



# Why is Generative Grammar Recursive?

Fintan Mallory<sup>1</sup> 

Received: 11 March 2021 / Accepted: 7 November 2021 / Published online: 24 November 2021  
© The Author(s) 2021

## Abstract

A familiar argument goes as follows: natural languages have infinitely many sentences, finite representation of infinite sets requires recursion; therefore any adequate account of linguistic competence will require some kind of recursive device. The first part of this paper argues that this argument is not convincing. The second part argues that it was not the original reason recursive devices were introduced into generative linguistics. The real basis for the use of recursive devices stems from a deeper philosophical concern; a grammar that functions merely as a metalanguage would not be explanatorily adequate as it would merely push the problem of explaining linguistic competence back to another level. The paper traces this concern from Zellig Harris and Chomsky's early work in generative linguistics and presents some implications.

## 1 Introduction

Why should a grammar be recursive? Recursion, we are told, is the fundamental property of natural language (Hauser et al. 2002).<sup>1</sup> The standard story is that grammars must be recursive if they are to capture the infinitude of natural languages—the fact that languages have unboundedly many sentences. This claim has come under scrutiny in recent years and no longer seems compelling enough to motivate any strong theoretical commitments. Does this undermine the claim that grammars should be recursive? This paper argues that it doesn't—because capturing infinitude was not the original reason why generative grammars were designed to be recursive. The original reason, which is both deeper and more serious, is independent of the debate surrounding infinitude. Even if there were a finite number of sentences in any

---

<sup>1</sup> I-language, to be exact. This is a claim about grammars and not sentences themselves, i.e., the claim that a grammar is recursive is separate from the claim that every sentence may be embedded in a further sentence.

---

✉ Fintan Mallory  
fintan.mallory@ifikk.uio.no

<sup>1</sup> University of Oslo, Oslo, Norway

language, or more plausibly, even if it were meaningless to talk about the number of sentences in a language, there may still be a good reason to characterise our knowledge of language this way. The central claim of this paper is that generative grammars were originally designed to be recursive because only a recursively formulated grammar could provide an analysis of what an ‘intelligent interpreter’ brought to understanding a language.

In the first section, I’ll present and criticise the familiar Infinitude Argument, the claim that a grammar must be specified recursively if it is to capture the unboundness of natural languages. I will then introduce the concerns about metalinguistic regress which motivated early generative linguistics and reconstruct a separate Explanatory-Recursion Argument.

Two points before starting.

1. I don’t mean to suggest that these philosophical concerns are what motivated a generation of linguists to adopt a new set of tools. The advance of generative grammar was not driven by philosophical arguments alone but by its descriptive power; the ability to describe phenomena like the English auxiliary system and unaccusative verbs, and to identify novel phenomena like control, raising, and island constraints. Nevertheless, I assume that a proper understanding of the philosophical underpinnings of the framework can be valuable when we contrast it to other systems of analysis.
2. In the following, I will do everything possible to avoid endorsing a particular grammatical framework. When this paper refers to a grammar being recursive, it means something like, specified in a formal system which contains an operation or function that is defined so that it can apply to its own output. This might be unification (e.g., HPSG, GPSG, LFG), merge (e.g., minimalism), substitution/adjunction (e.g., TAG), function application and  $\beta$ -reduction (e.g., categorial grammar) or something else.<sup>2</sup> The specific focus on Chomsky’s early work in this paper is due to his influence in the development of both the theoretical tools and metatheoretical goals of generative linguistics and is not intended to indicate that one grammatical framework is superior to others.

## 2 The Infinitude Argument

Let’s start with the familiar argument:

Knowledge: People know natural languages.<sup>3</sup>

Infinitude: Natural languages contain infinitely many sentences.

<sup>2</sup> This paper will do its best to avoid recent controversies about the use of the term ‘recursion’ in linguistics (though see Tomalin, 2011, and Lobina, 2014, 2017 for details). All the systems listed provide a description of some operation for building syntactic structures that can repeatedly apply to structures in order to incrementally generate more complex structures. They are also characterised with enough precision that they may be understood as providing a recursive enumeration of a set of syntactic structures when applied to an appropriately specified lexicon.

<sup>3</sup> If you really don’t like the word ‘know’ you can substitute the word ‘cognize’ as per (Chomsky, 1986).

Recursion: To finitely represent an infinite set, you must use a recursive device.

Conclusion: Therefore, people's knowledge of language must use a recursive device.

If you are inclined to minimalism, you can add an additional conclusion:

Minimalism: Methodological simplicity demands that we reduce knowledge of language to the simplest possible device for recursively constructing infinite sets, merge.

This argument is a relatively straightforward transcendental deduction. It begins with a claim about an object of cognition and derives some further claims about the necessary conditions for the possibility of representing that object. There is nothing intrinsically wrong with this; much like 'poverty of stimulus' arguments, transcendental deductions may have a role to play in the cognitive sciences. The main problem they face is the tendency of nature to give us a range of alternative conditions sufficient for the emergence of any given phenomenon, a tendency which manifests theoretically in different but extensionally equivalent theories. In any case, it's not the form of the argument that has been traditionally controversial but its second premise.<sup>4</sup>

The first cause for concern about the infinitude premise is its history. This premise started its life as an idealisation that was explicitly made to simplify the design of grammars and was only later treated as a basic property of language that must be explained.<sup>5</sup> Thus we have the shift from, '[I]n general, the assumption that languages are infinite is made for the purpose of simplifying the description. If a grammar has no recursive steps... it will be prohibitively complex' (Chomsky, 1956: 115–116), to the claim that, 'the most elementary property of the language faculty is the property of discrete infinity... this property is virtually unknown in the biological world' (Chomsky, 2000: 51).

Intuitively, idealisation and desideratum are on opposite ends of a methodological spectrum. Physicists don't attempt to capture how a surface in a model is frictionless (because real surfaces aren't) or how gas molecules are like billiard balls (because they aren't) and, while Turing's chemical theory of morphogenesis idealises cells as geometrical points, biologists needn't worry about how organs can be composed of ideal geometric points. These idealizations are decisions to be justified but not phenomena to be explained. Understood as an idealization, the issue of infinitude is an *external* question to linguistic theory. That is, it concerns how we formalize an everyday notion for the purposes of scientific inquiry but it is not something that

<sup>4</sup> In large part as a result of the papers in van der (Hulst, 2010), particularly, Pullum and Scholz, 2010 and Tiede and Stout, 2010. A recent discussion can be found in (Nefdt, 2019)

<sup>5</sup> This dual-status as both idealisation and desideratum can be found throughout generative writing. Cedric Boeckx in *Language in Cognition*, acknowledges that 'the property of infinity is a simplifying assumption. It's an idealisation (on a par with Turing's 'infinite tape'...), one that allows for simple questions to be asked, and hopefully simple theories to be constructed' (Boeckx 2010: 79) and in *Bare Syntax*, includes it among the four 'Basic desiderata' which any inquiry into the nature of the human language faculty must meet, writing that 'the key property of natural language syntax is discrete infinity' (Boeckx, 2008: 108).

can be deduced from pre-established facts.<sup>6</sup> It isn't the only principle in generative grammar that appears to be both a substantive claim and an idealisation. At times, it is unclear whether the *strong minimalist thesis*, the *uniformity principle* or the commitment to binary branching are discoveries or assumptions.<sup>7</sup> One may not take this to be a problem and instead reject the idea that there is a strong distinction between discovery and assumption since linguistics, like any other science, requires that we work with and within hypotheses.

If we accept that the infinitude premise is an idealisation, then we need to reinterpret our conclusions to reflect this. The most we would be permitted to conclude would be that our models of linguistic competence must use recursive devices, a claim which remains non-committal about ascribing the capacity for recursion to the language faculty. The minimalist conclusion, in turn, becomes a claim about the simplicity of our mathematical theory rather than the simplicity of a human cognitive system. This is much less interesting than a claim about the fundamental nature of the language faculty but may be more tenable.

It's not clear that we have a reason to treat it as more than idealization. While there have been some informal attempts to 'prove' that there are infinitely many sentences, the claim that there are infinitely many objects of any given kind rarely admits of non-circular proof. We can prove that a subset of some infinite set is infinite (e.g., the prime integers), but such claims ultimately depend on the assumption that there are infinitely many integers in the first place—something that follows from our definition of the integers. Consider the 'master argument' for the infinitude of language, 'There is no longest sentence (any candidate sentence can be trumped by, for example, embedding it in "Mary thinks that...")', and there is no non-arbitrary upper bound to sentence length. In these respects, language is directly analogous to the natural numbers' (Hauser et al., 2002). If, for any given sentence, one can always add some words, it would seem that languages are as extendable as the integers. As Pullum and Scholz point out, this is merely a statement of the claim that sentences can be continued unboundedly; it assumes exactly what is at issue (Pullum and Scholz, 2010).

The fact that this premise can't be deduced a priori is hardly damning and my point here isn't that we should think that languages are finite. I suspect most people would be happy to say that even if there are not strictly infinitely many sentences in a language, there are many more than could possibly be memorized, and therefore our

<sup>6</sup> Pullum presents the infinitude question as a red herring that has been used to motivate the generative conception of a grammar (Pullum, 2013). The alternative to the generative approach which treats a grammar as a device for recursively enumerating syntactic structures, is the model-theoretic approach which treats a grammar as a set of statements (i.e. constraints) which a language might either satisfy or not and which, importantly for Pullum, does not make any claims about the cardinality of language. There are many grammatical frameworks that employ constraints in some form but the purest example of model-theoretic syntax uses a formal language like (monadic) second-order logic to state the grammar.

<sup>7</sup> The Uniformity Principle states that 'In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances' while the strong minimalist thesis suggests that grammar provides an 'optimal solution' to the problem of mapping semantic/phonological information at different interfaces (Chomsky, 2001: 2).

knowledge of language will have to be represented recursively.<sup>8</sup> It is probably this weaker idea rather than any claims about infinitude that convinces most linguists. However, this claim faces its own problems. If we think that the need for recursion follows merely from the sheer number of sentences we could parse, we need to ask if every other capacity which involves cognizing or manipulating a large set of objects requires a recursive operation like merge. If we allowed for the same kinds of idealization that we permit for language, we could say that a human being can identify unboundedly many different fruits, dance unboundedly many different dances, and build unboundedly many different bagels. It is not a unique feature of the human language faculty that we can—when idealizing beyond physical and memory limitations—say that it exemplifies an infinite capacity. At the very least, when we abstract away from physical limitations, there is no clear bound on many of our capacities. For example, the lower bound for the number of possible games of chess restricted to 40 moves is around  $10^{120}$ . There are many more possible games of chess than atoms in the known universe and obviously more than can be memorised. So, either our capacity to play chess or build bagels must be recursive, or there is some number between the number of possible games of chess/kinds-of-fruit/bagels and the number of possible sentences at which recursive generation becomes a necessary component of a description of competence?<sup>9</sup>

If the answer is yes, then the recursion argument might be acceptable. If not, at the very least, we need to find a different premise to make it plausible. Assuming that there isn't such a number, there is no argument from the mere size of language to the claim that language is computationally unique or evolutionarily *sui generis* in the sense that it requires recursive devices while other complex abilities do not, and we should not conclude on account of the number of possible sentences that our knowledge of language requires methods not shared with our other capacities.

It isn't a serious problem that the arguments for the infinitude of natural language are circular. However, it does undermine the idea that they motivate the existence of a distinct and *sui generis* computational system underlying language (e.g., the 'faculty of language in a narrow sense' described in Hauser et al., 2002). If the argument for the use of recursive devices in linguistics relies on the infinitude claim, then it is not a strong argument. What is required is an argument that says that our knowledge of *language* must be represented by a recursive device, an argument that emphasizes the uniqueness of language and that doesn't appeal to idealizations that can't

<sup>8</sup> Pirahã allegedly has only finitely many sentences (Everett, 2012). However, it allows for unbounded apposition (Futrell et al., 2016: 17). The issue here appears to come down to how your theory of syntax handles apposition, which is beyond the scope of this paper.

<sup>9</sup> Defenders of the recursion argument are always in a position to either, deny that these activities are relevantly similar to language or argue that they are somehow grounded in the use of language. If one thinks that a computer can play chess, then it would seem that language is not a necessary condition for playing chess. Alternatively, one might say that a chess computer does use a (programming) language, in which case, the idea that complex abilities require recursion appears to collapse into a broader but less interesting commitment to the idea that computation is essentially linguistic. Either case just serves to show that it is not infinitude alone which motivates the use of recursive devices but infinitude combined with some further claims about the nature of language which distinguish it from other cognitive systems.

be non-circularly justified. In the next section, I will describe this argument and the role it played in the foundation of generative linguistics.

### 3 The Metalanguage Argument

Metalanguage arguments attempt to show that merely describing the structures of natural languages in a metalanguage supplied by grammatical theory leads to some kind of vicious regress. Even before infinitude became a theme in the development of generative linguistics, Zellig Harris had motivated his use of distributional methods by appealing to the idea that ‘language can have no external metalanguage, existing independently of the structure of language itself, in which the structure of language could be described’ (Harris, 1974).<sup>10</sup> While this is true in some senses, it isn’t immediately clear why it matters. To see this, it’s helpful to separate several different forms this argument might take.

**Ontogenetic metalinguistic regress:** Some languages are ‘specified’ into existence within a metalanguage. This is typical for formal languages like first-order logic or  $a^nb^n$ , as well as constructed languages like Jespersen’s Novial or Klingon. Natural languages cannot be like this *in general* as the language which specified them into existence would have to have been specified into existence and so on.

**Structural metalinguistic regress:** Facts about object languages are explained by facts about the grammars in which they are specified; they have the structures they do because a grammar dictates it. If a grammar is stated in an independent metalanguage, then facts about its structure must be determined by facts concerning the structure of the metalanguage and so on. Each explanation of linguistic structure pushes the question back to another level and so none is complete.

Harris used arguments like these to motivate an approach to linguistic structure that defined syntactic categories distributionally as patterns of co-occurrence in linguistic data, and then explained the emergence of linguistic structure probabilistically so that structure itself would be regarded as a deviation from equiprobability while structural relations between expressions correspond to the likelihood of co-occurrence. The details of Harris’s system aren’t the focus here. Instead, I want to consider the argument that motivated it. The idea that describing a language in a metalanguage leads to a regress. What is the nature of this regress and why should we care?

The first argument may have some significance for how we understand the evolution of language. Harris appealed to something like it to justify his claim that our account of language must explain how linguistic structure can emerge through processes of gradual self-organisation (Harris, 1976). Unlike formal languages, natural

<sup>10</sup> Early formulations of Harris’s point concentrated on the methodological danger of applying a fixed set of categories to different languages (Harris, 1951). Over time though the focus on metalanguages became more explicit. The discussion here draws from Harris’s own reflections on the dominant themes of his work (Harris, 1990). While such reflections should naturally be taken with a grain of salt, I believe that much of what he claims here was at the very least implicit in his earlier writings and explicit by the mid 60s. See (Ryckman, 2002) for an insightful account of Harris’s concerns with metalinguistic regress.

languages cannot have been stipulated into existence in general. While an individual language might have its structure specified within some cognitive system of representation (e.g., a language of thought), the structure of the language of thought could not have been specified into existence within some other language without leading to a regress.

If stipulation were the only model for how linguistic structure could emerge, it would seem that we have a serious problem of regress. However, stipulation is not the only model we have for the emergence of language and, in any case, the issue of whether the emergence of either the capacity for language or of individual languages themselves was sudden (as it would be if stipulated) or incremental is probably not an issue that can be determined by a priori reasoning.

One way of understanding the structural argument is in terms of whether the facts about a language's structure are endogenous or exogenous. Harris took the facts about linguistic structure to be endogenous, that is, they depend upon the distributional properties of the sentences of that language. It has since become more common to view these facts as exogenous. If the set of sentences of a language fail to generate a canonical structure (i.e., if we have multiple extensionally equivalent theories of about the structure of these sentences), then we must appeal to something outside of the set of sentences to fix the correct description. In the generative tradition, the dominant idea has been that grammatical structure is fixed by a mental grammar (sometimes defined with reference to an idealized speaker-hearer). On this account, a set of sentences does not have a canonical grammatical structure independent of a grammar. If we assume that structure is fixed by exogenous facts, then we cannot assume a hierarchy of metalanguages grounding those facts as the task would never be complete.

This argument is less innocent than it seems as it would demand a completeness from our descriptions of linguistic structure that we don't demand in other fields. We don't need to describe the syntax and semantics of the language of astrophysics to describe the structure of a pulsar. There is a strict sense in which the truth of the statements of astrophysics depends upon the syntactic and semantic properties of those sentences; sentences with different properties would lead to different claims. However, the propositions those sentences express are true or false irrespective of how they are expressed and so it would be an error to treat the structure of the object language as dependent upon the structure of the metalanguage. So assuming we can separate propositions from the sentences in which they are expressed, structural regress is not a serious challenge.

This brings us to a third metalinguistic regress which is much more serious than the other two.

Epistemic metalinguistic regress: To know a language, one must know the metalanguage in which it is specified, but to know this one would have to know the language in which that is specified and this leads to a regress. For example, to understand Jespersen's grammar of the English language, one must already understand the English in which it is written.

The rest of this section will argue that generative linguistics was developed in response to the threat posed by this kind of regress. Specifically, it will show that the objection to traditional grammars had nothing to do with their inability to capture

the infinitude of languages but was instead based on the claim that traditional grammars already assumed that anyone using those grammars grasped a natural language. As a result, traditional grammatical theory was unable to account for what it takes to know a language. The novel contribution of generative linguistics was how it accounted for an agent's knowledge of language without appealing to another language in which that knowledge was stated. This metalinguistic knowledge is often characterised as the agent's prior 'intelligence' or 'intuition'. We see this in the canonical definition of generative grammar:

'A grammar of a language purports to be a description of the ideal speaker-hearer's intrinsic competence. If the grammar is, furthermore, 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: 4).<sup>11</sup>

According to this definition, a generative grammar is a grammar that 'does not rely on the intelligence of the understanding reader'. There is no reference to infinitude here and the function of explicitness is not tied to any of its usual theoretical virtues but is introduced because it is a necessary condition for bypassing the need for human understanding. This was taken to address a problem in traditional grammars which had merely served as descriptive metalanguages. As the definition continues, '[p]erhaps we should call such a device a generative grammar to distinguish it from descriptive statements that merely present the inventory of elements that appear in structural descriptions, and their contextual invariants' (Chomsky, 1964: 9).

Some historical context should make this clearer. In 1961, Chomsky mentions that a grammar 'can' generate an infinite set of strings but no significance is made of this and no reference to infinitude is found among the list of requirements for a theory of grammar (Chomsky, 1961: 223). Instead, we find the idea that a grammar must characterize a function because 'it must be possible to determine what a particular grammar states about particular sentences without the exercise of intuition' (Chomsky, 1961). The appeal to 'intuition' remains the primary objection to traditional grammars in the following years. 'A traditional grammar has serious limitations so far as linguistic science is concerned. Its basic inadequacy lies in an essential appeal to what we can only call the 'linguistic intuition' of the intelligent reader' (Chomsky, 1962: 528). By 1964, we find the claim that the set of sentences of a language may be regarded 'for all practical purposes' as infinite, but again, the fundamental focus is on eliminating the role of 'intuition' from grammars. In 'Explanatory Models in Linguistics' (Chomsky, 1966b), traditional grammars are again criticised for appealing to the 'intuition' of an intelligent reader. It is clear from these early works which inaugurated the field of generative linguistics that the demand to eliminate the intuition required to interpret a grammar is what motivated the development of generative linguistics and not any claims about the infinitude of language.

---

<sup>11</sup> Note that neither this definition nor the other definitions we will see below make any reference to the logical notion of generation (recursive enumeration).



In any case, it would be inappropriate to object that traditional grammars failed to capture the infinitude of natural languages. There is absolutely nothing stopping us from taking the claims of Jespersen, Jakobson or Harris to hold for unboundedly many structures if one wanted to interpret them in that way. What matters is that this act of ‘interpretation’ itself would be left unanalysed. Reflecting on the works of Otto Jespersen, Chomsky claims, ‘In reality his [Jespersen’s] commentaries were not sufficient, because they appealed implicitly to the ‘intelligence’ of the reader to understand them and to use these examples and his often insightful commentary in the creation and comprehension of new forms, the reader had to add his own intuitive knowledge of language... This contribution of the intelligent reader, presupposed by previous grammars, must be made explicit if we hope to discover the basic principles of language. *This is the first goal of generative grammar.* In psychological terms, what is the nature of the intuitive unconscious knowledge, which (in particular) permits the speaker to use his language?’ (Chomsky, 1979: 109, emphasis added).<sup>12</sup> In other words, to understand the content of Jespersen’s claims, one must already understand English and so however accurate his account of English grammar is, it doesn’t provide an analysis of how one could know English grammar in the first place. An adequate grammar should account for the cognitive conditions which enable a reader to understand it.

The primary concern remains that, if we merely state a grammar in some metalanguage, we presuppose that we grasp such a metalanguage leaving unanswered the question of how a language is grasped. What is required is an account of how a metalanguage is to be interpreted. What Chomsky adds to Harris’s point is the idea that a grammar should explain the abilities required to interpret a language, something that is not captured by traditional grammars. Putting this point slightly differently, we might say that generative grammar was not developed to describe languages but to explain what it takes to understand a language at all and developments

<sup>12</sup> To labor this point: ‘It [generative grammar] attempts to go beyond traditional grammar in a fundamental way, however. As has repeatedly been emphasised, traditional grammars make an essential appeal to the intelligence of the reader. They do not actually formulate the rules of the grammar, but rather give examples and hints that enable the intelligent reader to determine the grammar, in some way that is not at all understood. They do not provide an analysis of the ‘*faculté de langage*’ that makes this achievement possible. To carry the study of language beyond its traditional bounds, it is necessary to recognize this limitation and to develop means to transcend it. This is the fundamental problem to which all work in generative grammar has been addressed’ (Chomsky, 1964: 11). ‘[I]t is important to bear in mind that even the more careful and complete traditional grammar relies in an essential way on the intuition and intelligence of the user, who is expected to draw the correct inferences from the many examples and hints (and explicit lists of irregularities) presented by the grammar. If the grammar is a good one, the user may succeed, but the deep-seated regularities of the language that he somehow discovers escape explicit formulation, and the nature of the abilities that enable him to perform this task remain a complete mystery’ (Chomsky, 1964: 16). ‘To return to the point, a generative grammar (that is, an explicit grammar that makes no appeal to the readers *faculté de langage* but rather attempts to incorporate the mechanisms of this faculty) is a system of rules that relates signals to semantic interpretations of these signals’ (Chomsky, 1966a: 12). While reference to the ‘intelligent reader’ became less frequent from the 70s onward, it has not disappeared: ‘As soon as serious descriptive work was undertaken, it was discovered that available accounts of language, however extensive, barely scratched the surface; even the most comprehensive grammar provided little more than hints that sufficed for the intelligent reader; the language faculty was tacitly presupposed (without awareness, of course)’ (Chomsky, 1994: 293).

in mathematical logic were seen to provide a solution to this. Importantly, all of these ideas are independent of the issue of whether natural languages have infinitely many sentences or whether this is a reasonable idealization to make. Even if natural languages were for some reason finite, we would still fail to provide an analysis of the capacity for language if we simply said that our finite languages are defined according to some grammar in our mind. To summarise, the infinitude claim is neither necessary for the introduction of recursive devices into linguistics and neither was it the actual basis for their introduction.

#### 4 The Explanatory-Recursion Argument

We need now to ask; how does a recursive device give an account of the intelligence a reader brings to interpreting a grammar? This is a difficult question, and I will limit myself to some tentative remarks. What the theory of recursive functions provides is a means of characterizing sets of objects without appealing a linguistic description of those sets. A recursive definition is complete in the sense that it determines an extension without relying on our pre-existing knowledge. It doesn't say what the objects have in common and it doesn't appeal to undefined concepts within an existing language. The result is that we can 'follow' a recursive definition to either determine if an object is a member of set without having any grasp of a concept that designates that set, or alternatively and more appropriately, we can use a function to recursively enumerate the members of a set without knowing the concept that picks out that set.<sup>13</sup> Such a concept needn't even exist in our everyday language. Without the need to grasp concepts, there is no need to grasp the logical relations in which those concepts stand, or to have any knowledge of a particular language or knowledge about the world. This is the unique contribution of the theory of recursive functions.

At this point, you might think that I am overplaying the significance of things. After all, aren't recursive functions defined within formal languages? For example, isn't the familiar Herbrand-Gödel schema for stating recursive definitions stated in something like first-order logic augmented with a theory of types (the language of *Principia Mathematica*)? To understand the significance of recursive definitions, it may be helpful to consider a similar concern that struck Gödel when he was attempting to generalize his incompleteness proofs beyond the system of the *Principia*. For several years after the incompleteness proofs were published, it was still hoped that incompleteness may be traced to particular properties of the system of the *Principia*, for example, the theory of types. What was required to generalize the proofs beyond this system was an account of what it meant for a function to be computable that was not tied to any particular formalism. Gödel eventually took this requirement to be satisfied by Turing's analysis of computation.

---

<sup>13</sup> While grammars may recursively enumerate (generate) sets of syntactic structures, they do not constitute recursive definitions in the first sense, i.e., they do not recursively define the characteristic function of the set of structures.

One way of understanding the significance of Turing's account of computation, is that it give us the means to step out of the regress of languages into something that is not itself a language but that we can view as formally equivalent to one.<sup>14</sup> A machine is nothing but a set of states, symbols, and actions and, while it is possible to describe a machine in some language, it is not necessary to do so for the machine to exist. The machine can be *implemented* rather than described and though description occurs within a language, implementation need not. Any rules that could be stated by a grammar could be implemented by a machine. By breaking the process of following a rule and responding to an input into its component actions, Turing managed to explain what was required to use a formal system in a way that doesn't presuppose a knowledge of a language (i.e. intuition). A system doesn't need to grasp what a function is to be accurately characterised as implementing one. Of course, many machines will implement some kind of machine language, but the point remains that Turing's analysis shows how these can bottom-out with the dispositional properties of implementable computational states rather than another set of statements. In other words, he provided the first explanatory analysis of a language in terms of something that was not a language.<sup>15</sup> Furthermore, every Turing machine computable function is recursive and vice-versa. As a consequence, if a grammar can recursively enumerate the syntactic structures of a language, this should be sufficient to explain our knowledge of language without assuming the existence of an independent metalanguage.<sup>16</sup>

In short, the question isn't, how do you represent an infinite set but how do you represent *any* set without using a language to do so? Generative grammar provides us with a means of recursively enumerating the syntactic structures of a language ensuring that the syntactic structure of each sentence can be provided without relying on any metalinguistic statements which themselves would require interpretation.

This is as much the case for model-theoretic approaches to syntax (MTS) as it is for more traditionally generative systems. MTS states a grammar as a set

<sup>14</sup> While there are a range of equivalence proofs linking automata to languages, the most relevant in this case is probably Kuroda's demonstration of the equivalence of context-sensitive languages and linear bounded automata (Kuroda, 1964).

<sup>15</sup> The idea that Turing had 'for the first time succeeded in giving an absolute definition of an interesting epistemological notion, i.e., one not depending on the formalism chosen' is discussed in (Gödel, 1946 and 1957). It was in virtue of Turing's analysis that Gödel was willing to generalize his incompleteness claim beyond the system of *Principia Mathematica* since to do so required 'a precise and unquestionably adequate definition of the general notion of formal system'. For a helpful discussion of this point, see (Sieg, 2006) and (Chen, 2017). The discussion to the Church-Turing thesis in this section is obviously a bit of a simplification. For a history of the development of the ideas here, (see Soare, 1996) and for a more recent and controversial discussion of some of these issues in the context of generative grammar, see (Lobina, 2017). While there is a striking resemblance between Chomsky's attempts to eliminate 'intuition' from linguistics and Frege's attempts to eliminate 'intuition' from mathematics, a full account of these similarities is beyond the scope of this paper.

<sup>16</sup> For example, claims like: 'The second component of UG, the theory of generative procedures, has been amenable to study, really for the first time, since the mid-twentieth century. By then the work of Gödel, Turing, Church, and others had established the general theory of computation on firm grounds, making it possible to undertake the study of generative grammar with a fairly clear understanding of what is involved' (Berwick and Chomsky, 2016: 91).

of constraints that license certain syntactic structures. In their purest form, these constraints are written in a formal language like monadic second-order logic. Yet the constraints don't say anything unless we have a means of interpreting them. The formulas of a logical language are marks on a page without a semantic theory putting them to use and this semantic theory will almost inevitably invoke a recursively defined interpretation function to map the statements of the grammar onto objects in some domain (e.g., multi-dimensional tree domains as in Rogers, 2003). Model-theoretic syntax does not escape the need to use recursion in its description of language. This is simply a formal analogue of the fact that a description of a language, no matter how clearly it is written, doesn't tell us anything unless we know how that description is to be interpreted, and this interpretation will require recursion in some form. The difference between model-theoretic and generative grammars is whether one places the recursively defined function at the heart of the theory—like unification or merge—or whether it is in the metalanguage stating the semantics of your model-theoretic grammar.

We are now in the position to formulate a different argument for the use of recursive devices in linguistics which is both more plausible and a more accurate reflection of the motivations behind early generative linguistics.

1. A grammar merely composed of metalinguistic statements would not be explanatorily adequate as we would require some account of how those statements are to be interpreted. If grammars were limited to such statements, they would generate a regress.
2. The theory of recursive functions provides us with a means to avoid this regress as a recursively defined function can enumerate the syntactic structures of a language without appeal to statements in a metalanguage (representing the knowledge an intelligent agent brings to interpretation).
3. Therefore, a suitably supplemented recursive grammar will be able to represent linguistic competence.

While I think this is a better argument than the recursion argument above it does have several weaknesses. The infinitude argument gave us a necessity claim; our grammar *must* be recursive if it is to capture the ideal infinitude of languages. The explanatory recursion argument gives us a weaker, partial-sufficiency claim. There may be alternative ways to characterise an agent's knowledge of language that don't appeal to a recursive specification of the grammar. Further, if some kind of Language of Thought Hypothesis (Fodor, 2008) is correct, then the grammar of a natural language might be stated in a mental metalanguage. Of course, this metalanguage would require a grammar but it's always possible that this grammar is the generative one (or maybe the grammar up from that). There is no a priori reason to believe that the natural language grammar must be the only level of representation. What matters is that the regress terminates somewhere. One virtue of the idea that the mind contains a hierarchy of grammars is that a syntactically less expressive language can state the grammar of a strictly more

expressive language (e.g., the grammars of recursively enumerable languages can be stated in context-free languages).

Alternatively, you might think that the first premise is grounded in an outdated conception of the range of computational models. Constraint-based systems like Optimality Theory have a relatively well-understood computational implementation in connectionist networks. It may be possible for such approaches to give full credit to the idea that a grammar can be stated in constraints while simultaneously explaining those constraints through appeal to something we may think of as non-symbolic such as vectors in a neural network. The first premise of the explanatory recursion argument seems to presuppose some form of symbolic realization of metalinguistic statements which may have been plausible in the early 60 s but now appears outdated.<sup>17</sup>

A final point is that the explanatory-recursion argument doesn't really support the idea that language is *sui generis*. While the infinitude argument (or at least an assumption behind the argument) suggested that language was unique among human capacities for its ability to give rise to infinitely many sub-capacities, the explanatory-recursion argument gives us something that looks much more like the standard case for a computational model of mind.

## 5 Conclusion

In this paper, I have extended previous criticisms of the infinitude argument. The upshot of these is that there seems to be little principled reason to claim that language is unique in its unboundedness as any claims for this rely upon idealisations which could just as well be applied to any cognitive or physical system. I have also suggested why the focus on infinitude is a distraction. Textual evidence shows that such claims played little role in the development of generative linguistics. Instead, we have seen that generative grammar was introduced to account for the 'intuition' that a reader brought to understanding a traditional grammar. It is because understanding a traditional grammar requires understanding the language in which it was written that traditional grammars do not provide adequate accounts of linguistic competence but instead generate epistemic regresses. If this paper is correct, the solution to these regresses was found in the theory of recursive functions as developed by Post and others which provided for the first time the means of characterising language by means of something non-linguistic. Or in a slogan, making linguistic use of non-linguistic means.

**Funding** Open access funding provided by University of Oslo (incl Oslo University Hospital).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as

---

<sup>17</sup> See (Lappin, 2021), for an overview of the current capacities of deep neural networks for learning syntactic structures.

you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Berwick, R. C., & Chomsky, N. (2016). *Why only us?* The MIT Press.
- Boeckx, C. (2008). *Bare syntax*. Oxford University Press.
- Boeckx, C. (2010). *Language in cognition*. Wiley-Blackwell.
- Chen, L. (2018). *Interpreting Gödel: Historical and Philosophical Perspectives*. PhD Thesis, King's College, London.
- Chomsky, N. (1956). Three models for the description of language. *RE Transactions of Information Theory*, 2(3), 113–124.
- Chomsky, N. (1961). On the Notion 'Rule of Grammar' (in R. Jakobson (Ed.) *Proceedings of Symposia in Applied Mathematics 12: Structure of Language and its Mathematical Aspects* (pp. 6–24). Providence, RI: American Mathematical Society)
- Chomsky, N. (1964). *Current Issues in Linguistic Theory*. Mouton.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press.
- Chomsky, N. (1966a). *Topics in the theory of generative grammar*. Mouton.
- Chomsky, N. (1966b). Explanatory Models in Linguistics. *Studies in Logic and the Foundations of Mathematics*, 44, 528–550.
- Chomsky, N. (1979). *Language and Responsibility*. Pantheon Books.
- Chomsky, N. (1986). *Knowledge of Language*. Praeger.
- Chomsky, N. (1994). On Linguistics and Politics (interview with Günther Grewendorf). *ProtoSociology*, 6, 293–303.
- Chomsky, N. (2001). Derivation by phase (in M. Kenstowicz, (ed.) *Ken Hale: A life in language*, (pp. 1–52). Cambridge, MA: MIT Press)
- Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*. Cambridge University Press.
- der Hulst, V. (Ed.). (2010). *Recursion and human language*. De Gruyter Mouton.
- Everett, D. (2012). What does Pirahã grammar have to teach us about human language and the mind? *Wires Cogn Sci*, 3, 555–563. <https://doi.org/10.1002/wcs.1195>
- Fodor, J. A. (2008). *LOT 2: The Language of Thought Revisited*. Oxford University Press.
- Futrell, R., Stearns, L., Everett, D.L., Piantadosi, S.T., Gibson, E. (2016). A Corpus Investigation of Syntactic Embedding in Pirahã. *PLoS ONE* 113
- Gödel, K. (1946). Remarks before the Princeton Bicentennial Conference on Problems (in S. Feferman, J. Dawson, S.C. Kleene, G.H. Moore, R.M. Solovay, and J. van Heijenoort, (eds.) *Mathematics. In Kurt Gödel: Collected Works, Vol. II*, pp. 150–53. Oxford: Oxford University Press)
- Gödel, K. (1957). Letter to Ernest Nagel of 25 February 1957. (in S. Feferman, J. Dawson, S.C. Kleene, G.H. Moore, R.M. Solovay, and J. van Heijenoort, (eds.) *Kurt Gödel: Collected Works, Vol. V*, (pp. 145–146) Oxford: Oxford University Press)
- Harris, Z. (1951). *Methods in Structural Linguistics*. University of Chicago Press.
- Harris, Z. (1976). On a theory of language. *The Journal of Philosophy*, 73(10), 253–276.
- Harris, Z. (1991). *A Theory of Language and Information: A mathematical approach*. Oxford & New York: Clarendon Press.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science*, 298(5598), 1569–1579.
- Kuroda, Y. (1964). Classes of Languages and Linear Bound Automata. *Information and Control*, 7, 207–223.
- Lappin, S. (2021). *Deep learning and linguistic representation*. CRC Press.
- Lobina, D. (2014). When linguists talk mathematical logic. *Frontiers in Psychology*, 5, 1–3.
- Lobina, D. (2017). *Recursion: A Computational Investigation into the representation and processing of language*. Oxford University Press.
- Nefdt, R. (2019). Infinity and the Foundations of Linguistics. *Synthese*, 196(5), 1671–1711.

- Pullum, G. K. & Scholtz, B. (2010). Recursion and the Infinitude Claim (in van der Hulst (ed) *Recursion and Natural Language*, (pp.1–9) New York: De Gruyter Mouton).
- Pullum, G. K. (2013). The central question in comparative syntactic metatheory. *Mind and Language*, 284, 492–521.
- Rogers, J. (2003). wMSO theories as grammar formalisms. *Theoretical Computer Science*, 293, 291–320.
- Ryckman, T. (2002). Method and theory in Harris's Grammar of Information in (B. E. Nevin (ed.) *The Legacy of Zellig Harris: Language and information into the 21st century. Volume 1: Philosophy of science, syntax and semantics* (pp. 19–37) Amsterdam: John Benjamins Company
- Sieg, W. (2006). Gödel on computability. *Philosophia Mathematica*, 14(2), 189–207.
- Soare, R. (1996). Computability and recursion. *Bulletin of Symbolic Logic*, 2(1996), 284–321.
- Tiede, H. & Stout L. (2010). Recursion, infinity, and modelling in (van der Hulst (ed) *Recursion and Natural Language*, H. (pp.147–158) New York: De Gruyter Mouton)
- Tomalin, M. (2011). Syntactic structures and recursive devices: A legacy of imprecision. *Journal of Logic, Language and Information*, 203, 297–315.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.