Language Sciences 100 (2023) 101585

Contents lists available at ScienceDirect

Language Sciences

journal homepage: www.elsevier.com/locate/langsci

Generative linguistics: ‘Galilean style’

John Collins

School of Politics, Philosophy, Language and Communication, University of East Anglia, Norwich, NR4 7TJ, UK

ARTICLE INFO

Article history:
Received 20 July 2022
Received in revised form 24 July 2023
Accepted 23 August 2023
Available online xxx

Keywords:
Chomsky
Centre-embedding
Competence/performance
Computation
Galilean style
Galileo

ABSTRACT

Generative linguistics is often claimed by Chomsky to have a ’Galilean style’, which is intended to position linguistics as a science continuous with standard practise in the natural sciences. These claims, however, are more suggestive than explanatory. The paper will, first, explain just what a Galilean style is. It will then be argued that its application to two key notions in generative linguistics - the competence/performance distinction (with reference to centre-embedding) and the notion of computation - demands a departure from what we might expect of a Galilean style. In this sense, the epithet is misleading. It will also be shown, however, that the ’Galilean’ label is appropriate once we factor in the difference between a science concerned with kinematics (the relations between objects in space and time) and one concerned with language.

1. Introduction

Chomsky often speaks of generative linguistics as possessing a ‘Galilean style’ (e.g., Chomsky, 1980, 2002). The epithet, due to Husserl (1936/70), is intended to position linguistics as a genuine scientific endeavour that targets an aspect of human cognition in the same manner as undisputed successful science—post-Galilean physics—targets its phenomena. Generative linguistics is understood to be a branch of cognitive science, and so, ultimately, a branch of human biology. Its topic is the acquisition and maintenance of language as a species property. Its chief hypothesis is that such phenomena are subserved by a *language faculty*, a computational system that, abstractly specified, realises a function or procedure that generates structures (syntax) that encode the properties that allow a speaker hearer to pair sign (phonology/hand gestures) with meaning.

Chomsky’s invocation of Galileo has attracted by both critical (Botha, 1982; Behme, 2014) and positive (Rey, 2020; Allott et al., 2021) responses. The emphasis of this work has been on what Botha refers to as Galileo’s putative ‘epistemological tolerance’, the idea that theories ought to be allowed a certain slack when it comes tolerating contradictory data or lack of coverage of data (cf., Riemer, 2009; Kertész, 2012). In distinction, my focus is on the central Galilean theme of explanation being grounded in an abstract mathematically specified system to which observable phenomena only approximate. To be sure, this results in a certain ‘epistemological tolerance’, but that is a consequence of a profounder idea concerning the nature of proper explanation. This emphasis is aligned with Weinberg’s (1976) interest, whom Chomsky (1980, pp. 8–9) quotes when first invoking Galileo, and is the chief concern of Husserl (1936/70), who first coined the epithet of ‘Galilean style’. As we shall see, it is also central to Galileo’s actual approach. This appeal to the abstract, however, raises some deep interpretive problems for central commitments of generative linguistics: the competence/performance distinction and the computational view of

E-mail address: John.collins@uea.ac.uk.

https://doi.org/10.1016/j.langsci.2023.101585
0388-0001/

the language faculty. The past emphasis on 'epistemological tolerance', therefore, is not exactly a red herring, but does elide what, I intend to show, is the more central and far-reaching core proposal of the 'Galilean style'.[1]

More specifically, the paper's aim is twofold. First, the notion of 'Galilean style' will be explained. We shall see that Chomsky's simple invocation of the style masks a certain complexity in regard to what we take the topic of generative linguistics to be, *viz*. something 'real' or something 'ideal'. Secondly, the sense in which generative linguistics has a 'Galilean style' will be interrogated with respect to the two central ideas just above. I shall focus here on two foundational moves from the opening chapter of *Aspects* (Chomsky, 1965): the competence/performance distinction and the computational conception of competence. The basic problem in the first case is that if we abstract away from contingent factors, in a Galilean fashion, to have as our topic an 'ideal speaker-hearer', then a grammar would appear to be an idealised performance system, but this is a conception Chomsky explicitly rejects. The natural move is thus to view the necessary abstraction as not to an ideal speaker-hearer, but to a specific component of the mind/brain viewed in isolation. This view was only to be become explicit in Chomsky (1986), but is already present in *Aspects* in the form of the competence/performance distinction. The problem now arises, however, of the 'psychological reality' of the hypothesised component from the perspective of the 'Galilean style'. I shall suggest, in a Galilean manner, that there is a natural way to understand a grammar to be both a real mental component *and* abstract.

2. Chomsky's appeal to Galileo

Chomsky's initial appeal to Galileo is via Clavelin, (1968/74) and Weinberg (1976). Chomsky (1980, pp. 8–9) quotes the latter as follows:

> The Galilean style … [amounts to] making abstract mathematical models of the universe to which at least the physicists give a higher degree of reality than they accord the ordinary world of sensation (Weinberg, 1976, p. 28).

Clavelin argues in the same direction, claiming that Galileo initiated a new standard of 'scientific intelligibility' that prioritises constants abstracted from observational phenomena at the expense of full coverage of data. It is here where 'epistemological tolerance' arises, but only as a consequence of the initial move to think of the explanatory goal as directed towards delineating an underlying abstract system. As Chomsky (2002, pp. 98–9) elaborates:

> What was striking about Galileo … was that he dismissed a lot of data … The Galilean style … is the recognition that it is the abstract systems you are constructing that are really the truth; the array of phenomena is some distortion … [due to] too many factors … And so it makes sense to disregard phenomena and search for principles that really seem to give some deep insight into why some of them are that way, recognizing that there are others you can't pay attention to.

Weinberg's appeal to 'greater degree of reality' suggests the tradition in Galileo scholarship that depicts him as a Platonist (Koyré, 1938/78), and Chomsky's latter elaboration of Galileo's apparently blithe attitude (at best) to evidence is reminiscent of Feyerabend (1975), which has been taken up by Botha (1982) and Behme (2014). Chomsky's invocation of Galileo, however, is not principally animated by an attitude towards data (Botha's 'epistemological tolerance'), but the view that a certain abstract conception of what explains the phenomena is and ought to be given more importance than the phenomena themselves, and the conception will shape decisions concerning what phenomena are central and what phenomena are peripheral to the explanatory project.[2] In short, the phenomena to be explained are partly *selected* in terms of being those that crucially support the hypothesised abstract system rather than satisfying any other criteria.[3]

For orientation, we may consider three examples of how linguistic theory is not designed to fit all and any phenomena (more complex phenomena will be considered in §5). First, consider ordinary conversational discourse. From an intuitive perspective, this might seem to be just what language amounts to. Yet such use of language is characteristically messy, strewn

[1] Botha (1982) does give due weight to the abstractness of the 'Galilean style', but gives greater focus to 'epistemological tolerance' as a putative feature of the style in linguistics.

[2] It is now clear that Galileo did conduct experiments and was not the least bit blithe about evidence, with the proviso that it was by no means straightforward to design replicable paradigms for the kind of phenomena that concerned Galileo (for historical discussion, see Wootton, 2010; Hellbron, 2012).

[3] It is an error, therefore, to think of the 'Galilean style' as some kind of licence to dismiss data or evidence. Behme (2014, p. 687) writes:

> Chomsky has moved further and further toward 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).

Behme (op cit., p. 687, n. 19) suggests 'that Chomsky's 'Galilean style' is based not on Galileo's work but on Feyerabend's misinterpretation of that work'. While I think Behme is correct about Feyerbend, it is a mistake to saddle Chomsky with Feyerbend's views. Chomsky's first invocation of 'Galilean style' is via Weinberg and Clavelin, both of whom stress Galileo's prioritising of mathematically specified relations over coverage of data, which is central to Husserl's (1936/70) conception of Galileo, who first coined the epithet, not Chomsky (see §3). Thus, Chomsky is right, following many others, beginning with Husserl and Koyré in the 1930s, that the Galilean style involves the dismissal of data, but not, absurdly, in a corrupt way, but in the way in which Galileo sets aside 'accidental' phenomena, such as friction, air pressure, etc (see §3). See Rey (2020, chp. 1) and Allott et al. (2021) for discussion of Chomsky's approach to evidence.

with half-finished sentences, false starts, and 'errors' (e.g., agreement attraction). Such phenomena are treated as 'distortions', as not characterising the abstract syntactic system the theory hypothesises. Secondly, just as much goes on in our use of language that is eschewed, so there is much a theory should capture that doesn't occur in the data. Most trivially, the principles that can be seen to generate a single sentence will generate denumerably many others. That an infinite number of such sentences cannot be witnessed doesn't mean that a theory is free to exclude them, for excluding them might result in a theory that lacks the simplicity and elegance of explanation that animates the very idea of positing such a system. A theory is not a description of a data-set. Thirdly, simple hypotheses that promise to offer insight in one area are not necessarily rejected for creating trouble elsewhere. For example, a system of grammar that allows for *n*-ary branching is guaranteed to capture any apparent phrasal structure, but at the cost of triviality. A common hypothesis since the mid-1970s is that all branching is binary, which was enshrined in the so-called 'X-bar scheme' that provided a universal template for all phrases. Binarity itself makes for an elegant theory, rendering syntax as a minimal kind of algebra (what is called a magma). It also allows for generalisations otherwise unavailable and offers explanation of puzzling phenomena, such as double-object constructions, but it comes at a cost of forcing one to reanalyse co-ordination and stacking of adnominal modifiers.[4]

Some of these points will be properly explained in subsequent sections. Before that, I want to elaborate on just what Galileo proposed, which creates some interpretive problems, as advertised in my introductory remarks, when applied to linguistics.

3. Galilean style

The term 'Galilean style' was coined by Husserl (1936/70) to characterise Galileo's distinctive explanatory approach, which was to define the break with Aristotle that heralded modern philosophy and science.[5] There is an irony here, unremarked upon by either Weinberg or Chomsky. For Husserl, the new Galilean way of thinking is essentially problematic in initiating a 'crisis', for the world as we experience it (*lebenswelt*) is radically dissociated from the world described by the new science. This 'crisis', according to Husserl, was generated by a kind of forgetfulness that the very mathematical means (geometry, for Galileo) that powered the new science were grounded in experience itself. I have not the space to assess this charge, although it becomes supererogatory given the subsequent development of analytical geometry and calculus even within the 17th century, to say nothing of later developments. In short, although the Galilean style might have originated with intuitive geometry, it is not defined by it. What, then, is this style?

The rise of mechanical philosophy in the 17th century had its explicit progenitors in the work of Epicurus and Lucretius, the latter rediscovered in the late medieval period. Epicurean metaphysics was monist, with *all* phenomena being the result of different kinds of interaction between 'atoms' in motion in a void: nature is a purposeless set of interactions, whose precise specification is to be discovered. Modern thinkers differed along various dimensions, such as the proper place of God and minds, the nature of divisibility (/infinity), and whether there is a void or a plenum. Still, non-human nature, most thinkers agreed, is a teleology-free realm supported by a monistic metaphysical base. The explicit target of the new philosophy was a general Aristotelianism, which for present purposes might be characterised as a view of nature as imbued with purpose or design, where things have a proper place and behave in ways appropriate to what is proper to them (heavenly bodies move in circles; terrestrial bodies in straight lines). Moreover, the whole architecture of nature is available to observation, grossly divided between a few elements and their compounds, and their proper regions, such as on Earth, in the sublunary region, or in a fixed position in the heavens, or beyond. As Boyle (1686/1996), observed, *Nature* on this Aristotelian conception was not so much an unknown realm whose character is to be revealed by inquiry, but a kind of purposeful system, whose aims are open to view, discerned from observation.

[4] To see the point here, it is natural to read phrases of the form 'A and B' as being symmetrical, which appears to be in accord with their semantics. But uniform binary branching appears to leave but two options, neither of which is symmetrical:

(i) a [A [and B]]
b [[A and] B]

In distinction, a double object construction *looks* flat, but we know that the direct and indirect objects are differentiated:

(ii) a Bill gave Mary flowers
b *Bill gave flowers Mary
c Bill showed the boys themselves
d *Bill showed themselves the boys

See Zhang (2009) for discussion.

[5] Of course, no 'revolution' is entirely due to one person. Galileo's work was in some sense the apogee of an anti-Aristotelian tradition informed by humanism, the reformation, and the rise of neo-Platonism in the previous century (see, for example, Greengrass, 2014, chp. 7).

Galileo overthrew this philosophy, but not so much via an endorsement of an Epicurian mechanical metaphysics, as (more or less) proffered by Gassendi, Hobbes, Boyle, and Descartes, although Galileo was assuredly an atomist. For Galileo, method and explanation came first; whether the approach issued in a mechanical metaphysics or not was a secondary issue.

The Galilean style can be described as three moments:

(i) Complex observable events are decomposed.
(ii) The decomposition is in terms of mathematically specifiable lawful phenomena and local contingencies.
(iii) The observable events are reconstructed as the interaction effects of the formally specifiable phenomena with contingencies.[6]

As remarked upon above, Koyré (1938/78) famously read Galileo as a Platonist in opposition to his Aristotelian forebears, with the formally specifiable phenomena being real in contrast to the illusion of the directly observable. Parking issues of Galileo's actual intent, the style itself certainly does not entail such a view.

Consider perhaps Galileo's most famous methodological statement:

> Philosophy is written in this grand book, the universe, which stands continually open to our gaze. But the book cannot be understood unless one first learns to comprehend the language and read the letters in which it is composed. It is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures without which it is humanly impossible to understand a single word of it; without these, one wanders about in a dark labyrinth (1623/1960, pp. 183–4).

Galileo is here both appealing to the old scholastic model of treating the natural world symbolically, Scripture-like, as if made to be understood by us in ways that require a hermeneutics, and overturning that very model, for the required hermeneutics is the pure abstraction of mathematics.[7]

Further, Galileo in the same work introduced a version of what came to be called the 'primary/secondary quality' distinction (the distinction also appears in Lucretius). The distinction certainly appears to render as an illusion what Husserl (1936/70) sought to preserve as the reality of our *lebenswelt*:

> Without the senses to guide us, reason or imagination alone would perhaps never arrive at such qualities. For that reason I think that tastes, odors, colors, and so forth are no more than mere names so far as pertains to the subject wherein they reside, and that they have their habitation only in the sensorium. Thus, if the living creature were removed, all these qualities would be removed and annihilated (1628/1960, p. 309).

The point throughout the passages, however, is the *explanation* of heat, sound, odours et al. in terms of the relation between physically recognisable properties (motion and various media) and our senses. Galileo's primary interest is not metaphysical. If natural events are to be understood, then one must decompose them to identify the mathematically lawful phenomena, such as of motion; a mere classification of the observable via 'substantial forms' or the imputation of ends and purposes will not suffice. In this sense, there are not two realms, one perfect and unchanging, the other a sensory flux, but a single theoretical orientation, which finds necessity and lawfulness both in mathematics and in the natural world itself. Nature, in this light, is elevated more than it is disparaged. Thus, the observable world is not rendered as mere illusion but explained as an interaction between our minds and objects outside of us. Qua an interaction effect, the colour of an object, say, is no more an illusion than is the *actual* motion of a falling body.

It bears emphasis that the Galilean requirement for abstraction and idealisation (frictionless planes, free fall, etc.) is a condition on the explanation of actual events rather than the pursuit of a transcendent truth *a la* Neoplatonism. Consider the following passage:

> Just as the computer who wants his calculations to deal with sugar, silk, and wool must discount the boxes, bales, and other packings, so the mathematical scientist, when he wants to recognize in the concrete the effects which he has proved in the abstract, must deduct the material hindrances, and if he is able to do so, I assure you that things are in no less agreement than arithmetical computations. The errors, then, lie not in the abstractness or concreteness, not in geometry or physics, but in a calculator who does not know how to make a true accounting. Hence if you had a perfect

[6] Husserl's (1936/70, p. 33) central claim concerning the 'style' is:

> By means of pure mathematics and the practical art of measuring, one can produce, for everything in the world of bodies which is extended in this way, a completely new kind of inductive prediction; namely, one can "calculate" with compelling necessity, on the basis of given and measured events involving shapes, events which are unknown and were never accessible to direct measurement. Thus ideal geometry, estranged from the world, becomes "applied" geometry and thus becomes in a certain respect a general method of knowing the real.

The point here is that the abstract or perfect mathematical model, which admits precise reasoning and measurement, has application to the 'sense' phenomena that were abstracted away to arrive at the model (i.e., objects are construed solely in saptio-temporal terms). The ideal thus becomes a measure of the real, as Weinberg (1976) stresses.

[7] For background on the hermeneutical approach to the natural world, see Gaukroger (2006, chp. 4).

> sphere and a perfect plane, even though they were material, you would have no doubt that they touched in one point (Galileo, 1632/1967, pp. 207–8)

I think there is a natural way to understand this passage, which relates to Galileo's wider cosmological perspective. Methodologically, Galileo made the sublunary superlunary (cf., Cassirer, 1942, 1985). As the latter used to be considered unchanging and perfect, so Galileo reversed this by identifying mathematically precise constants, unchanging fulcrums of order hidden by observational contingencies. In this sense, the abstract models depict the reality of our world, albeit not experientially. The models might be understood as specifying modal or counterfactual properties of the actual world, how the world would be, if free of 'accidenti'. The ideal model does not falsify the world, but simplifies it by 'deduct[ing] the material hindrances'.

Consider two examples. Galileo's inclined plane 'experiment' compared the rate of acceleration of a free-falling body on a vertical AB and the rate of motion along the hypotenuse AC. If both movements were actual, different effects would be in play (the pressure of the atmosphere, the friction of the plane), but the rate is measured as the same, i.e., the distance travelled is proportional to the square of the time (on the incline, a body will have covered 25 of the distance in 5 of the time); it is not proportional to the mass of the bodies. Thus, the idealisation is aimed at capturing an invariance, what we would now describe as acceleration under gravity, in terms of which different events can be theorised as compositions of different factors and the same forces. The alternative, for Galileo, is to be faced with infinite difference, much as one actual event is always different from another:

> Of these properties [accidenti] of weight, of velocity, and also of form, infinite in number, it is not possible to give an exact description; hence, in order to handle this matter in a scientific way, it is necessary to cut loose from these difficulties; and having discovered and demonstrated the theorems, in the case of no resistance, to use them and apply them with such limitations as experience will teach (Galileo, 1638/1954, pp. 252–3).

The full power of this thought can only be appreciated once it is recognised that, at the time, it was wholly unclear, empirically, whether one could abstract from 'accidenti' and be left with any phenomena at all. The existence of a vacuum, for instance, was hotly contested, and would remain so for the next 50 years. In other words, Galileo is not presupposing that there is a vacuum, still less did he seek to demonstrate its existence, as Boyle and others attempted to do, no more than Galileo presupposes the existence of a perfect material sphere that might be in contact with a surface at a single point. Galileo, rather, is insisting that science is only possible once we abstract to 'perfect' mathematical systems. Don't ask how things actually behave, upon observation, but how they would behave were factors essential to observation removed. Only then might we return to the observable with a proper understanding.

Similarly, Galileo's series of 'experiments' concerning inertial frames of reference and the apparent rectilinear motion of bodies are designed to show that there is an invariant inertial state that is the same between rest and uniform velocity. The 'appearance' of being at rest is saved, but is rendered as equivalent to uniform velocity such that falling bodies from the mast of a ship sailing by a coast behave the same as falling bodies on the coast. The invariance is supported by inertia, which cannot be observationally isolated, for some body will always be acting on some other body, but inertia still has its real effects in the difference between rest and motion being rendered as relative to a system of measurement. In this context, Galileo does not set out to show that the earth is in motion, only that all observations that suggest it is at rest equally indicate that it is in uniform motion.[8]

With free fall and inertia in place, Galileo was able to specify the motion of a projectile as inscribing a parabola. The reigning Aristotelian conception proposed that an impetus was somehow imparted to a projectile, and when, somehow, it was exhausted, the body would fall vertically. If, however, a projectile is subject to both inertia, which dictates that the body continues in its line of flight, and free fall, which dictates that the body will fall perpendicular to its origin, then the composition of forces will produce a curve. Thus, the metaphorical opacity of imbued impetus, which gives false predictions, is replaced with a mathematically precise specification in terms of unobservable forces that issue in accurate predictions.

In the remainder of this paper, we shall see how the Galilean style is witnessed, and liable to be misunderstood, in two areas of generative linguistics that have proved contentious at a methodological level—the competence/performance distinction and computation.

4. Competence vs. performance

Chomsky's (1965) celebrated competence/performance distinction marks a division between use (performance) of language and a putative underlying grammar known or cognized by the speaker (competence). The natural charge has been that a focus on the latter neglects data and the integration of language into wider cognition and human sociality. It is hobbled by a syntacocentrism that places a formal system at the heart of language to the neglect of everything else. From a Galilean perspective, however, the impugnation is misplaced. What explanation involves is precisely a decomposition of complex phenomena in the search of formally precise invariances. Such an approach does not essentially neglect phenomena, but hopes to capture them as interaction effects, much as Galileo saw the fall of actual objects in real circumstances as an interaction effect. Chomsky's introduction of the competence/performance distinction is explicit on this score:

---

[8] Galileo continually sought evidence of the Earth's motion in tidal phenomena (Galileo, 1616/1989; 1632/1967, pp. 416–465). The decisive evidence was to arrive two centuries later with observation of parallax phenomena and Foucault's pendulum experiment.

> Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance … To study actual linguistic performance, we must consider the interaction of a variety of factors, of which the underlying competence of the speaker-hearer is only one. In this respect, study of language is no different from empirical investigation of other complex phenomena (Chomsky, 1965, pp. 3–4)

Note two things. First, while Chomsky's appeal to the 'ideal speaker-listener' is akin to Galileo's ideal pendula and ideal motion (which free fall amounted to), it is in tension with his latter 'internalism', where the object of study is the internal states of a speaker-hearer (Chomsky, 1986). We shall see how this tension plays out, but, minimally, appeal to an ideal speaker-listener is a way of isolating an invariant phenomenon from general factors that variably contribute to actual use of language, much as Galileo eschewed the infinite accidents that attend motion. Secondly, Chomsky is not suggesting that 'performance' is to be forgotten about; to the contrary, an adequate theory of performance will view it as an interaction effect. But there is a disanalogy, which makes the 'Galilean' epithet misleading.

Recall the point stressed in the previous section. For Galileo, the abstract models are real, not in any Platonic sense, but in the sense that they are explanatory of the phenomena, which is rightly seen as a composition of factors, including the abstract properties. A pendulum, say, viewed as having a period subject to the length of the massless rod and gravity alone is ideal, much as Chomsky's 'ideal speaker-listener' is.[9] Further, Galileo is offering an account of actual pendula, which fall short of the ideal. If one could have massless rods and a vacuum, the period of a pendulum would be precisely described by the equation. By parity of reasoning, Chomsky should be offering a theory of a speaker-hearer; that is, under 'ideal' conditions, the theory would explain the actions of the subject, just as Galileo's equation captures the movement of the ideal pendulum. In other words, competence would be a competence to produce and consume linguistic material, albeit under ideal conditions (cf., Devitt, 2006). Yet this is not so. Chomsky is clear that 'a generative grammar is not a model for a speaker or a hearer' (1965, p. 4).

The point might be put as follows: with the pendulum, if we abstract to the ideal, we still have a pendulum, just as if we abstract to a vacuum, we still have motion, or if we abstract to a perfect sphere, it would still be in contact with a surface at a single point, but with the speaker-hearer, if we abstract to just the generative grammar, we don't have a speaker-hearer at all. This difference is due to the difference between the respective explananda. Galileo's concern is for kinematics (motion and interaction in space), regardless of the internal constitution of the objects. Chomsky is concerned with internal states. Thus, the production–consumption profile of Chomsky's 'ideal speaker-listener' is *not* the target of the theory; rather, such an ideal agent would exhibit a profile that would provide a clearer view of what the generative grammar—the actual target of the theory—contributes to performance. In this sense, performance is essentially bound up with factors that fall outside of the purview of a grammar, even under ideal conditions. What is not noted by Chomsky is the tension between appeal to the idealised speaker-hearer and the insistence that the theory is *not* a model of the speaker-hearer, not even an ideal one, presumably (cf., Rey, 2020, p. 20). The crucial point here is that if the theory targets internal states, albeit abstractly specified, then the ideal speaker-hearer is the wrong kind of Galilean idealisation, as it were. We need to be abstracting away from interactions between systems, some of which might be necessary for speaking-hearing, rather than be idealising to a perfect speaker-hearer. Let us consider a specific phenomenon to bring the issue here properly to light.

Miller and Chomsky (1963) had observed that speaker-hearers cannot parse, and so they judge to be unacceptable, centre embeddings of relative clauses:

(1) Men women cats dogs bit like marry hate pets

Even with overt complementizers, unacceptability still results:

(2) a Men that women that cats that dogs bit like marry hate pets
b The novel that the horror author who the publishing company had recently fired had typed quickly was banned by the local library

Cases such as (2b) give rise to an intriguing 'illusion' (Huang and Phillips, 2021). For example, if the VP *had typed quickly*, which has *the horror author* as its subject, is elided, acceptability improves, even though the result is ungrammatical.

(3) The novel that the horror author who the publishing company had recently fired was banned by the local library

[9] Galileo's equation for pendula is:

$$T \propto 2\pi\sqrt{\frac{l}{g}}$$

The crucial feature here is that it is only the length of the rod (l) that is the significant variable, not the mass of the bob.

There is also controversy over what the limit of acceptable embedding is.[10] These complexities should certainly inform any adequate account of the competence/performance distinction. My present purpose, though, can be forwarded by just consideration of (1)-like cases. I assume that the eschewed complexity will not cut against me.

What is clear is that cases such as (1) appear to be word-salad. From a usage perspective, therefore, one might think that (1) is simply not part of the language, ungrammatical. If so, then whatever recursive procedure generates relative clauses must be bounded, presumably due to general factors and experience (Christianson and MacDonald, 2009). Such an inference would be erroneous.

Consider that the simplest specification of a recursive procedure or definition is unbounded, and so it is with relative clauses. Their syntax remains controversial, including the possibility that there is movement out of an initial clause to form the head, but to simplify greatly, suppose we depict the underlying facts in terms of a pair of re-write rules:

(RC) (i) NP → N RC
(ii) RC → NP VP

So, if we have the NP *men women marry* by (i), then the NP of the relative clause by (ii) can contain a relative clause: *men women cats like marry*. And so on. Suppose that unacceptability kicks in on the third iteration. If we were to posit a bound on relative clause formation, however, then such a bound would vacuously feature in the very rules that generate the acceptable cases, and, in fact, would never apply, unless a speaker finds herself studying linguistics. If we are to posit a bound, therefore, it would be most sensible not to have it affect the procedure that pays no heed to such a bound, but to have it as a factor that constrains the application of the procedure. But this is precisely Chomsky's claim: however we are to theorise the boundedness of centre embedding, it does not affect the basic linguistic principles, which, qua recursive, remain unbounded, but only their application in potential performance. This split between the rules and their application is not noticed because the data are nigh-on always bounded, not the rules themselves. The general point here is that a recursive procedure cannot be learnt stepwise, for if we posit a rule with an intrinsic bound on the depth of its application, which is to be removed or extended during learning or evolution, the new rule will be equivalent to the unbounded recursive rule plus a bounding condition, and so be actually more complex than the simple recursive rule itself.

Of course, the usage-based theorist would forswear the invariant rule, even if such a rule is adequate for the actual data on which people operate. Recursivity, in this sense, would be learnt from the data and always be bounded—hence its break down when the data goes beyond the learnt limits (Christianson and MacDonald, 2009). What is clear, at least, is that how recursivity shows itself in a particular language must be acquired from data. For example, the Saxon genitive allows embedding (*Bill's friend's dog*), but no such construction is available in the corresponding Spanish. This difference, however, pertains to where recursivity shows up, not whether recursivity as such is learnt from and reflects the limitations of the data, for whenever a recursive principle is acquired, it does not reflect the limitations of the data.[11] If we were to suppose that, at least in the case of centre embedding, the limitations of the depth of relative clause formation reflect the limitations of the data, we are faced with two quandaries.

First, (1) and its ilk appear not to be straight-up gibberish; indeed, (1) has a precise interpretation, which can easily be explained, and which follows from the kind of interpretation acceptable relative clause constructions possess. It might be thought that this level of understanding is due to some form of analogical reasoning.[12] I don't see any way to decide this issue at the level of judgement data, but two theoretical considerations weigh heavily against the anological route. First, if (1) is ungrammatical, then all of our accounts of relative clause formation are wrong, for all allow for centre-embedding. Secondly, no information beyond (1) is called for to explain its grammatical status; all that is required is the indication of the relative clause boundaries. It is difficult, therefore, to see the debriefing as analogical, for there is no required analogue.

The second reason to doubt that limitations in the data determine limitations in depth of embedding is that if recursivity is learnt from data limited to non-centre-embedding in relative clause formation, then one should expect recursivity to be similarly limited in other cases where the data is relevantly limited. That is to say, if boundedness of recursion follows from the character of the data, then equivalent boundedness should be found wherever data is relevantly limited. This is not the case. Consider:

(4) a. The boy behind the girl behind the man behind the dog … is blonde
b. 1 is the predecessor of 2, which is the predecessor of 3, which the predecessor of 4 …
c. The woman believes that the man hopes that the girl wants that the shopkeeper … is home
d. Bill's sister's friend's dog's … dinner

[10] See Karlson (2007) for discussion.

[11] In other words, the language-acquiring child needs data on when recursion may apply. Daoxin et al. (2021) propose an interesting answer to the quandary. The child identifies a pair of nouns, for example, where one my substitute for the other (*A friend's sister's* … to *A sister's friend's* …). This triggers the generalisation that, as we would say, a genitive complement can contain a genitive, which is not limited to the depth of embedding the data happens to exhibit.

[12] We needn't doubt that analogical reasoning can rescue many constructions that are ill-formed (*vide* the famous, *Many more people have been to Paris than I have*).

Park the question of the syntax of these constructions. Also park the fact that one's capacity to understand them gives out, which everyone happily acknowledges, although a case like (4a-b) can be extended indefinitely. The crucial thing to note that is that these kinds of constructions are not used up to whatever limit our understanding packs up. If recursivity in these cases were leant from the data, then we should expect the limit of our understanding to be aligned with the constructions' occurrence, but it is not. Multiple embeddings of all kinds are uncommon in the primary linguistic data. It also bears emphasis that children show sensitivity to embedded clauses in terms of nonlocal dependencies in questions (Perkins and Lidz, 2021). This indicates that children are not using data to determine the possibility of embedding, but are able to interpret embedding when it arises.

The difference between (1) and (4) remains unsettled, but let me advertise a lovely idea from Fodor et al. (2019).[13] The recursion involved in relative clauses applies centrally, not rightward. Traditionally, approaches to the centre-embedding puzzle have sought to relate this feature of recursion to constraints on working memory. Another approach, however, is to note that the centre-embedding cases make for prosodic difficulties, which is the basis of the present account. If we think of a sentence in terms of prosodic chunks, then how the chunking is done affects comprehension:

(5) a The pipes the unlicensed plumber/the new janitor reluctantly assisted/tried to repair burst
b The rusty pipes the plumber/I hired/had tried to fix continue to leak occasionally

This pair displays the so-called 'pronoun effect', where reducing the morphological load of the second relative clause aids comprehension. Surprisingly, balance between phrases also helps:

(6) The rusty old ceiling pipes/the plumber my dad trained fixed/continue to leak occasionally

Centre-embedded relatives, therefore, turn out not to be essentially unparsable, and do not relate to working memory or processing constraints. Instead, the right kind of prosodic properties need to be realized, i.e., three balanced chunks reflecting natural prosodic breaks. But such stars being aligned is not common or easily achievable in normal language use. Crucial for our purposes is that such a use of prosody in comprehension is pure performance in the sense that neither syntax nor semantics is mirrored in prosody. It is because prosody is not recursive or compositional, but parsing relies upon it, that comprehension is out of step with competence. This is familiar form other cases. For instance, while prosody (placement of focus) can serve to disambiguate in some cases, it cannot in all:

(7) a Billy knows how $[\text{good}]_F$ meat tastes
b Billy knows how good $[\text{meat}]_F$ tastes
c Visiting relatives can be boring

Thus, while a difference in focus can readily locate *good* as an adnomial adjective or an adverbial, no focus or other prosodic feature can resolve the ambiguity in (7c). Similarly, prosody gives no indication of whether the accusative form in exceptional case marking constructions belongs to the matrix verb phrase or the subordinate infinitive:

(8) a Billy believed him to be best for the job
b Billy considered her to be reckless

If some such account is on the right lines, then all phenomena are accounted for, including the interpretability of centre embeddings, consistent with recursion being unbounded independent of whether the data is relevantly limited or not; indeed, a prosody-based account gives an explanation for why performance errors, at least with centre embeddings, arise: comprehension makes uses of prosody, which is independent of syntax and semantics.

The bottom-line here is that consumption and production makes essential use of extra-grammatical properties, such as prosodic structure, that simply fall outside of the purview of a competence theory. We cannot abstract away from them and still be left with a model of consumption-production, no more than it makes sense to think of consumption-production without an external vehicle, some kind of material sign.

The basic point here has become stark as the generative enterprise has progressed. If one has a conception of a grammar as a set of rewrite rules with linear order a feature, then it at least makes sense to think of a grammar much like Galileo thought of pendulums, for the grammar will generate strings recognisable as English, say, even if they lack prosodic information. Rewrite rules were moribund, however, by the early 1980s, and utterly forsaken by the time of the advent of the so-called minimalist programme (Chomsky, 1995). The problem with the rules is that they simply amount to stipulations as to the structure of the language, without them following from or being explained by anything more basic. For example, we might want to say that 'VP $\rightarrow$ V(DP)' is a descriptively accurate statement about English, but it offers no enlightenment. Firstly, that

[13] It had occurred to me when teaching on centre-embedding that normal prosody breaks down in the standard examples. I asked Janet Fodor why that might be. It turns out that she had an answer.

VPs are headed by Vs is simply assumed, for there is nothing in the nature of a rewrite rule that entails such a truth. Secondly, it is equally assumed that the object (if any) follows the verb. Again, there is nothing in the logic of the system that entails that. Of course, in some languages, the object precedes the verb, and so the connection between the verb and its object simply becomes a matter of rejigging a definition. A deeper alternative, in general terms, will leave questions of what the head of a phrase is and linear order a language realises as phenomena to be explained. Bare syntactic structure, then, might leave these as options, only fixing that a head is contained in the phrase, and that the object of a verb is a sister of the verb. If that is right, then, again, any consumption-production profile has been abstracted away, for anything produced or consumed must be linear.

This raises the issue of how, in a Galilean spirit, we are to think of the internal system and the generative grammar that is a theory of it. Here we have an internal Galilean attitude, as it were, where the system is ideal, not with respect to external signs but the organisation of the human brain.

5. Computation as generation

Competence, or the 'knowledge' of a language, which is the topic of inquiry, rather than the use we may make of it, is understood to be computational. The notion is ambiguous, however, between a formal and a process construal. There is a further ambiguity between the theorist's perceptive and the perspective of the object of the theory. This section will attempt to disentangle these ambiguities in a way that will reveal them to be intimately related. In short, I shall argue that treating competence as computation is a reflection of the *kind* of theory that proves adequate, rather than a metaphysical commitment to the kind of states that a speaker-hearer realises. Again, this is a Galilean attitude, which priorities a certain kind of explanation over any metaphysical presumptions.

Mathematically, computation is the mechanical or stepwise determination of the values of a function. Viewed *extensionally*, a function is a set of ordered pairs, corresponding to the values of the independent and dependent variables, respectively. Viewed *intensionally*, a function is a procedure that, for any first member of a pair, will issue in the second member. Distinct procedures can be extensionally equivalent. For trivial examples, '$2x^2 - 2$' and '$x^2 - 1$' have the same values, as do '$x^2 - 2x + 1$' and '$x - 1$'.

If we consider a grammar, then, to be a function from primitive lexical items to sentences, we can ask if the function is to be understood extensionally or intensionally. The former conception is curiously too strong and too weak.

It is too strong: since English (whatever that might be) is an empirical phenomenon, we simply don't know what the phenomenon is (what should definitely count as English) prior to having a theory. It would thus be impossible to say whether the procedure was correct or not in the absence of an independent measure of what is to count as correct or not, which is exactly what a theory is supposed to provide (Chomsky, 2005).[14]

Equally, the extensional approach is too weak, for it is explanatorily empty. Consider the following grammatical chestnuts:

(9) a The shooting of the soldiers scared the tourists
b The barking of the dogs scared the tourists
(10) a Sally is eager to please
b Sally is easy to please

All of these belong to English, as it were, but such a gross classification leaves the interesting phenomena unexplained (9a) is ambiguous between the soldiers shooting or them being shot (9b) is not similarly ambiguous, for one cannot bark a dog. In other words, the root verbs that appear under progressive aspect are, respectively, transitive and intransitive.[15] A mere classification cannot inform us of this, but nor is the difference extra-linguistic information a grammar may eschew, for the resolution of the ambiguity or lack of it determines the grammatical position in which the respective PP adjuncts are interpreted. In distinction, neither of (10) is ambiguous, but the surface subject *Sally* is interpreted differently relative to the subordinate infinitive. In (10a), *Sally* is the understood subject of *please*, whereas in (10b) *Sally* is the understood object of *please*. Again, this is not extra-linguistic information a grammar may neglect. Such cases can be multiplied indefinitely.[16]

In this light, the standard conception emanating from Chomsky (1957) treats a grammar as a function in intension that generates structures that might be explanatory over the observed phenomena of the kind just sketched. The structures just

[14] Claims of indeterminacy regarding semantics and syntax arise from prioritising an extensional conception (Quine, 1960, 1972; Stich, 1972). The original sin is the assumption that it is possible to fix the extension of the relevant linguistic categories without some intensional specification of what such categories are.

[15] Verbs enjoy some flexibility of construal, of course. One can use *bark* as a transitive with the right background in place. Imagine that a dog needs to be made to bark as part of its treatment. It would be natural to describe the required activity as a barking of a dog. The point of the traditional example, however, is the contrast between the underlying transitive and intransitive root, which doesn't rely on the transitive/intransitive distinction being lexically fixed.

[16] A number of complex issues arise here concerning the development of syntactic theory; in particular, 'rules' which apply to and change structures into other structures gave way to general principles and conditions, which constrain the production and interpretation of structures that are otherwise freely generable (for discussion of these developments, see Chomsky, 1995, chp. 1). For present purposes, these internal developments do not affect the general problem I am addressing, which is most vivid in the context of the *Aspects* discussion.

are reflections of a specific procedure that generates them. Although in the case of a natural language the nature of the procedure is open to empirical inquiry, the general feature follows from the mathematics. For instance, in model theory, a set of natural numbers is not posited as primitive elements; rather, 0 and the successor function are specified, from which all positive integers are recursively enumerable: $1 = \text{suc}(0)$; $2 = \text{suc}(\text{suc}(0))$; $3 = \text{suc}(\text{suc}(\text{suc}(0)))$. In effect, the integers become structured objects, or members of a sequence, in terms of a number n being the nth successor of 0. Just so in the case of linguistic structure.

A sentence might be defined as a linear arrangement of words, but such is not what a grammar generates. It generates structures containing other structures, each of which corresponds to a step in the procedure which generates the whole. Such structures are potentially many-one related to sentences as intuitively understood.

Although the bare bones as sketched reflects early generative theory, the same underlying ideas remain in the latest work. One component of current theory is the status of the basic combinatorial operation, called Merge, that underlies syntactic organisation. On a standard conception, all structures are binary with each merged object being open to merge with another object, either simple or complex; the grammatical function thus generates the transitive closure of immediate 2-membership containment defined over lexical items (Chomsky, 1995).

We now arrive at a first ambiguity. Assuming that a grammar is a function in intension that generates a class of structures, is the procedure an actual algorithmic process neurologically realised? Chomsky foreswears the 'psychological reality' of a grammar *in this sense*. A grammar is simply one that is explicit, in the same way as a recursive definition can be made explicit: it enumerates the members of class by specifying them in terms of a procedure for their generation, free of any further assumptions or conditions or intuitive criteria:

> If the grammar is … perfectly explicit—in other words, if it does not rely on the intelligence of the understanding reader but rather provides an explicit analysis of his contribution—we may (somewhat redundantly) call it a *generative grammar* (Chomsky, 1965, p. 4, p. 4)

Although this purely formal characterisation does not preclude the grammar specifying an actual process, Chomsky is clear that it is not intended to detail any process:

> [A] generative grammar is not a model for a speaker or a hearer. It attempts to characterize in the most neutral possible terms the knowledge of the language that provides the basis for actual use of language by a speaker-hearer. When we speak of a grammar as generating a sentence with a certain structural description, we mean simply that the grammar assigns this structural description to the sentence. When we say that a sentence has a certain derivation with respect to a particular generative grammar, we say nothing about how the speaker or hearer might proceed, in some practical or efficient way, to construct such a derivation. These questions belong to the theory of language use—the theory of performance (*op cit.*).

Chomsky is here meaning to exclude the identification of a grammar with a parser, although a parser will make use of the information a grammar provides.

There is, however, a fundamental issue elided: psychological reality. If we have an artefact that is a computer, then the function in intension is part of the programme or software, not the machine itself; indeed, the function can be specified independent of any machine at all. If we are thinking of natural systems, then the function is to be discovered and its specification will prove explanatory *only if* it is, as it were, run on the system. In this sense, the function *must* be an account of the underlying causal architecture, albeit one that does not entail a definite mechanism, i.e., it can be multiply realised. This is just what it means to think of a natural system in computational terms. If we are not committed to the function being realised or run, then the theory appears only to offer a *description* of the output of the system (Pylyshyn, 1991; Soames, 1984; Devitt, 2006). To be sure, the description might be assessed for empirical coverage, but it remains, so the thought goes, unclear how it could be genuinely explanatory or count as *the* correct theory. I shall offer a provisional answer to this quandary, which will be resolved once we have the second ambiguity on the table.

Chomsky (1980, p. 106–9) rejects any complaint that a generative grammar is not psychologically real as an appeal to a 'mysterious property'. Chomsky suggests that the complaints are predicated on the presupposition that, to be a real, a grammar, must be constrained by certain kinds of evidence, such as from parsing or behaviour or neuroscience, and typical hypotheses are not so constrained. Yet a grammar *is* open to any and all conceivable evidence. The problem, such as it is, for generative theory and all other theories of language, is that it is hard to marshal evidence that is *both* relevant to the hypotheses of the function in intension *and* suitably low level or concrete to count as indicative of 'reality'. Indeed, Chomsky suggests that this plight is utterly typical, and even applies to physics, where entities and structures are posited in order to explain certain phenomena even though there is currently no evidence for them, and it might even be unclear how evidence could be secured:

> Of course, there are differences [between linguistics and physics]; the physicist is actually postulating physical entities and processes, while we are keeping to abstract conditions that unknown mechanisms must meet. We might go on to suggest actual mechanisms, but we know that it would be pointless to do so in the present stage of our ignorance concerning the functioning of the brain. This, however, is not a relevant difference of principle (Chomsky, 1980, p. 197; cf., Chomsky, 1988, p. 152; 2004, p. 380; 2005, p. 2)

More recently, Chomsky has written:

> [Linguistics] is taken to be the study of a real object, a biological organ, comparable to the visual or immune systems, the systems of motor organisation and planning, and many other subcomponents of the organism that interact to yield the full complexity of thought and action, abstracted for special investigation because of their apparent internal integrity and special properties (Chomsky, 2007, p. 2, p. 2).

Chomsky's philosophical point here is that a lack of neural evidence for a grammar, say, does not entail, or even suggest, that it is a theory of something else, such as behaviour or an extra-mental realm, whether abstract or concrete, as variously entertained by Katz (1981), Soames (1984), and Devitt (2006). What a theory is *about* is not the same as what evidence happens to be available for the theory. Still, while Chomsky is certainly correct that the conflation of evidence with ontology has erroneously inflected the 'psychological reality' issue, an underlying problem remains. If a grammar (a function in intension) is intended to be about the brain rather than merely a descriptive specification of idealised linguistic output, then it *must*, it seems, be a specification of a causal or mechanical process realised in the brain, and not a mere formal entity. This has been the long-standing position of Fodor (1975): computation *just is* a mechanical process conditioned to respect informational transitions. If so, then Chomsky cannot be so sanguine about endorsing a formal understanding of computation, which would render the function a mere mathematical entity. I shall offer a resolution of this issue, which relates to a second ambiguity.

When developing theories in physics or chemistry, say, there is no temptation to attribute the theories to what the theories are about, as if matter were obliged to contain or represent the field equations of general relativity for the theory to be true. The situation is different in linguistics and other areas of psychology, where the theories are not about the behaviour of a system, but the organism's information and the use it makes of it. Bare use of 'grammar' thus creates an ambiguity:

> [A] child who has learned a language has developed an internal representation of a system of rules that determine how sentences are to be formed, used, and understood. Using the term "grammar" with a systematic ambiguity (to refer, first, to the native speaker's internally represented "theory of his language" and, second, to the linguist's account of this), we can say that the child has developed and internally represented a generative grammar, in the sense described … [W]e are again using the term "theory"—in this case "theory of language" rather than "theory of a particular language"—with a systematic ambiguity, to refer both to the child's innate predisposition to learn a language of a certain type and to the linguist's account of this. (Chomsky, 1965, p. 25)16

The ambiguity raises a simple question: What do the two notions have in common? Why is a theorist's theory a theory of the speaker-hearer's theory? One answer is that, to put it crudely, it is the *same* theory; that is, what the theorist develops, if right, *just is* the theory of the language the speaker-hearer naturally acquires by age five (more or less) (Fodor, 1983). This can't be right, however. Much of what goes into the theorist's theory is not anything that a speaker-hearer needs to know in any sense, no matter how etiolated. The theory provides explicit statements *about* the speaker-hearer's grammar, not statements *of* the grammar. For example, the theorist's theory will contain in a nigh-on explicit form that, say, all branching is binary, that the syntactic computation does not lose information, that copying extends a structure at the root, that theta-roles are assigned as the recipient item enters a derivation, that reflexives are bound locally, and so on and so forth. These are statements *about* the system, concerning what constrains it, not statements of the system that need to be represented for the use of the system. In other words, these are things, if correct, we have discovered about the system, but they correspond to no item of knowledge a speaker-hearer possesses, although they can be made aware of the effects of the system being so constrained.

It might seem that this is mere a formatting issue; that is, of course the theorist's theory is in a propositional format, but the speaker-hearer has the same information albeit realised in some more implicit manner.[17] If that is the right answer, however, then the two theories are not the same at all, in form or ontology. The theorist's theory is *about* the system of the speaker-hearer rendered in explicit propositional form, whereas the speaker-hearer's theory is not about anything at all, but amounts to, as Chomsky (1980, p. 197), quoted above, puts it: 'abstract conditions that unknown mechanisms must meet'. In this light, it is a *façon de parler* to depict the speaker-hearer as a *theorist* of language. Nonetheless, it was a natural construal, and one in which Chomsky (1975, p. 11) himself indulged.[18] Still, it encourages the same mistake as thinking of a grammar as a perfect production system: an idealised view of the output is attributed to the system that produces the actual imperfect output.

[17] Approaches of this kind have been variously entertained in relation to both grammar and semantic theory; see Stabler (1983), Peacocke (1986, 1989), and Davies (1986, 1989). The core idea is that a theory can be said to be true of a speaker-hearer, not because she represents its axioms or generalisations, still less that she behaves in certain ways, but because the structure of the theory will be 'mirrored' in the causal structure of the speaker-hearer. In this light, Peacocke (1986) thinks of the 'information' a relevant theory attributes as being at 'level 1.5', meaning it is neither merely physical nor strictly represented. For critical discussion in relation to language, see Collins (2004). Note that the principal issue here is not the status of content (how mental states can be *about* things), but the status of a computational theory (how a physical system makes true a computational theory). Thus, as Egan (1995, 2003) stresses, one can have various sceptical attitudes towards content while still being 'realist' about computation.

[18] The developments in the theory, especially pertaining to the design of the language faculty, made it less natural to depict the language-acquiring child as a theorist.

Now, if the moral that holds for the idea that speaker-hearers possess a 'theory' hold too for computation, then all quandaries are resolved; that is, we have an answer to why a particular computational theory might count as true of the system rather than just a formalisation of the system's idealised output. From the theorist's perspective, the speaker-hearer realises a computational system, because a computational system is the perfect system to which, in a Galilean spirit, we abstract, much as we view a pendulum as possessing a massless rod or an inclined plane as being frictionless. Yet the move from deflating 'theory' to deflating 'computation' cannot be so swift.

First, it might be complained that here we are merely raising the traditional 'vacuity' complaint, that virtually any physical system could be said to be a computer, and so a thorough anti-realist attitude towards computation is apt (Putnam, 1988; Searle, 1992). This charge can be seriously doubted, for the ascription of computation certainly imposes *some* conditions upon the underlying system, even if there is no consensus what such conditions precisely are.[19] The view being entertained, however, does not deny the existence of specific conditions, but only claims that they are 'unknown' in the case of language, and so only discriminated given the explanatory purposes of the theory. Such a position is perfectly realist to the extent that we should pursue inquiry on the presumption that there is a single correct theory to be found. The crucial rider is that the conditions for the theory to be correct do not involve or entail any underlying states specifiable as computational independent of the theory. This position might be thought to be too weak.

Fodor (1975), Pylyshyn (1984), and Rey (1997) appeal to Turing's fundamental idea that computation involves local mechanical state transitions whose discreteness is such as to make the states serve as symbols, i.e., mechanical transitions realise informational relations. One can accept this view, however, while still thinking that what makes a system a computer is that only a computational theory is adequate for its explanation, independently of whether or not any physical states are discriminable as realising the computation. Indeed, this was precisely Turing's own view. He writes:

> [D]iscrete machines are the machines which move by sudden jumps or clicks from one quite definite state to another. These states are sufficiently different for the possibility of confusion between them to be ignored. Strictly speaking there are no such machines. Everything really moves continuously. But there are many kinds of machines which can profitably be thought of as being discrete state machines. For instance in considering the switches for a lighting system it is a convenient fiction that each switch must be definitely on or definitely off. There must be intermediate positions, but for most purposes we can forget about them (Turing, 1950: p. 439).

Note that this is not suggestive of the anti-realist vacuity claim, still less an endorsement of it. The claim is only that it profits thought, or explanation in my parlance, to impose discreteness. If this is not done, then generalisations and predictions are lost. In the case of language, the combinatorial novelty of linguistic consumption and production, massively underdetermined by any external factors, demands a computational approach. In this sense, Turing's notion of a 'convenient fiction' underplays the demands, at least if generalised. It is more that some phenomena involve a necessary idealisation, just as Galileo pursued. Treating walls or rocks as computers, might be permissible, but is a *useless* fiction.

It might be protested that while Galileo's pendula, say, are genuinely abstract in the sense of being unrealisable (massless rods are impossible), computers are perfectly realisable; indeed, the physically producible computer amounts to a proof of concept that cognition is computation. But Galilean idealisation needn't abstract to the impossible. As made clear in §3, Galileo abstracted to motion *in vacuo*, regardless of whether a vacuum is possible or whether a plane can be frictionless or not. The same holds for computation. The computational theory abstracts to discrete states, and while states answering to such discreteness must be in place ('unknown mechanisms'), their discreteness holds from the perspective of the computational explanation of the system.

This gap between theory and reality is precisely the kind of unknown that Weinberg (1976), as quoted by Chomsky (1980, p. 8), takes to be akin to Wigner's (1960) puzzle concerning the unreasonable effectiveness of mathematics. Somehow, formal theories manage to reflect underlying systems, and such theories earn their keep by accounting for phenomena, not by a confirmation that the underlying system is as the theory says, but only because the theory that trades in the relevant states is indispensable. All the action thus devolves upon what reasons we have to commend one theory over another, with their respective Galilean content of abstract, perfect systems. There is no metaphysical route to determining which one is correct, and the models give us no intimation of how reality is anyway.

I assume we just don't know at a given stage of inquiry what is '*really* real' and what is 'merely instrumental'. What appears to be true is that if a formal theory is explanatorily successful, then we take the world to somehow satisfy whatever conditions the theory essentially expresses (invariant over notational or representational differences).

6. Conclusion

The Galilean style has been crucial to the growth of natural science since the seventeenth century. It has, we might say, worked. There is no a priori inevitability about such success, however. Kant pondered, much like Wigner a couple of centuries later, on the effectiveness of the theory of conic sections to specify orbits. Kant's answer was to render mathematics as synthetic a priori. Such a resolution is no longer plausible given the scope and success of the Galilean style that far outstrips

[19] For rejection of the vacuity claim, but disagreement over just what physical conditions computation imposes, see Chalmers (1996), Copeland (1996), and Piccinini (2008, 2015).

what experience might offer us. Linguistics offers us a curious case. On the one hand, adopting the style has made for great progress, but here as elsewhere there is no guarantee of success. Once we recognise how queer any scientific success is from an a priori perspective, we might perhaps be less ready to impugn what modest but genuine success linguistics has enjoyed.[20]

Data availability

No data was used for the research described in the article.

References

Allott, N., Lohndal, T., Rey, G., 2021. Chomsky's 'Galilean' explanatory style. In: Allott, N., Lohndal, T., Rey, G. (Eds.), A Companion to Chomsky. Wiley-Blackwell, Oxford, pp. 517–528.
Behme, C., 2014. A 'Galilean' science of language. J. Linguist. 50, 671–704.
Botha, R.P., 1982. On "the Galilean style" of linguistic inquiry. Lingua 58, 1–50.
Boyle, R., 1686/1996. In: Davis, E.B., Hunter, M. (Eds.), A Free Inquiry into the Vulgarly Received Notion of Nature. Cambridge University Press, Cambridge.
Cassirer, E., 1942. Galileo: a new science and a new spirit. Am. Scholar 12, 5–19.
Cassirer, E., 1985. The idea and problem of truth in Galileo. Man World 18, 353–368.
Chalmers, D., 1996. Does a rock implement every finite-state automaton? Synthese 108, 309–333.
Chomsky, N., 1957. Syntactic Structures. Mouton, The Hague.
Chomsky, N., 1965. Aspects of the Theory of Syntax. MIT Press, Cambridge, MA.
Chomsky, N., 1975. The Logical Structure of Linguistic Theory. Chicago University Press, Chicago.
Chomsky, N., 1980. Rules and Representations. Columbia University Press, New York.
Chomsky, N., 1986. Knowledge of Language: Its Nature, Origin, and Use. Praeger, Westport.
Chomsky, N., 1988. Language and the Problem of Knowledge: The Managua Lectures. MIT Press, Cambridge, MA.
Chomsky, N., 1995. The Minimalist Program. MIT Press, Cambridge, MA.
Chomsky, N., 2002. On Nature and Language. Cambridge University Press, Cambridge.
Chomsky, N., 2004. Language and mind: current thoughts on ancient problems. In: Jenkins, L. (Ed.), Variation and Universals in Biolinguistics. Elsevier, Oxford, pp. 379–406.
Chomsky, N., 2005. Three factors in language design. Ling. Inq. 36, 1–22.
Chomsky, N., 2007. Approaching UG from below. In: Sauerland, U., Gärtner, H.-M. (Eds.), Interfaces + Recursion = Language? Mouton de Gruyter, Berlin, pp. 1–30.
Christianson, M.H., MacDonald, M.C., 2009. A usage-based approach to recursion in sentence processing. Lang. Learn. 59, 126–161.
Clavelin, M., 1968/74. The Natural Philosophy Of Galileo: Essays On the Origins and Formation of Classical Mechanics. In: Trans. by A. J. Pomerans. MIT Press, Cambridge, MA.
Collins, J., 2004. Faculty disputes. Mind Lang. 19, 503–533.
Copeland, J., 1996. What is computation? Synthese 108, 335–359.
Devitt, M., 2006. Ignorance of Language. Oxford University Press, Oxford.
Daoxin, L., Grohe, L., Schulz, P., Yang, C., 2021. The distributional learning of recursive structures. In: Dionne, D., Vidal Covas, L.-A. (Eds.), Proceedings of the 45th Annual Boston University Conference on Language Development. Cascadilla Press, Somerville, MA, pp. 471–485.
Davies, M., 1986. Tacit knowledge and the structure of thought and language. In: Travis, C. (Ed.), Meaning and Interpretation. Blackwell, Oxford, pp. 127–158.
Davies, M., 1989. Tacit knowledge and subdoxastic states. In: George, A. (Ed.), *Reflections* on C*homsky*. Blackwell, Oxford, pp. 131–152.
Egan, F., 1995. Computation and content. Phil. Rev. 104, 181–203.
Egan, F., 2003. Naturalistic inquiry: where does mental representation fit in? In: Antony, L., Hornstein, N. (Eds.), Chomsky and His Critics. Blackwell, Oxford, pp. 181–203.
Feyerabend, P., 1975. Against Method: Outline of an Anarchist Theory of Knowledge. Verso, London.
Fodor, J.A., 1975. The Language of Thought. Harvard University Press, Cambridge, MA.
Fodor, J.A., 1983. Modularity of Mind. MIT Press, Cambridge, MA.
Fodor, J.D., Macaulay, B., Ronkos, D., Callahan, T., Peckenpaugh, T., 2019. Center- embedded sentences: an online problem or deeper? In: Carlson, K., Clifton, C., Fodor, J.D. (Eds.), Grammatical Approaches To Language Processing: Essays In Honour Of Lyn Frazier. Springer, Dordrecht, pp. 11–28.
Galileo, 1616/1989. Galileo's discourse on the tides. In: Finocchiaro, M. (Ed.), The Galileo Affair: A Documentary History. University of California Press, Berkeley, pp. 119–133.
Galileo, 1623/1960. The assayer. In: Drake, S., O'Mally, C.D. (Eds.), The Controversy of the Comets, vol. 1618. University of Pennsylvania Press, Philadelphia, pp. 151–336.
Galileo, 1632/1967. Dialogue Concerning The Two Chief World Systems. In: Ed. and trans by S. Drake, revised second ed. California University Press, Berkeley.
Galileo, 1638/1954. Dialogues Concerning Two New Sciences. In: Trans. By H. Crew and A. de Salvio. Dover, New York.
Gaukroger, S., 2006. The Emergence of a Scientific Culture: Science and the Shaping of Modernity. Oxford University Press, Oxford.
Greengrass, M., 2014. Christendom Destroyed: Europe 1517-1648. Penguin, London.
Hellbron, J., 2012. Galileo. Oxford University Press, Oxford.
Huang, N., Phillips, C., 2021. When missing NPs make double center-embedding sentences acceptable. Glossa 6. https://doi.org/10.5334/gjgl.1292.
Husserl, E., 1936/70. The Crisis of European Sciences and Transcendental Phenomenology. In: Trans by D. Carr. Evanston. Northwestern University Press.
Karlson, F., 2007. Constraints on multiple center-emdedding. J. Linguist. 43, 365–392.
Katz, J., 1981. Language and Other Abstract Objects. Oxford University Press, Oxford.
Kertész, A., 2012. The 'Galilean style in science' and the inconsistency of linguistic theorising. Found. Sci. 17, 91–108.
Koyré, A., 1938/78. *Galileo Studies*. Trans. By John Mepham. Harvester Press, Hassocks.
Miller, G., Chomsky, N., 1963. Finitary models of language users. In: Luce, R.D., Bush, R., Galanter, E. (Eds.), Handbook of Mathematical Psychology, Volume II. John Wiley and Sons, New York, pp. 421–491.
Peacocke, C., 1986. Explanation in computational psychology: language, perception and level 1.5. Mind Lang. 1, 101–123.
Peacocke, C., 1989. When is a grammar psychologically real? In: George, A. (Ed.), *Reflections* on *Chomsky*. Blackwell, Oxford, pp. 111–130.
Perkins, L., Lidz, J., 2021. 18-month-old infants represent nonlocal syntactic dependencies. Proc. Natl. Acad. Sci. USA 118. https://doi.org/10.1073/pnas.2026469118.
Piccinini, G., 2008. Computation without representation. Phil. Stud. 137, 205–241.

[20] Thanks to two referees for construactive criticism. An early version of this paper was given at Philosophy and Generative Grammar in May 2022, organised by David Rey. I thank David and audience.

Piccinini, G., 2015. Physical Computation: A Mechanistic Account. Oxford University Press, Oxford.
Putnam, H., 1988. Representation and Reality. MIT Press, Cambridge, MA.
Pylyshyn, Z., 1984. Computation and Cognition: Toward a Foundation for Cognitive Science. MIT Press, Cambridge, MA.
Pylyshyn, Z., 1991. Rules and representations: Chomsky and representational realism. In: Kasher, A. (Ed.), The Chomskyan Turn. Blackwell, Oxford, pp. 231–251.
Quine, W.V.O., 1960. Word and Object. MIT Press, Cambridge, MA.
Quine, W.V.O., 1972. Methodological Reflections on Current Linguistic Theory. In: Davidson, D., Harman, G. (Eds.), Semantics of Natural Language. Reidel, Dordrecht, pp. 442–454.
Riemer, N., 2009. Grammaticality as evidence and as prediction in a Galilean linguistics. Lang. Sci. 31, 612–633.
Rey, G., 1997. Contemporary Philosophy of Mind: A Contentiously Classical Approach. Wiley-Blackwell, Oxford.
Rey, G., 2020. Representation of Language. Oxford University Press, Oxford.
Searle, J., 1992. Rediscovery of Mind. MIT Press, Cambridge, MA.
Soames, S., 1984. Linguistics and psychology. Ling. Philos. 7, 155–179.
Stabler, E., 1983. Are grammars represented? Behav. Brain Sci. 6, 391–402.
Stich, S., 1972. Grammar, psychology, and indeterminacy. J. Philos. 69, 799–818.
Turing, A., 1950. Computing machinery and intelligence. Mind 59, 433–460.
Weinberg, S., 1976. The forces of nature. Am. Acad. Arts Sci. Bull. 29, 13–29.
Wigner, E., 1960. The unreasonable effectiveness of mathematics. Commun. Pure Appl. Math. 13, 1–14.
Wootton, D., 2010. Galileo: Watcher of the Skies. Yale University Press, New Haven.
Zhang, N., 2009. Coordination in Syntax. Cambridge University Press, Cambridge.