 'discovery procedures', or rules for using a collection of phonetically characterized utterances to induce phonemic, morphemic and – it was hoped – finally syntactic analyses of the target language (see Sampson, 1980, 76ff). Such discovery procedures have fallen by the wayside, but many contemporary linguists would still agree with Bloomfield that linguistics seeks generalizations that both emerge from, and provide good coverage of, the data of language use.[2]

Those who have made such arguments tend to regard Chomsky's Galilean style as manifestly, even outrageously, unscientific. Chater et al. (2015) claim it "allows a serious researcher to ignore data which is incompatible with his theory":

> There is a very troubling and disturbing problem we encounter as soon as we undertake to ignore data – does it need to be spelled out? The problem with this methodology is this: that each of us…is, perhaps, confident to identify true conjectures which do not appear to be supported by the data, but we are not always so confident about the next person's. And who knows? Someone

[1] Some of the material in this chapter draws on Allott (2019) and Rey (2020).
[2] Such empiricist leanings are prevalent among corpus linguists and usage-based linguists (regarding whom see Newmeyer this volume) as well as those who explicitly describe themselves as empiricists (e.g. Chater et al. 2015).

> else might feel exactly the same way about it, only in reverse. *That's* the problem. (Chater et al. 2015, 95 – their italics)

Indeed, "there is no philosophy of science that allows one to ignore data" (Chater et al. 2015, 96). Similarly, Knight (2016, 171) claims that "[w]hen the facts contradict a theory, most scientists would agree that the theory needs to go. But not Chomsky." And Seuren (2004, 29) claims that Chomsky, following the Galilean style, is "free to ignore any kind of contrary or refractory evidence to uphold his theory […] One wonders what would become of science if everyone took similar liberties."

In our view, such criticisms are founded on misunderstanding – not so much of Chomsky's position, although there is a good deal of that – as of good scientific method. It's striking that the main features of Chomsky's Galilean style are independently taken to be rather obvious features of scientific method in contemporary philosophy of science. It shouldn't be surprising, for example, that generative syntacticians make use of data that are only collected as a result of the theory under consideration: the same is true of other sciences – from the particle accelerators which allow physicists to produce interactions to order, to the artificial stimuli and controlled conditions of experiments in cognitive psychology.[3] And generativists, like other scientists, are interested in data that bear on their theoretical claims. So quite rationally, scientists discount observations that do not bear on the system under investigation because they hypothesize that they are likely caused by other systems.

In what follows we first look in more detail at the roles that idealization plays in contemporary generative grammar (section 2), and then proceed to explore the congruence between the Galilean style in linguistics and aspects of good scientific methodology (section 3). Finally, we set out and respond briefly to some of the arguments that have been raised against the Galilean style (section 4).

2. Idealization in the sciences

There are several claims at the heart of Chomsky's Galilean style.[4] These include i) the conception of science as a quest for underlying explanatory structure; ii) the distinction between superficial generalizations and deeper explanatory principles; and iii), the claim that the crucial relationship between data and theory is that whatever data is cited should support inferences about the phenomena under investigation, and the hypotheses about other, interfering causes are correct (see Pietroski and Rey 1995 for how this is a substantive commitment), not that the theory should cover *all* the data in a domain

2.1. Science as quest for underlying structure

> The great successes of the modern natural sciences can be attributed to the pursuit of explanatory depth which is very frequently taken to outweigh empirical inadequacies. This is the real intellectual

[3] Critics sometimes point out that that Galileo was more of an experimentalist than was thought by Koyré, whose seminal work from the 1950s seems to have influenced Chomsky on this point (see Riemer 2009, 627-8, and Sharrett 1994, 75-9). Whatever the historians of science decide, there is obviously no argument here against Chomsky's methodology in linguistics. In any case, generativists don't oppose experimental work. Indeed psycho- and neurolinguistic work relevant to generative grammar has mushroomed in recent decades, largely driven by the advances made by theoretical syntacticians. (See Kush and Dillon this volume; Zaccarella and Trettenbein this volume).

[4] On the Galilean style, see Galileo (1610/2016); Koyré (1943); Weinberg (1976); Chomsky (1978a, 9–10); Chomsky (1978b, 14); Chomsky (1980, 218); Chomsky (2002, 98–100).

revolution of the seventeenth century" (Chomsky 1978a, 9–10)

A scientist aiming to frame deep explanatory principles is not well served merely by collecting surface data, since their relation to underlying structure is typically quite indirect. Most of the phenomena we ordinarily observe are the result of interaction effects involving a huge variety of underlying systems, perhaps including interacting masses, frictional planes, biological creatures and perceptual processes. Any serious science must study each system largely in highly idealized isolation from the others (hence the 'different disciplines' of science and the effort and expense of 'controlled experiments').

Indeed, many of the crucial data for generative grammar are not to be found in everyday language use or corpora. Some are negative data, that is data about sentences that people seldom or never produce, since they are not ones that are generated by their language faculty. Others structures that are infrequently used, such as a passive interacting with clausal embedding (see Baker this volume, on such data). Generative research often relies on data that are only discovered in connection with the theories it is considering. Focusing only on spontaneously produced utterances that have been pre-theoretically collected would significantly handicap inquiry.

This focus on crucial data is connected to Chomsky's most famous, but only initial idealization: his (1965, 4) distinction between 'competence' and 'performance.' [5] In contrast to the Bloomfieldian concern with the actual use, or *performance* of language, Chomsky is interested in the underlying grammatical *competence* system, what he (1986) came to call the 'I-language',[6] and with what he argues is the innate system that allows an I-language to develop, Universal Grammar. On his view, the I-language underlies linguistic parsing and production, but its principles may not be clearly manifested in particular instances of language use, which, again, involves the interaction of many factors: not only the parser or sentence production processor, but also memory and pragmatic inference systems, and, in production, the speaker's preferences, articulatory abilities, speech pragmatics and so forth. The involvement of this multitude of diverse factors is why Chomsky doubts performance is a fruitful domain for general theoretical insight:

> [A] formalized grammar, regarded as a predictive theory, is an idealization in at least two respects: first, in that it considers formal structure independently of use; and second, in that the items that it generates will not be the utterances of which actual discourse is composed. Actual discourse consists of interrupted fragments, false starts, laps, slurring, and other phenomena that can only be understood as distortions of an underlying idealized pattern.
> (Chomsky 1962, 531)

On this view, the use of language, like the trajectories of leaves or automobiles, is a massive interaction effect, far too complex to hope to cover with a single theory.

There are actually several claims historically associated with the Galilean method. We've touched so far only upon the methodological one concerning what scientists should seek. Galileo also thought that mathematics is 'the language of nature' which scientists often assume to be mathematically elegant (see Koyré 1943, 336ff; Weinberg 1976). Although Chomsky (1983/2004, 154ff) also takes

---

[5] Ferdinand de Saussure (1914/77) made a similar distinction between 'langue' and 'parole' (cf. Chomsky 1968/2006, 17-18; 1986, 32ff; see also Newmeyer 1986, 71, for differences between Saussure's and Chomsky's notions).
[6] Chomsky's original term "competence" led to confusion between mere behavioral dispositions and the underlying system responsible for those dispositions. Note that an I-language is not an idiolect, but *a computational system*, i.e., not the sort of thing one might speak. On the I-language as a mental system, see Allott and Smith (this volume).

this view seriously, it is secondary and independent of the point about idealization, which could of course be to mathematically inelegant systems. We'll return to it in section 3.[7]

2.2. Superficial generalizations and explanatory principles

Not every general truth is explanatory. Chomsky distinguishes between generalizations over data and explanatory principles and argues that the point of idealization is not to arrive at the former but to discover the latter (Chomsky 1978, 23). Generalizations are useful, not because they offer insight into phenomena, but because they describe and make clearer the phenomena that we should try to explain in terms of deeper rules or principles.

For example, the philosopher of science, Nancy Cartwright (1983) notes a distinction between two sets of laws that both accurately capture an effect discovered by Faraday. The physicist, Francis Everitt,

> "distinguishes Airy's phenomenological [i.e. merely descriptively accurate] law from the later theoretical treatment of Lorentz, […] because the electron theory explains the magneto-optical effect and Airy's does not. Phenomenological laws describe what happens. […] For the physicist […] the distinction between theoretical and phenomenological […] separate[s] laws which are fundamental and explanatory from those that merely describe." (Cartwright 1983, 1–2)[8]

A good example of the same issue in linguistics is afforded by 'island constraints'. In work by Ross and others in the 1960s and afterwards, it was discovered that there are a number of kinds of phrase from which constituents cannot be extracted. Compare (1b) and (1d) below. On generative accounts, (1d) is ungrammatical because *what* cannot be extracted from the clause *who saw what*. This is an instance of the generalization that *wh*-clauses resist extraction, but this is not a deep explanation. One wants to know why such phrases do not allow extraction, preferably in terms of a theory that also explains why various other kinds of constituents (e.g. subjects and adjuncts) are also islands.

(1) a. Anne knows [that Bill saw the getaway car].
b. What does Anne know [that Bill saw ___]?
c. Anne knows [who saw the getaway car].
d. *What does Anne know [who saw ___]?

Syntacticians, foremost among them Chomsky (1973), have proposed grammatical principles such as 'subjacency' to explain island phenomena. This is not the place to review the proposals (see Müller, this volume). The point here is rather that the principles proposed are all constraints on the kinds of computations (i.e. manipulations of representations) that the I-language is capable of, and they try to explain islands of apparently diverse types by isolating properties they share and showing that those properties block certain kinds of (movement) operations in the grammar.

[7] There is, of course, a difficult question in philosophy of science about the ontological status of the generalizations and laws reached by abstraction, but there is no reason to think this question is more problematic for linguistics than for physics, chemistry or biology. Views about the status of idealizations range from Nancy Cartwright's (1983) famous claim that they are false, albeit often useful and interesting, to Chomsky's view that "it is the abstract systems that you are constructing that are really the truth; the array of phenomena is some distortion of the truth because of too many factors" (Chomsky 2002, 99; see Frigg and Hartmann (2020) for an excellent overview of the issues). Weinberg (1976, 28) claims that physicists "give a higher degree of reality [to] abstract mathematical models of the universe […] than they accord the ordinary world of sensation," but again, such metaphysical claims are independent of the methodological one we are discussing here.

[8] Of course, what is explanatorily deep needn't be only fundamental physics; cf. fn. ?? below.

2.3. Relation between data and theory

As noted above, if explanatory principles rather than data coverage are the aim, then the data that matter will be those that help to distinguish between theories. Here, too Chomsky's methodology is in accord with contemporary philosophy of science. As the eminent philosopher of science James Woodward puts it, expanding on Bogen and Woodward's classic (1988):

> What matters is not that we be able to infer 'downward' to features of the data from theory and other assumptions; rather, what matters is whether we are able to infer 'upward' from the data (and other assumptions) to the phenomenon of interest. Data are scientifically useful and interesting insofar as they provide information about features of phenomena. (Woodward 2011, 168)

To take a famous example, the astrophysicist Arthur Eddington photographed stars visually near the sun during an eclipse and

> infer[red] *from* features of his photographs (and various background assumptions) *to* a value for the deflection of starlight [by gravity] rather than trying to infer or derive characteristics of the photographs from other assumptions. (Woodward 2011, 168 – his italics)

That is, Eddington did not aim to explain all of the data provided by his photographs. He was concerned with the crucial data that were predicted by Einstein's theory but not Newton's. As the philosopher Jerry Fodor aptly put it:

> What goes on in science is not that we try to have theories that accommodate our experiences; it's closer that we try to have experiences that adjudicate among our theories (Fodor 1991, 202–3)

Thus, a generative syntactician is not primarily concerned to explain all – or even a significant portion – of language use; rather, just crucial data that are predictable from her but not rival theories. Of course, once deep explanatory principles are understood, they can be used to shed light on 'surface' phenomena, typically in conjunction with other theories. The paths of ordinary objects can in principle be predicted using Newton's laws in conjunction with assumptions about friction, wind speed and so on (although the calculations required are often prohibitively complex). Similarly, the growing understanding of I-language predicts certain facts about usage fairly directly – for example, aside from papers on syntax, examples of certain island violations are seldom, if ever encountered – but they invite the investigation of more complex aspects of usage, e.g. the study of how language is processed in real time (see Kush and Dillon this volume).

A worry many people have voiced is that generativists make a great deal of use of speakers' intuitions, and using such possibly unreliable intuitive verdicts as evidence seems a dubious methodology. Speaker verdicts are however not being taken to afford any special *insights* into grammar. Rather, they are just *data*: readily observable facts about human reactions that, like any other data, might help to distinguish between theories. One could also note differences in reaction times or pupillary dilation: anything that would betray a distinctive reaction. In this respect, they are like subjects' reactions to illusions in vision experiments. Of course, a confound that can occur is the influence of a linguist's commitment to his or her theory, or a hearer's acquiescence to a linguist's claims. This is one reason why linguists routinely ask for the (repeated) verdicts of "naive" *native* speakers of a language. Note that all such data could be subjected to controlled experiment, e.g. using Likert scale questionnaires. Recent studies have demonstrated high reliability between linguists and large sets of experimentally controlled subjects, particularly ones, such as beginning cognitive science students, who have some conception of what is being asked but are not yet in thrall to any particular theory (see Gross this volume, for discussion and references).

3. How the Galilean style works in generative linguistics

The Galilean style is manifested in linguistic research in two distinct ways. Per our discussion so far, it can be taken as primarily methodological in import; but, as we mentioned, it can also be taken as involving two ontological claims that are Galilean in flavor: that nature is mathematically organized, and that the underlying principles of nature are in some sense elegant or simple. The second of these claims in particular has been pursued in recent generative linguistics, but, again, it is independent of the methodological point: mere methodological idealization provides no guarantee that the underlying principles or system which it helps to lay bare will turn out to be particularly simple or elegant, although it does presume that they are simpler than the mass of surface phenomena.[9]

Strong claims of ontological simplicity seem to be what Boeckx (2006, 115) has in mind when he labels the Minimalist Program in generative syntax 'Galilean' (on the Minimalist Program, see Alexiadou and Lohndal this volume). They also seem to be behind what Chomsky (1995, 2000, 2001) calls the 'Strong Minimalist Thesis', which holds that language is an optimal solution to the task of relating sound and meaning, that is, to satisfying demands imposed by the interfaces. This is the claim that the operations of the language faculty are as efficient as they could possibly be, given the role it plays in cognition, a proposal reminiscent of claims of mathematical elegance in physics. However not all Minimalists share this assumption. One leading theorist, Norbert Hornstein (2013), suggests that the goal of Minimalism is, or should be, rather to recast theories of syntax so that they postulate as little as possible that is proprietary to the language faculty. This version of Minimalism eschews strong ontological simplicity claims regarding the nature of the computational system, but is entirely compatible with the Galilean style as a methodology.

As we have discussed above, generative linguistics has been methodologically Galilean from Chomsky's early work onwards. The primary example of this is what we referred to in section 1 as Chomsky's most famous idealization, namely the competence-performance distinction. It's important to see, though, that idealization goes well beyond this distinction in guiding the work of linguists. In studying, say, a syntactic property, it is crucial to attempt to isolate this property not only from other aspects of cognition but also as far as possible from other parts of the linguistic system, such as lexical, semantic, phonological, morphological and discourse properties, and even from other syntactic properties. Typically, generative syntacticians attempt to work with examples that hold all these other properties constant while putatively varying just in whatever syntactic property is at issue.

Other theorists may want to abstract and idealize in different ways for different purposes. Everything depends on how theoretically fruitful a particular idealization proves. In this respect, Chomsky's

[9] A related issue concerns the relationship between unification of theories and reduction to a 'more fundamental' theory (e.g. neuroscience), which Boeckx (2006, 131-133) discusses with reference to physics and biology. Reduction is one type of unification, but there are other types of unification that do not entail reduction, and the Galilean style is not committed to either.

proposals have fared extremely well: several classes of previously unsuspected phenomena have been uncovered, including syntactic islands as discussed above, but also Binding, Control and Raising phenomena (for discussion see Hornstein 2013; D'Alessandro 2019). And, of course, his approach offers the beginning of an understanding of the extraordinary speed and stability with which children attain grammatical competence.

4. Criticisms of the Galilean style in linguistics

> To tell an Aristotelian not to bother about friction or air-resistance was like telling him to abandon the real world we all inhabit in favour of a mathematical fantasy. (Sharratt 1994, 74)

Most critics of Galilean idealisations have rather strong empiricist and externalist leanings. In theory they may agree that the objective of linguistics is to look for hidden internal structures that give rise to observable phenomena, yet they strongly prefer coverage of surface facts about language(s) and language use – echoing Carnap, Hahn and Neurath's "there are no depths; there is surface everywhere" (1973, 306) [10]. Many also take the object of study to be not I-language and the language faculty, but external languages as shared social objects that are partially learned by each speaker.

The stress on the importance of surface facts sometimes amounts to a denial of the assumption that what we observe is determined by the grammar interacting with other factors: e.g., when Riemer writes (2009, 622): "A theory which often predicts the ungrammaticality of clearly acceptable strings, or often predicts the grammaticality of unacceptable ones, is in trouble." This claim rests on the assumption that acceptability is (very largely) determined by the grammar: i.e. that interpretation is (nearly) free of interaction effects – a very dubious assumption, given what is known about parsing and pragmatics. A different way to make the same point is to note that strictly speaking, theories of grammar by themselves do not predict facts about acceptability. Only in conjunction with various auxiliary assumptions about parsing and so on do predictions of acceptability emerge.

In this section, we will discuss some of the criticisms of the Galilean style in linguistics, beginning with two core claims made by critics: (i) good science does not ignore data, (ii) the drive for abstraction in generative research has lost sight of the huge linguistic diversity that is evident in the languages of the world.

Let us deal with the second issue first. This claim is often made, e.g., by Evans & Levinson (2009) and Haspelmath (2007). However, as Baker (this volume) makes clear, not only has there been a great deal of work on non-English-like languages in the history of generative linguistics, but this comparative work has been crucial for progress in the field, particularly since the emergence of the principles and parameters program in the 1970s. And conversely, the success of generativist syntax here makes at least a *prima facie* case for the need for abstraction in comparative work: that is, it seems that the Galilean approach has enabled scholars to probe the grammatical architecture of diverse languages in ways that would have been impossible without idealizing (Sheehan this volume provides several examples). Idealization does not diminish the significance of linguistic diversity; rather, it enables us to study it using ordinary scientific methodology.

[10] Writing about science, of course. The idea that "nothing is hidden" in the mind has been a recurrent theme in work of various kinds of behaviorists, from Quine and Skinner to Wittgenstein and Ryle, and, even more recently, in non-non-behaviorists such as Dennett (1991, 461ff.) and Chater (2018).

The other major criticism is that generativists ignore data. We suspect that those making this complaint assume that the domain of data to be accounted for by linguistic theory is something given, which ought to be common ground among linguists. But as we've explained above, the data that matter are those that help us accept or reject a theory or distinguish between theories. Thus, depending on what your research questions are, the data of interest will be substantially different. To return to our example, if a scholar is interested in studying the grammatical constraints on argument *wh*-movement in a given language, it is probably not especially relevant to study frequencies of occurrence (say, of different types of arguments) in a given corpus. So even though such data could be interesting with a different research question in mind, grammatical constraints on *wh*-dependencies can (it seems) be studied without any need to look at them. The same goes for other kinds of data, such as how these *wh*-dependencies are used in discourse. That is an interesting question, but (as far we know) not one that generative syntacticians need presently be interested in since such usage data do not currently look likely to shed light on the structure of our mental grammar. Obviously none of this should prevent others from using whatever frequency or usage data bear on their theoretical claims, and almost any data *could* be relevant to any theory (if it turns out that they do bear on a claim it makes),[11] but the mere fact that some generativists do not use particular kinds of data does not warrant accusations of data ignorance. To establish that, the critics would have to show that *data that are relevant to claims made by generativists* are being ignored (deliberately or negligently), and as far we are aware, they have not done so.

Another complaint is that generativists ignore computational parts of linguistics that the critics regard as having good data coverage, including Artificial Intelligence (AI) work with corpora which is dismissed as "engineering solutions" (Chomsky 1998). It is true that most generative syntacticians have not taken such work to be very useful for their research, but with good reason. We've explained above why they are not in the business of explaining the diverse performance phenomena that a corpus captures. There's a second reason why the AI work strikes them as of limited relevance. AI might solve similar problems to human intelligence, but it generally does it differently, just as airplanes and birds do not fly in the same way. Contemporary 'deep learning' algorithms appear to acquire linguistic generalizations very differently from humans, needing to be trained on vast input corpora. (See Adger 2019, ch. 8 for an accessible summary of the evidence.)

One final objection is of broader scope. This is the claim that linguistics is not a science, or at the least not a hard science, and that the quest for physics-like exceptionless laws is quixotic, leading necessarily to distortion of the phenomena since they are intrinsically less hard-edged or clear-cut. (Riemer 2009, 630; Evans and Levinson 2009.) It sometimes seems that those taking this line simply haven't understood the idealizations that are being made in generative linguistics (see for example Harbour 2009, on Evans and Levinson's confused claims about number in Kiowa). Be that as it may, the arguments for this view, such as they are, tend to beg the question against idealization in linguistics. The methodology will be judged by how fruitful it proves to be. This paper is not the place to discuss in detail the successes of Chomsky's program (among many chapters, see Cheng and Griffiths this volume; Hornstein 2013; D'Alessandro 2019) but it is obviously relevant that there are no non-generative explanations for many of the crucial data and generalizations of modern linguistics.

References

Adger, D. 2019. *Language Unlimited: The Science Behind Our Most Creative Power*. Oxford: Oxford University Press.

[11] See Antony (2003).

D'Alessandro, R. 2019. "The achievements of Generative Syntax: a time chart and some reflections". *Catalan Journal of Linguistics Special Issue* 7-26.
Allott, N. (2019). "Scientific tractability and relevance theory". In K. Scott, R. Carston & B. Clark (Eds.), *Relevance: Pragmatics and Interpretation* (pp. 29–41). Cambridge: Cambridge University Press.
Antony, L. 2003. "Rabbit-pots and supernovas: On the relevance of psychological data to linguistic theory". In *Epistemology of Language*, edited by A. Barber, pp. 47–68. Oxford: Oxford University Press.
Bloomfield, L. 1933. *Language*. New York: Henry Holt.
Boeckx, C. 2006. *Linguistic Minimalism: Origins, Concepts, Methods, and Aims*. Oxford: Oxford University Press.
Boeckx, Cedric. 2010. *Language in Cognition*. Wiley-Blackwell: Malden.
Bogen, J., and J. Woodward. 1988. "Saving the phenomena". *The Philosophical Review* 97(3): 303–352.
Botha, R. P. 1982. "On 'the Galilean style' of linguistic inquiry". *Lingua* 58(1): 1-50.
Carnap, R., H. Hahn, and O. Neurath. 1973. "The scientific conception of the world: The Vienna Circle". In *Empiricism and Sociology*, edited by M. Neurath and R. S. Cohen, 299-316 Dordrecht: Reidel.
Cartwright, N. 1983. *How the Laws of Physics Lie*. Oxford: Clarendon Press.
Cartwright, N. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.
Chater, N. 2018. *The Mind is Flat: The Illusion of Depth and the Improvised Mind,* New York: Penguin Random House
Chater, N., A. Clark, J. A. Goldsmith, and A. Perfors. 2015. *Empiricism and Language Learnability*. Oxford: Oxford University Press.
Chomsky, N. 1962. "Explanatory Models in Linguistics". In *Logic, Methodology and Philosophy of Science*, edited by E. Nagel, P. Suppes, & A. Tarski, 528-550. Stanford, CA: Stanford University Press.
Chomsky, N. 1978a. "A theory of core grammar". *Glot* 1: 7-26.
Chomsky, N. 1978b. "Interview with Sol Saporta". *Washington State University, Dept. of Linguistics Working Papers in Linguistics, supplement* 4: 1-26.
Chomsky, N. 1980. *Rules and Representations*. New York: Columbia University Press.
Chomsky, N. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.
Chomsky, N. 1995. Language and nature. *Mind* 104(413): 1-61.
Chomsky, N. 2000. "Linguistics and brain science". In *Image, language, brain: papers from the First Mind Articulation Project Symposium*, edited by A. Marantz, Y. Miyashita, and W. A. O'Neil, 13-28. Cambridge, MA: MIT Press.
Chomsky, N. 2002. *On Nature and Language*. Cambridge: Cambridge University Press.
Chomsky, N., and J. Andor. 2005. "The master and his performance: An interview with Noam Chomsky". *Intercultural Pragmatics* 1(1): 93-111.
Chomsky, N., R. Huybregts, and H. C. V. Riemsdijk. 1982. *Noam Chomsky on the Generative Enterprise: A Discussion With Riny Huybregts and Henk Van Riemsdijk*. Dordrecht, Holland; Cinnaminson, N.J: Foris Publications.
Chomsky, N. 1980. *Rules and Representations*. Columbia: Columbia University Press.
Dehaene, S. 2007. "A few steps toward a science of mental life". *Mind, Brain, and Education* 1: 28-47.
Dennett, D. 1991. *Consciousness Explained*. Boston: Little Brown & Co.
Dupré, J. 1993. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge, MA: Harvard University Press.

Evans, N., and S. C. Levinson. 2009. "The myth of language universals: language diversity and its importance for cognitive science". *Behavioral and Brain Sciences* 32(5): 429-448.
Fodor, J. A. 1991. "The dogma that didn't bark (a fragment of a naturalized epistemology)". *Mind* 100(2): 201–220.
Freidin, R., and J-R. Vergnaud, J-R. (2001). "Exquisite connections: some remarks on the evolution of linguistic theory". *Lingua* 111: 639-666.
Frigg, R. & Hartmann, S. 2020. "Models in science". *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/models-science/
Galilei, G. 1610/2016. *Sidereus Nuncius, or The Sidereal Messenger*. Translated with Commentary by Albert van Helden. Chicago: University of Chicago Press.
Godfrey-Smith, P. 2003. *Theory and Reality: An Introduction to the Philosophy of Science*. Chicago: University of Chicago Press.
Harbour, D. 2009. "The universal basis of local linguistic exceptionality". *Behavioral and Brain Sciences* 32(05): 456–457.
Haspelmath, M. 2007. "Pre-established categories don't exist: Consequences for language description and typology". *Linguistic Typology* 11(1): 119-132.
Hornstein, N. 2013. "Three grades of grammatical involvement: Syntax from a minimalist perspective". *Mind & Language* 28(4): 392–420.
Knight, C. 2016. *Decoding Chomsky: Science and Revolutionary Politics*. New Haven and London: Yale University Press.
Koyré, A. 1943. "Galileo and the scientific revolution of the seventeenth century". *The Philosophical Review* 52(4): 333-348.
Losee, J. 1980. *A Historical Introduction to the Philosophy of Science* (2nd ed.). Oxford: Oxford University Press.
McMullin, E. 1985. "Galilean idealization". *Studies in History and Philosophy of Science Part A* 16(3): 247-273.
Pietroski, P., and G. Rey. 1995. "When Other Things Aren't Equal: Saving Ceteris Paribus". *British Journal for the Philosophy of Science* 46: 81–110.
Sampson, G. 1980. *Schools of Linguistics: Competition and Evolution*. London: Hutcheson.
Seuren, P. A. M. 2004. *Chomsky's Minimalism*. Oxford: Oxford University Press.
Riemer, N. 2009. "Grammaticality as evidence and as prediction in a Galilean linguistics". *Language Sciences* 31(5): 612-633.
Rey, G. 2020. *Representation of Language: Philosophical Issues in a Chomskyan Linguistics*. Oxford: Oxford University Press.
Sharratt, M. 1994. *Galileo: Decisive Innovator*. Oxford: Blackwell.
Uriagereka, Juan. 1998. *Rhyme and reason*. Cambridge, MA: MIT Press.
Weinberg, S. 1976. "The forces of nature". *Bulletin of the American Academy of Arts and Sciences* 29(4): 13-29.
Weinberg, S. 1998. "The revolution that didn't happen". *The New York Review of Books*, *October 8, 1998*. Retrieved from https://www.physics.utah.edu/~detar/phys4910/readings/fundamentals/weinberg.html
Weinberg, S. 1996. "Sokal's hoax". *The New York Review of Books*, *August 8, 1996*.
Woodward, J. 2011. "Data and phenomena: a restatement and defense". *Synthese* 182: 165–79.