



RESEARCH ARTICLE

Pullum's philosophy of linguistics: towards a unified framework

Ryan Mark Nefdt^{1,2} 

¹University of Cape Town, Rondebosch, South Africa

²University of Bristol, UK

Email: ryan.nefdt@uct.ac.za

Received: 10 January 2024; **Revised:** 26 May 2024; **Accepted:** 20 June 2024

Keywords: Pullum; philosophy of linguistics; linguistic infinity; model-theoretic syntax; normativity

Abstract

Geoffrey Pullum has produced countless contributions to linguistic theory over his 50-year career in the field. Given this exceptional scientific achievement, his philosophical work often goes underappreciated. In this article, I discuss and critique three themes from Pullum's philosophy of linguistics, namely, cardinality neutrality, model-theoretic syntax and normativity in language. I conclude by showing how these seemingly disparate elements might indeed be connected in terms of a normative constructivist approach to linguistics.

1. Introduction

The philosophy of linguistics has emerged as a robust philosophy of science (Scholz et al. (2022)). Prominent topics include the ontology of language, realist versus antirealist conceptions of standard linguistic entities (Rey (2020)) and the commensurability of competing theoretical frameworks (Pullum (2013)). Geoffrey Pullum has contributed not only to the theory of language and formal linguistics but also to its philosophical analysis. The three main themes of his philosophy of linguistics have been the metatheory of formal syntax, the role and reality of linguistic infinity and the normative force of linguistic rules. In this article, I explore each of these topics in turn with a special eye toward a unified philosophy of linguistics along constructivist and even possibly finitist lines.

In [Section 2](#), I outline the basics of the model-theoretic alternative to generative syntactic theory. I argue that Pullum is too quick to draw the theoretical consequences he does from this model of grammar, from the perspective of the philosophy of science. In [Section 3](#), I discuss one particularly poignant such consequence in his rejection of the linguistic infinity postulate. Here, I suggest, with Langendoen (2010), that Pullum & Scholz (2010) provide only the first half of the critique. In [Section 4](#), I address Pullum's recent claims as to the normative grounding of linguistic theory (Pullum (2018, forthcoming)). Finally, in [Section 5](#), I motivate a novel unification of these themes.

2. A model theory for syntax

The origins of modern generative grammar are proof-theoretic in nature (Tomalin (2006), Pullum (2011), Lobina (2017)). This involves the recursive enumeration of a set by means of a finite set of rules over a finite vocabulary. To be a well-formed formula or grammatical string of a language is to be generated by the rules over the vocabulary, that is, to be an output of the grammar. Phrase structure grammars, categorial grammars, minimalist grammars and many others fall within this broad mathematical framework.

Model-theoretic syntax (MTS) replaces the centrality of that proof-theoretic notion of syntax with a concept associated more with semantics, drawing from first-order logic and model theory. Grammar formalisms such as Head-Driven Phrase Structure Grammar (HPSG) (Müller et al. 2021) and Sign-based Construction Grammar (Michaelis 2009) are prominent examples of this approach. They are also contrasted with the derivational approach of generative grammar. Instead of a grammar generating the well-formed formulas of a language, a grammar models the rules of that language via constraints. ‘An MTS [model-theoretic syntax] grammar does not recursively define a set of expressions; it merely states necessary conditions on the syntactic structures of individual expressions’ (Pullum & Scholz 2001: 19). In this sense, a sentence or expression is well-formed iff it is a model of the grammar (defined in terms of constraints which act as the axioms of the formalism). A grammar is then just a finite, unordered set of MTS rules, which in turn are just statements about the structure of expressions like ‘every direct object noun phrase in a transitive verb phrase immediately follows the verb’ (Pullum 2007). To be a model of the grammar is to be an expression that satisfies the grammar or meets its constraints.

Pullum (2013) draws out a number of comparative advantages of adopting an MTS framework over a generative one. In fact, he states that ‘[t]he question I regard as most central, in that it arises at a level more abstract than any comparison between rival theories, concerns the choice between two ways of conceptualizing grammars’ (Pullum 2013: 492). The general thrust of the argument is to establish the superiority of MTS over generative syntax on a number of significant theoretical desiderata, such as incremental grammaticality, incorporation of neologisms and language acquisition (see Steedman (this volume) for more on these). What is interesting from our perspective is the property of ‘cardinality neutrality’.

Grammars of this sort [MTS] are entirely independent of the numerosity of expressions (though conditions on the class of intended models can be stipulated at a meta-level). For example, suppose the grammar of English includes statements requiring (i) that adverb modifiers in adjective phrases precede the head adjective; (ii) that an internal complement of know must be a finite clause or NP or PP headed by of or about; (iii) that all content clause complements follow the lexical heads of their immediately containing phrases; and (iv) that the subject of a clause precedes the predicate. Such conditions can adequately represent facts like those in (1). But they are compatible with any answer to the question of how many repetitions of a modifier an adjective can have, or how deep embedding of content clauses can go, or how many sentences there are. The constraints are satisfied by expressions with the relevant structure whether there are infinitely many of them, or a huge finite number, or only a few (Pullum & Scholz 2010: 123).

I think Pullum might be reading too much into the structure of model-theoretic grammars here. Admittedly, cardinality neutrality might be the more conservative interpretation.

However, as Langendoen & Postal (1984) argue somewhat controversially in their ‘vastness proof’, non-denumerable infinity of language might also warrant model-theoretic treatment.

Nevertheless, I am going to take a different route here. Rather than emphasise the differences between generative and model-theoretic grammars, I am going to highlight their similarities. In fact, according to the scientific modelling perspective proffered by Tiede & Stout (2010) and Nefdt (2019), these are merely alternative models of grammar potentially modelling the same linguistic reality. Pullum is correct that a divergence in ontological commitment on a particular property might occasion neutrality and the possibility that the given property is an artefact of the model. However, the history of science has shown us that artefacts of models can prove to be genuine aspects of the target. For example, multiple revisions, in terms of physical interpretations, of the same mathematical formalism in classical mechanics led to the discovery of the positron (Bueno & Colyvan (2011)). Dirac initially considered negative energy solutions to be mere features of the mathematical model. However, after finding physical interpretations of these solutions, he was led to revise his entire theory and predict the existence of a novel particle.

In terms of convergence, formal language theory, the mathematical theory of the structure of language, was concerned with defining languages of nested complexity known as the Chomsky hierarchy. One way of doing this is by defining languages in terms of the production rules of grammars of certain varieties. For instance, regular grammars or familiar finite-state grammars were argued by Chomsky (1956) to be inadequate as models of natural language syntax since they cannot capture certain common linguistic patterns (such as $a^n b^n$). The context-free languages contain regular languages as a proper subset (their grammars can generate those languages) as well as more complex structures. Similarly, as we move up the hierarchy, we approach the context-sensitive class and, above that, the recursively enumerable, which subsumes all those below. Each grammar type can be further associated with a type of recognising machine or automaton, with the last ring corresponding to a Turing machine. The specifics, however fascinating, will not detain us here.¹

The important part for our purposes is that this nested hierarchy of formal languages used to model the grammatical rules and features of natural languages can be defined in more than one way. Classical formal language theory pursued this task in terms of grammars and generated classes of formal languages. But there is a model-theoretic alternative. For a flavour of the results: Büchi (1960) showed that a set of strings forms a regular language if and only if it can be defined in the weak monadic second-order theory of the natural numbers with a successor. Thatcher & Wright (1968) then showed that context-free languages ‘were all and only the sets of strings forming the yield of sets of finite trees definable in the weak monadic second-order theory of multiple successors’ (Rogers 1998: 1117). It seems that the Chomsky Hierarchy is not immune to equivalent model-theoretic characterisation.

At the slightly lower level, various specific grammatical formalisms under the umbrellas of the generative grammar and model-theoretic syntax, respectively, have been shown to be weakly equivalent or characterise the same sets of strings (Vijay-Shanker & Weir 1994). The differences between derivational approaches and declarative (or constraint-based) ones have sometimes even been assumed to be only apparent or ‘notational variants’ (Chomsky 2000b). This might be too strong a claim.² Nevertheless, herein lies the argument: if

¹ See Levelt (1974) for a classic introduction and Hunter (2021) for a more recent treatment.

² See Borsley & Müller (2021) for a rigorous comparison of Minimalism and HPSG.

generative syntax and MTS can characterise the same formal systems, output the same stringsets and converge on various characterisations of linguistic phenomena, then there seems to be some invariant structure across these formalisms. In other words, they might be more compatible and less obvious in competition than Pullum acknowledges.³

Identifying the common structures could give us traction on which parts of the models are genuinely reflective of the target, that is, natural language, and which parts are merely model artefacts. However, some features of particular models might indeed reflect linguistic reality. The argument from model theory to linguistic cardinality is missing a step. At this point, we move on to Pullum's argument to the effect that the linguistic infinity postulate is unsupported by the philosophy of linguistics.

3. Pullum on Linguistic Infinity

The idea that natural language is somehow an infinite capacity or even an infinite set has become more than just a metaphor in contemporary linguistics. It has informed positions on the kind of science linguistics is (Katz 1996, Chomsky 2000a, Postal 2009), views on what language itself might be (Chomsky 1995) and strong claims about the evolution of human language (Berwick & Chomsky 2016). It is also often linked with debates on linguistic creativity, compositionality and recursion (Chomsky 1980, Hauser et al. 2002).

There are a number of popular avenues for establishing linguistic infinitude (although it is often just stated without argument). The first, and earliest, attempts are proof-theoretic reconstructions of an insight, attributed to Willem von Humboldt, that language is 'infinite use of finite means'. This strategy relies on the concept of recursion and produces an analogy with the natural numbers. In the following section, I will explore this option, with special emphasise on the critique produced by Pullum & Scholz (2010).

3.1. The argument from recursion and generative grammar

The literature on recursion in linguistics is complicated. Specifically, it is unclear at which level (or even which object) the property is meant to be interpreted. I cannot embark on a mission to disentangle the many vines of this particularly thorny issue here. Basically, the mathematical definition involves a property of self-reference, usually in two steps: one that specifies the condition of termination of the recursion (or the base case) and the recursive step, which reduces all other cases to the base (Tomalin (2006)).⁴ For example, addition in Peano arithmetic can be defined recursively as $x + 0 = x, x + Sy = S(x + y)$ (where 'S' denotes the 'successor of' one place relation). Similarly, the rewrite rules for a context-free grammar are recursive: $S \rightarrow ab$ or $S \rightarrow aSb$ (where 'S' is now a unique start symbol or category and 'a' and 'b' are terminal elements like words). In early generative grammar, phrase-structure rules were instructions to rewrite symbols or strings into other strings, such as $S \rightarrow NP VP$ (Chomsky 1957). In fact, as noted in Pullum (2011), generative grammars are

³ Johnson (2015) has something similar in mind when he connects notational invariance in linguistics (when grammars agree on structure) to symmetries in physics, arguing that they 'determine the empirically 'real' or 'meaningful' structure of any one of the theories taken individually' (164).

⁴ Lobina (2017) further distinguishes recursion from iteration, where the latter is a cost-saving mechanism that grounds the former in the case of linguistic cognition.

special cases of Post production systems, which are proof-theoretic devices capable of characterising the recursively enumerable sets. More importantly, they are finite rule systems that can generate infinite sets, hence fulfilling von Humboldt's prophecy.

This brings us to the first argument in favour of discrete infinity, that is, that generative grammars output infinite expressions or grammatical structures. If a grammar G is a set of finite rules for defining stringsets (or sequences of symbols), then a language is the set of all the strings that G generates. In Post production system style, ' G will be said to generate a string w consisting of symbols from Σ [the alphabet] if and only if it is possible to start with S and produce w through some finite sequence of rule applications' (Jäger & Rogers 2012: 1957). What does this mean for linguistic infinity? Well, it comes down to what is known as the 'membership problem'. Generative grammars have sharp boundaries. Some string w is either generated (or derived) by the rules of G or it is not. It is either in the language or it is not. There is no in between. This means that the language that the grammar generates 'has an exact cardinality: it contains either some finite number of objects or a countable infinity of them' (Pullum 2013: 496). Furthermore, the reason for opting for a countable infinity specifically comes from the data of natural language itself. As Chomsky states,

[T]he rules of the grammar must iterate in some manner to generate an infinite number of sentences, each with its specific sound, structure, and meaning. We make use of this "recursive" property of grammar constantly in everyday life. We construct new sentences freely and use them on appropriate occasions (Chomsky 1980: 221).

The best way to mount an analogy between language and arithmetic is by establishing a connection between the set of natural language expressions and the set of natural numbers. Firstly, we need an analogy with the successor function in Peano arithmetic. Recall that the successor function applies to an object in the set and produces another object in that same set endlessly. Thus, we would need a similar kind of operation in language. The usual candidates come from syntax. Consider the following examples:

- (1) (a) The linguist is tired.
- (b) The linguist is very tired.
- (c) The linguist is very very tired.
- (d) The linguist is very very very tired.

The idea is something like: English syntax is 'closed under adverbial modification'. In other words, if (a) is in the set of English sentences or, if you prefer, generated by the rules of English, then so are (b), (c), (d) and so on *ad infinitum*. The same strategy can be employed for other common examples involving subordination, centre embedding, conjunction etc. Pullum & Scholz (2010) call this the 'no maximal length claim' (NML) or the claim that there is no longest English sentence (i.e. you can always create a longer one). The problem they point out is that to establish this analogy, one would have to assume both that every English expression has a successor, and that this function is injective somehow. But doing so would essentially be begging the question. What they do not show is how to account for either the syntactic facts or the intuition that prompts claims like NML. Instead, they offer three different alternative conceptions of grammar that are neutral on the cardinality question, such as the model-theoretic account discussed above.

However, it should be noted that Chomsky (2008) does attempt to make a direct argument concerning the Merge operation of Minimalism and arithmetic. He outlines the following procedure for mimicking the successor function within linguistics.

Suppose that a language has the simplest possible lexicon: just one LI [lexical item], call it “one”. Application of Merge to that LI yields {one}, call it “two”. Application of Merge to {one} yields {{one}}, call it “three”. Etc. In effect, Merge applied in this manner yields the successor function. It is straightforward to define addition in terms of Merge(X, Y), and in familiar ways, the rest of arithmetic (138).

The idea is that Merge applied over a lexicon with a single entry results in the successor function. Let us set aside for the moment the issue of whether Merge can operate on one element as it is usually defined in binary terms (perhaps Merge ('one', \emptyset) is assumed). The problem is that Merge, or the set-theoretic operation of taking two objects and creating a new labelled unit with the two objects as (unordered) members, has nothing to do with language in and of itself.⁵ Making the argument that natural language syntax can be characterised with this simple operation is a separate task (one that Minimalists have taken up for 20 years). But merely showing that you can model the successor function with a recursive set-theoretic operation is trivial. Merge, so stated, can also model the heartbeat of mammals, the rhythm of drums or the barking of dogs. Nothing about these phenomena is intrinsically connected to language or the axiomatisation of arithmetic, for that matter.

This brings us to an important assumption about the nature of generative grammars imbued with recursive rules, namely, that they necessarily yield or entail a discrete infinity of linguistic outputs. Again, Pullum & Scholz (2010) show that this entailment does not hold by means of a toy example using a context-sensitive grammar. The grammar contains a recursive rule ($VP \rightarrow VP VP$), which is only capable of a single application given the specified lexicon but nevertheless remains non-trivially recursive. Thus, the recursive element of the grammar does not ensure infinite output (even if it can create infinite structures).

3.2. Two options for infinitude

Despite the convincing challenge laid out by Pullum & Scholz (2010), there remains a number of options for motivating a natural language infinitude. I will discuss one prominent such option and another less discussed alternative. Both seem to require some notion of ideal competence. For instance, Huybregts (2019) insists that infinite cardinality is the default or null hypothesis of formal grammars. He specifically challenges the idea that there could have been a step-wise or gradualist approach to infinite language along a cultural or evolutionary path (starting out in linguistic finitude of some sort). He reiterates NML as a datum. What is interesting is where he locates the infinitude. He states '[w]e should be careful, however, not to confuse the infinite productivity of the generative procedure (a function in intension) with the cardinality of the decidable set of structures it generates (the well-formed formulas defined in extension)' (Huybregts 2019: 3).⁶ Appreciating this caveat allows for

⁵ See Mukerji (2022) for the same admission towards a very different conclusion.

⁶ Pullum has repeatedly insisted that the idea that generative grammars are functions (which take arguments) has little coherence.

explanations of languages with finite output or no overt recursive constructions, without giving up on infinity. Hence, the counterexample provided by Pullum & Scholz (2010) is neutralised, since the grammar ‘may generate an infinite language but only produce a finite subset of it’ (Huybregts 2019: 3).

In a related fashion, Everaert et al. (2015) have re-emphasised that structures and not strings are the driving force behind mental linguistic computations. This harks back to the distinction between the strong generative capacity of grammars (the structures or ‘structural descriptions’ they produce) and the weak generative capacity (the stringsets they generate). As Chomsky (1965: 61) notes, ‘discussion of weak generative capacity marks only a very early and primitive stage of the study of generative grammar’. He goes on to claim that ‘questions of real linguistic interest’ are those pertaining to the strong generative capacity or descriptive adequacy as well as explanatory adequacy (i.e. language acquisition).⁷

Another strategy, which dovetails with the scientific modelling perspective discussed above, is given in Langendoen (2010). He describes the situation in the following way.

For any natural language L , let L^w represent the finite set of expressions known to belong to L on the basis of such direct evidence. Given that L^w provides indirect evidence for genuinely recursive size-increasing operations, standard generative models project a denumerably (countably) infinite set L° of ‘possible’ members of L . By not distinguishing the models from the language, proponents of such models conclude that $L = L^\circ$ and so is infinite. As noted above, that conclusion may be correct, but an argument is still needed to show that the models do not overgenerate. In the absence of such an argument, all we can conclude is that L lies somewhere between L^w and L° , and so may be either finite or infinite.

To establish a cardinality claim, Langendoen relies on an antecedent (from Sapir) of Katz’s ‘effability principle’: that every language can represent every meaning or proposition. Sapir’s notion of ‘formal completeness’ is based on mathematically ‘complete’ systems, such as geometry or arithmetic. The idea is that every natural language possesses a complete system of reference, independently of the resource (or cognitive) limitations of human cognisers. Just as geometry needs to have the resources to characterise any possible space, natural languages need the resources (in principle) to characterise any thought or concept. What this means for the infinity claim is that closure operators (like adverbial modification) can be assumed to get us from L to L° in the natural language case by the same logic that takes operators like the successor function from direct evidence of numerosity to countable infinity.⁸

Langendoen then makes an almost Lewisian move (in Lewis (1995[2010])) in allowing for different languages to be modelled with different cardinalities. This process is determined by the needs of the linguistic community. When a particular language only requires a subset of the formal expressive power of another, then that language might indeed be considered finite in comparison. This would distinguish his view from Huybregts (2019). However,

⁷ I thank an anonymous reviewer for stressing this point to me.

⁸ Of course, linguistics, geometry, arithmetic and other sciences might just be utilising infinite idealisations, a common practice in various fields. Savitch (1993) argues that a discrete infinity of languages can be assumed as a tool for smoothing the mathematics. Nefdt (2019) shows that similar techniques are used in fields like population genetics, in which the target domain is clearly finite in magnitude.

I think it would also move us further from the essence of Katz's effability principle and similar universalist conceptions of the commonality of human language. In terms of our dialectic, it is not clear that Pullum would accept that some languages should be modelled with generative grammar and others with MTS, resulting in differing cardinalities.

The place at which we have arrived is one in which Pullum's critique compels adherents of linguistic infinity to motivate their claims more perspicaciously. This is to the clear benefit of the philosophy of linguistics. However, his arguments do not quite establish that natural languages are not discretely infinite by themselves.

4. Normativity

The last component of Pullum's philosophy of linguistics, which we will explore here, is his insistence that grammar is not about abstract rules grounded in human psychology or biology but rather descriptive of sets of norms of speakers.

Before we continue, it should be noted that by normativity, Pullum is careful to distinguish his view from prescriptivism about language. In fact, Pullum has been a champion of the anti-prescriptivism movement against the grammatical scruples often associated with linguistics (see Pullum (2014)). Pullum (*forthcoming*) argues that prescriptivism is guilty of what he calls 'dialect chauvinism' or 'the touting of one dialect as clearly (almost morally) better than another' (4). At the heart of most 'grammar rules' of this variety is an injunction to accept the standards of one dialect (usually one associated with the 'upper class') over others considered to be slang or informal. If anything, linguistic theory has grown out of an opposition to unnatural rules of behaviour imposed on speakers for the sake of social identification. Generative grammar replaced this conception with one that takes grammar rules to be features of the maturation of an innate system common to all human beings. These rules characterise not the linguistic performance of individuals but their abstract competence in their languages. They are descriptive of a state of the mind/brain.

Despite the rejection of prescriptivism, Pullum insists that any grammar worth its salt must have normative force. The rules of grammar are meant to define standards of correctness for sentences, according to this view. Furthermore, he follows Kripke (1982) in questioning how such standards or norms can arise from the kinds of descriptive facts Chomskyans have attributed to brain states: 'the world, whether on a geophysical or intracranial scale, cannot be said to be right or wrong' (Pullum *forthcoming*: 4).⁹ Grammatical rules, on the other hand, are more like the rules of chess or driving, for example, 'the bishop moves diagonally', 'right lane must turn right' etc. You can choose to follow them and thus conform to the conventions of that practice or withdraw from the practice altogether. There is a Wittgensteinian, almost inferentialist, flavour to Pullum's views on grammar here. He has even suggested a mechanism, borrowed from moral theory, to include corpus linguistics and usage data into the theoretical fold. The proposed epistemology of syntax involves a reflective equilibrium between theory, judgements and corpora:

The goal is an optimal fit between a general linguistic theory (which is never complete), the proposed rules or constraints (which are not quite as conformant with

⁹ Millikan has, over the years, offered a theory of biological proper function which does embrace normativity and deals with linguistic conventions among other things. See Millikan (1984, 2005).

the general theory as we would like), the best grammaticality judgments obtainable (which are not guaranteed to be veridical), and facts from corpora (which may always contain errors) (Pullum 2007: 37).

In order to proffer this alternative picture, he dismantles the concept of ‘rule’ in generative grammar. A generative grammar, for him, is neither a function nor a set enumerator, as some in the literature have suggested. It is a holistic, non-deterministic, random, construction system (Pullum (forthcoming)). In other places, he traces the concept to Emil Post’s work on production systems (Pullum (2011, 2018)). Whatever the interpretation of generative grammar, transformational, rewriting systems and so on, the nature of a rule is so far divorced from the ordinary notion that it is not fit for purpose.

The ordinary intuitive understanding of a rule is something that we can follow (that is, behave in a way that complies with it) or break (that is, violate or disobey it). It defines a regular pattern or practice, a way of “going on in the same way”. But nothing of the content of a generative grammar has anything like this character (Pullum 2018: 202).

It is not hard to imagine what Pullum thinks does have such a character, namely, model-theoretic syntax. We will return to that point in the next section. For now, it is clear that he does not prescribe to any concept of an ‘internalised, intensional, individualistic’ grammar or I-language. For Pullum, grammar is external, social and normative. But despite antagonism from both sides of the I-language-E-language divide, there is no reason to dismiss the possibility of a union of the two (Nefdt (2023)). There is some neurolinguistic evidence for modularity and domain-specific linguistic mechanisms in the brain. In a recent review, Tuckute et al. (2024: 281) claim, among other things, that ‘the brain’s language system constitutes a distinct component of the mind and brain that is specific for language processing, separable from other cognitive systems, and relatively functionally homogeneous across its regions’. Others, such as Friederici et al. (2017), identify the Broca’s area (BA 44) among others in syntactic processing (as well as the dorsal pathway). Describing the features of this brainstate at a more abstract level in terms of hierarchical trees and rules for deriving structures might be less about conforming to the ordinary usage of the term ‘rule’ and closer to what generative theory has deemed parameters or settings. When this modular system’s output is externalised to the extracranial linguistic environment, then it enters the world of norms and rules in the traditional sense. This option would, of course, be theory-dependent and synchronic. One might argue, with inferentialism, that language starts with external socio-linguistic rules undergoing internalisation over time, that is, from E-language to I-language.¹⁰ The options abound. In fact, as Seuren (2013) notes (channelling De Saussure), Chomsky’s I-language conception can only be a partial part of the story of the social complexity of language. ‘If it were the whole picture...it would be impossible for an individual speaker to discover, say, a new word in L, since the speaker, being competent in L, should already know the word’ (Seuren 2013: 13). Of course, Chomsky’s point is more structural than lexical, but the same reasoning could apply that in order to identify all the structures of a given language, we need to collate evidence from multiple I-languages. This introduces sociality, and with it comes normativity.

¹⁰ See Peregrin (2014) for this specific argument.

In sum, Pullum might be correct that grammar has a normative force. However, this possibility does not rule out that something like an I-language could be characterised by a generative system of rules conceived of in a different way. Naturalism and normativity need not be in conflict.

5. A Unified Outlook on Language

In this last section, we will attempt a unification of the different strands of Pullum's philosophy of linguistics so far discussed. Given space limitations, I will only provide a sketch here. What makes Pullum such an inimitable scholar is his considered adherence to positions not commonly held together in the literature. Pullum is a formalist about linguistics. His model-theoretic alternative for syntax has been developed by him and others to a high level of mathematical precision. Despite this, he rejects formal properties such as discrete infinity as appropriate properties of natural languages. He believes in formally characterisable linguistic rules but attributes normative force to them and jettisons a psychological or biological grounding of such rules for language.

The formalism and the model theory are certainly not incompatible. As a guide to his interpretation of the term 'formalism', in Scholz & Pullum (2007), they distinguish between 'formal' as denoting systems which abstract over meaning and 'formalization' as a term indicating the conversion of statements of theory into precise mathematical representation. Pullum's model-theoretic syntax is formal in both senses. The framework translates statements of grammatical constraints of a language into axioms or models of the grammar (via feature structures or attribute matrices, depending on the particular framework). In so far as these statements concerning syntactic norms or rules, such as 'adverb modifiers in adjective phrases precede the head adjective in English', abstract over the meaning, and focus on the distributional patterns. They are formal in the sense of an uninterpreted calculus.

Pullum (forthcoming) further links model-theoretic syntax to the normativity of language. The best way to capture the rules of language in an analogous manner to other types of common rules is via constraints on linguistic structures. 'What I mean by a constraint is simply a statement that may be true or not true when evaluated within a certain kind of object' (Pullum forthcoming: 12). These statements can be stated as constraints on formal tree representations, such as 'if any node is a parent of both an NP node and a VP node, the NP node precedes'. What this kind of rule captures about Standard English is a norm of an expression type. It is an abstraction from social practices of rightness and wrongness of usage, not an abstraction of a brain state of the cogniser. Here, there is genuine disagreement between Pullum and the generativists. Trees are not in the head, but constraints on trees do characterise the linguistic norms of speakers (and signers) in the world. It is ultimately a model of behaviour.

The emerging picture is related to a view in the philosophy of logic and mathematics, which, in my view, connects a number of streams in Pullum's philosophy of linguistics, including his formalism, his penchant for Wittgensteinian rule-following arguments, normativity and the opposition to infinity. Constructivism encompasses combinations of all of the above to varying degrees. Pinning Pullum down is tricky here. Let's start with the basics. According to Dummett (1975: 403), constructivism in mathematics is a theoretical framework with the following properties.

[T]he meaning of all terms, including logical constants, appearing in mathematical statements must be given in relation to constructions which we are capable of

effecting, and of our capacity to recognize such constructions as providing proofs of those statements.

Constructivism takes on many forms. The most extreme version, strict finitism, interprets ‘capable of’ in terms of practice and not principle. For a proof to be legitimate, it needs to be more than just finite but also actually surveyable and intelligible in its details. Naturally, the infinite is on the chopping block. The idea that infinitary proofs in mathematics correspond to actual (complete) totalities has been something that scholars have wrestled with since Aristotle.¹¹ Finitism has been advocated by many prominent mathematicians and philosophers, including Hilbert and Wittgenstein. Hilbert, for example, aimed to justify Cantorian set theory by means of a finitary arithmetic in the metatheory, whereas strict finitism rejects the transfinite conception of the natural numbers itself. Not all versions of constructivism block countable infinities, though. Intuitionistic set theory merely reinterprets its quantifiers in its expression of the axiom of infinity from ZFC. Larger cardinalities are, however, banned, which causes ripple effects for the axioms of power sets and replacement used to ensure their existence. What does link constructivist approaches to Pullum’s claims about linguistic infinity is that they question the need for the axiom of infinity. ‘That we need an Axiom of Infinity for a theory of natural numbers and standard arithmetic turns out, however, to be an artificiality of (systems like) ZFC’ (Bremer 2007: 134). Similarly, as has been shown above, Pullum considers the linguistic infinitude claim to be an artefact of generative grammars. MTS can thus be seen as a constructive manoeuvre directed at its avoidance. Even the resistance to a mathematical induction for the set of natural languages discussed in Pullum & Scholz (2010) finds a voice in constructivism, wherein mathematical induction itself is challenged as a methodological tool.¹²

What about normativity? Here, Wittgenstein’s (positive) thoughts on mathematics are relevant. For Wittgenstein, mathematics concerns a rich ‘network of norms’ (Wittgenstein 1975: VII Section 67). Contra Platonism, mathematics is not the study of objective facts or facts of the matter, but rather the study of norms such as those that constitute the rules of language. ‘Let us remember that in mathematics we are convinced of *grammatical* propositions; so the expression, the result, of our being convinced is that we *accept a rule*’ (Wittgenstein 1975: III Section 27). Thus, mathematical statements or sentences are perceived as primarily normative in nature.

To see how this connects to Pullum’s view of MTS axioms, consider the possible link between Wittgenstein’s metamathematics and Hilbert’s axiomatic method. According to Friederich (2010), Wittgenstein’s notion of mathematical sentences as norms can be specifically defended in terms of Hilbert’s notion of axioms as implicit definitions. The idea is that axioms defined implicitly do not report objective mathematical facts, but rather constitute normative conventions for the use of mathematical concepts. This presents grammatical axioms such as ‘if any node is parent of both a V node and an NP node, the V node precedes’ as similar to arithmetic axioms such as ‘every natural number has a unique successor’. Instead of stating a fact or a primitive of the system, this definition partly defines both ‘natural number’ and ‘successor’, that is, it establishes a connection (or dependence)

¹¹ See Moore (1990) the actual versus potential infinity distinction.

¹² There is a competence-performance distinction in mathematics here as well. Constructivists seem to range from intuitionists (with a mentalist perspective) to strict finitists who ground mathematics in actual practice as opposed to ideal competence.

between objects. Similarly, model-theoretic axioms in syntax under this guise connect VPs to NPs to PPs in terms of dependence relations.

Again, Friederich (2010) argues that the normative element of the axioms *qua* implicit definitions is brought out by noticing the equivalent ‘grammatical form’ of the axioms in terms of a covert *Let* in front of them. For example, the sentence/axiom ‘every natural number has a unique successor’ should be understood as ‘*Let* every natural number have a unique successor’. Thus, ‘the grammatical form of this sentence makes it clear that its role is that of stating a norm for the usage of the concepts ‘natural number’ and ‘successor’ and not that of describing anything’ (Friederich 2010: 10). Pullum’s claims about grammaticality, model theory and normativity dovetail with this general Wittgensteinian picture combined with a specific form of constructivism about formal methods.¹³

Of course, constructivism in mathematics, especially strict finitism, is a contested theory.¹⁴ The same applies to Wittgenstein’s views on the normativity of mathematical statements. Nothing I have stated so far hinges on their truth or acceptance. Nevertheless, the parallel between these accounts in the philosophy of mathematics and Pullum’s views in the philosophy of linguistics remain present and compelling, in my view.¹⁵

6. Conclusion

In this article, I have described and interrogated a number of themes in Pullum’s philosophy of linguistics. I questioned whether model-theoretic syntax is really in competition with generative approaches. I discussed Pullum & Scholz (2010) on linguistic infinity and, in spite of general agreement with their challenges, showed how the infinitude claim might still survive as a theoretical posit. Lastly, I outlined his views on the normativity of grammar and contrasted them with the claims of biolinguistics. The hard-fought conclusion is a surprising union of these seemingly disparate elements in the philosophy of linguistics in terms of constructivism in the philosophy of mathematics, coupled with a Wittgensteinian approach to rule-following.

Acknowledgements. My appreciation for Geoff Pullum cannot be understated. Not only did he read (and correct) an earlier draft of this article, but he has been a constant source of inspiration, critique and support in my career. I would also like to thank Agustin Rayo and fellow speakers and audience members at the Pullum tribute conference in Edinburgh in 2023 for useful feedback and engaging commentary on some of the ideas that eventually made their way into the present work.

References

- Berwick, Robert & Noam Chomsky. 2016. *Why only us: Language and evolution*. Cambridge, Massachusetts: MIT Press.
- Borsley, Robert D. & Stefan Müller. 2021. HPSG and Minimalism. In *Head-Driven Phrase Structure Grammar: The handbook*, 977–1030. Germany: Language Science Press.

¹³ Pullum’s suggestive comments on the benefits of MTS for language acquisition also seem to resemble the incremental, active emphasis of constructivist learning theories to a non-negligible extent.

¹⁴ See Dummett (1975) or Wang’s paradox for a specific argument against strict finitism (and Magidor (2007) for a challenge to this argument). The larger problem for finitary or formalist approaches to proof is due to Gödel’s incompleteness theorem which shows among other things that a full axiomatisation of a formal system such as arithmetic is not possible since all the truths of arithmetic can never be captured by the axioms (even if we keep adding to them).

¹⁵ It should be noted that I am not claiming Pullum is a constructivist about mathematics or needs to be one.

- Bremer, Manuel. 2007. Varieties of finitism. *Metaphysica*. 8(2). 131–148.
- Büchi, J. Richard. 1960. Weak second-order arithmetic and finite automata. *Mathematical Logic Quarterly*. 6(1–6). 66–92.
- Bueno, Otávio & Mark Colyvan. 2011. An inferential conception of the application of mathematics. *Noûs*. 45(2). 345–374.
- Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory*. 2. 113–124.
- Chomsky, Noam. 1957. *Syntactic structures*. Berlin: Mouton.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA, USA: MIT Press.
- Chomsky, Noam. 1980. *Rules and representations*. New York: Columbia University Press.
- Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000a. Minimalist inquiries. In R. Martin, D. Michaels & J. Uriagereka (eds.), *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000b. *New horizons in the study of language and mind*. New York: Cambridge University Press.
- Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*. Cambridge, MA: The MIT Press.
- Dummett, Michael. 1975. Wang's paradox. *Synthese*. 30(3–4). 201–232.
- Everaert, Martin B. H., Marinus A. C. Huybregts, Noam Chomsky, Robert C. Berwick & Johan J. Bolhuis. 2015. Structures, not strings: Linguistics as part of the cognitive sciences. *Trends in Cognitive Sciences*. 19(12). 729–743.
- Friederich, Simon. 2010. Motivating Wittgenstein's perspective on mathematical sentences as norms? *Philosophia Mathematica*. 18(3). 1–19. doi:10.1093/philmat/nkq024.
- Friederici, Angela, Noam Chomsky, Robert C. Berwick, Andrea Moro & Johan J. Bolhuis. 2017. Language, mind and brain. *Nature Human Behaviour*. 1. 713–722.
- Hauser, Marc, Noam Chomsky & Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science*. 298(22). 1569–1579.
- Hunter, Tim. 2021. The Chomsky hierarchy. In Nicholas Allott, Terje Lohndal and Georges Rey (eds.), *A companion to Chomsky*, 74–95. United States: Wiley.
- Huybregts, M. A. C. 2019. Infinite generation of language unreachable from a stepwise approach. *Frontiers in Psychology*. 10. doi:10.3389/fpsyg.2019.00425
- Jäger, Gerhard & James Rogers. 2012. Formal language theory: Refining the Chomsky hierarchy. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 367(1598). 1956–1970.
- Johnson, Kent. 2015. Notational variants and invariance in linguistics. *Mind and Language*. 30(2). 162–186.
- Katz, Jerrold J. 1996. The unfinished Chomskyan revolution. *Mind and Language*. 11(3). 270–294.
- Kripke, Saul A. 1982. *Wittgenstein on rules and private language: An elementary exposition*. Cambridge: Harvard University Press.
- Langendoen, D. Terence. 2010. *Just how big are natural languages?* 139–146. Berlin, New York: De Gruyter Mouton.
- Langendoen, D. Terence & Paul M. Postal. 1984. *The vastness of natural languages*. Oxford: Basil Blackwell.
- Levitt, Willem. 1974. *An introduction to the theory of formal languages and automata*. Berlin, Boston: De Gruyter Mouton.
- Lewis, David K. 1995[2010]. Languages and language. In Darragh Byrne & Max Kölbel (eds.), *Arguing about language*, 3–35. United Kingdom: Routledge.
- Lobina, David J. 2017. *Recursion: A computational investigation into the representation and processing of language*. Oxford: Oxford University Press.
- Magidor, Ofra. 2007. Strict finitism refuted? *Proceedings of the Aristotelian Society*. 107(1pt3). 403–411. doi:10.1111/j.1467-9264.2007.00230.x.
- Michaelis, Laura A. 2009. Sign-based construction grammar. In *The Oxford handbook of linguistic analysis*, 139–158. Oxford: Oxford University Press.
- Millikan, Ruth Garrett. 1984. *Language, thought, and other biological categories: New foundations for realism*. Cambridge, MA: MIT Press.
- Millikan, Ruth Garrett. 2005. *Language: A biological model*. Oxford, GB: Clarendon Press.
- Moore, Adrian W. 1990. *The infinite*. New York: Routledge.
- Mukerji, Nirmalangshu. 2022. *The human mind through the lens of language*. New York: Bloombury Press.
- Müller, Stefan, Anne Abeillé, Robert Borsley & Jean-Pierre Koenig (eds.). 2021. *Head-Driven Phrase Structure Grammar (Empirically oriented theoretical morphology and syntax 9)*. Berlin: Language Science Press.

- Nefdt, Ryan M. 2019. Infinity and the foundations of linguistics. *Synthese*. 196(5). 1671–1711.
- Nefdt, Ryan M. 2023. *Language, science, and structure: A journey into the philosophy of linguistics*. Oxford: Oxford University Press.
- Peregrin, Jaroslav. 2014. *Inferentialism: Why rules matter*. London and New York: Palgrave-Macmillan.
- Postal, Paul. 2009. The incoherence of Chomsky's 'biolinguistic' ontology. *Biolinguistics*. 3. 104–123.
- Pullum, Geoffrey K. 2007. The evolution of model-theoretic frameworks in linguistics. In James Rogers & Stephan Kepser (eds.), *Model-theoretic syntax at 10 (Proceedings of the MTS@10 workshop)*, 1–10. Dublin, Ireland: Trinity College Dublin.
- Pullum, Geoffrey K. 2011. On the mathematical foundations of syntactic structures. *Journal of Logic, Language and Information*. 20(3). 277–296.
- Pullum, Geoffrey K. 2013. The central question in comparative syntactic metatheory. *Mind and Language*. 28(4). 492–521.
- Pullum, Geoffrey K. 2014. Fear and loathing of the English passive. *Language & Communication*. 37. 60–74.
- Pullum, Geoffrey K. 2018. Formalism, grammatical rules, and normativity. In James McElvenny (ed.), *Form and formalism in linguistics*, 197–223. Berlin: Language Sciences Press.
- Pullum, Geoffrey K. forthcoming. Rules of grammar. In Ryan M. Nefdt, Gabe Dupre & Kate Stanton (eds.), *The Oxford handbook of the philosophy of linguistics*, Oxford: Oxford University Press.
- Pullum, Geoffrey K. & Barbara C. Scholz. 2001. On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In Philippe de Groot, Glyn Morrill & Christian Retoré (eds.), *Logical aspects of computational linguistics*, 17–43. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Pullum, Geoffrey K. & Barbara C. Scholz. 2010. Recursion and the infinitude claim. In Harry van der Hulst (ed.), *Recursion and human language*, 111–138. Germany: De Gruyter Mouton.
- Rey, Georges. 2020. *Representation of language: Philosophical issues in a Chomskyan linguistics*. Oxford: Oxford University Press.
- Rogers, James. 1998. A descriptive characterization of tree-adjointing languages: (project note). In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics - volume 2 ACL '98/COLING '98*, 1117–1121. USA: Association for Computational Linguistics.
- Savitch, Walter J. 1993. Why it might pay to assume that languages are infinite. *Annals of Mathematics and Artificial Intelligence*. 8(1–2). 17–25.
- Scholz, Barbara & Geoffrey Pullum. 2007. Tracking the origins of transformational generative grammar. *Journal of Linguistics*. 43. 701–723.
- Scholz, Barbara C., Francis Jeffrey Pelletier, Geoffrey K. Pullum & Ryan Nefdt. 2022. Philosophy of linguistics. In Edward N. Zalta & Uri Nodelman (eds.), *The Stanford encyclopedia of philosophy*, (Spring 2024 edn). Stanford: Metaphysics Research Lab, Stanford University.
- Seuren, Pieter. 2013. *From Whorf to Montague*. Oxford: Oxford University Press.
- Thatcher, James W. & Jesse B. Wright. 1968. Generalized finite automata theory with an application to a decision problem of second-order logic. *Mathematical Systems Theory*. 2. 57–81.
- Tiede, Hans-Jörg & Lawrence Neff Stout. 2010. Recursion, infinity, and modeling. In Harry van der Hulst (ed.), *Recursion and human language*, 147–158. Berlin, New York: De Gruyter Mouton.
- Tomalin, Marcus. 2006. *Linguistics and the formal sciences: The origins of generative grammar*. Cambridge: Cambridge University Press.
- Tuckute, Greta, Nancy Kanwisher & Evelina Fedorenko. 2024. Language in brains, minds, and machines. *Annual Review of Neuroscience*. 47. 277–301.
- Vijay-Shanker, K. & David Weir. 1994. The equivalence of four extensions of context-free grammars. *Mathematical Systems Theory*. 27(6). 511–546.
- Wittgenstein, Ludwig. 1975. *Wittgenstein's lectures on the foundations of mathematics, Cambridge, 1939*. Chicago: University of Chicago Press.