

# Infinity and the Foundations of Linguistics

Ryan M. Nefdt\*

Received: date / Accepted: date

**Abstract** The concept of linguistic infinity has had a central role to play in foundational debates within theoretical linguistics since its more formal inception in the mid-twentieth century. The Conceptualist tradition, marshalled in by Chomsky and others, holds that infinity is a core explanandum and a link to the formal sciences. Realism/Platonism takes this further to argue that linguistics is in fact a formal science with an abstract ontology. In this paper, I argue that a central misconstrual of formal apparatus of recursive operations such as the set-theoretic operation *merge* has led to a mathematisation of the object of inquiry, producing a strong analogy with discrete mathematics and especially arithmetic. The main product of this error has been the assumption that natural, like some formal, languages are discretely infinite. I will offer an alternative means of capturing the insights and observations related to this posit in terms of scientific modelling. My chief aim will be to draw from the larger philosophy of science literature in order to offer a position of grammars as models compatible with various foundational interpretations of linguistics while being informed by contemporary ideas on scientific modelling for the natural and social sciences.

**Keywords:** *infinity; linguistics; methodology; scientific modelling; mathematics; philosophy of science*

## 1 Introduction

The ontological basis of the linguistic enterprise has been contested since something resembling an official stance was adopted through the generative or biolinguistic tradition in the late 1950's. On this view, natural languages are states of the mind/brain and thus part of cognitive-psychological reality. An alternative approach (proffered by Katz 1981, Langendoen and Postal 1984, Soames 1984, Katz and Postal 1991, and Postal 2003) considers natural languages (and the sentences of which they are comprised) to be abstract objects, in the sense of being mind-independent and non-spatio-temporally extended. Linguistics on this view is a formal science on the level of logic or mathematics. Arguments have gone back and forth in favour of specific views without much common ground, to the effect that these linguists and philosophers often seem to talk past each other. In this paper, I will argue for an alternative relationship between methodology and ontology which redraws the lines between foundational views and offers a different interpretation of the role of mathematics in linguistics. Specifically, I will target the postulate of linguistic infinity so prevalent in the theoretical linguistics literature and pedagogy as a conduit to a larger discussion of the foundations of linguistics. My chief aim will be to offer a view with minimal ontological consequences which draws on insights from computational linguistics and the philosophy of science.

---

\* Forthcoming in *Synthese*, special issue on infinite idealizations in the sciences.

I shall first identify the problem of ‘mathematisation’ as I see it in section 2. In section 6, I introduce a framework in terms of three grades of mathematical involvement for the grammars of linguistic theory. These grades involve the methodological attitudes linguists take towards their grammars. More specifically, the grades correspond to possible positions on the nature of the mathematical apparatus used in the grammar and its relation to the nature of natural language itself. I will show that infinity claims only cause problems for the last two grades. Nevertheless, although each grade has its own set of difficulties, I suggest that the first grade of involvement offers us the most neutral and reasonable approach to linguistic methodology (while avoiding certain issues within its ontology).

## 2 The *Mathematisation* of Natural Language

### 2.1 Formalisation, Mathematisation and Conservativeness

I begin this section with a distinction. The distinction is between the concepts of *formalisation* and *mathematisation* respectively. Formalisation is the familiar tool of simplifying natural structures or phenomena for the purpose of making them more amenable to precise characterisation, often in terms of the language of first-order logic. This tool is certainly wide-spread in linguistics and philosophy. Mathematisation, on the other hand, can be seen as the process of rendering natural phenomena into mathematical structures or entities by either formal modelling, analogy or alleged proof. It marks an ontological shift in the target system (or sometimes omission of the target altogether). In other words, it approaches the subject matter of linguistics as a mathematical puzzle capable of precise mathematical characterisation and resolution without an attempt to interpret any features back into the target system. In what follows I shall argue, for instance, that an example of the formalisation is the use of lambda abstraction as a means of variable binding and functional application in semantics. While an example of mathematisation is present in the assumption of discrete infinity of natural language.

In this section, I want to provide some details concerning the concept of mathematisation, specifically within the context of the linguistic project. The concept will remain in the background of most of the discussion to follow. Unfortunately, it has not received any proper definition, despite some scattered usage across reflective theoretical work in the philosophy of science. I will provide a working definition for present purposes. This characterisation should not be considered definitive but rather illustrative. For the sake of contrast, I will begin with the more familiar concept of formalisation.

The concept and technique of formalisation, its features and fecundity have been addressed in philosophy, logic and other fields. In general, despite cautionary tales, formalisation is considered benign and mostly useful. As Pullum (2013) states “formalization is the use of appropriate tools from mathematics and logic to enhance explicitness of theories [...] Any theoretical framework stands to benefit from having its content formalized” (493). One important feature of formalisation is conservativeness, a property which received a more controversial treatment in Field (1980) and the mathematical nominalism with which it came. Contrary to popular indispensability arguments as to the essential place of mathematics within the natural sciences, Field proposed a thoroughgoing fictionalism about mathematical entities related to the concept of conservativeness given below.

A mathematical theory  $S$  is *conservative* if, for any nominalistic assertion  $A$  and any body of such assertions  $N$ ,  $A$  is not a consequence of  $N+S$  unless  $A$  is a consequence of  $N$  alone (Field, 1985: 240).

The basic idea is that nothing that can be proven with mathematics cannot be proven without it within scientific investigation. Whether or not this principle holds or can be shown to hold for the relationship between mathematics and the sciences remains to be seen (see Shapiro 1983, 1997). However, formalisation is certainly conservative in the way discussed above. Formal languages are devices used for the representation and abbreviation of an intended target domain. As the term suggests, the technique is meant to home in on the “form” of a problem and highlight the relations in a non-obfuscating manner. In fact, formalisation is related to an emphasis on a syntactic analysis of mathematical and scientific discourse, initially proposed in the philosophical project that came along with Hilbert’s programme in the first half of the 20th century.

## 2.2 Hilbert, Bar-Hillel and Chomsky

The beginnings of linguistics in the 20th century owe much to Hilbert’s programme and Formalism in the philosophy of mathematics. The issue of mathematical infinity preoccupied these philosophers and mathematicians and I will argue left this preoccupation with the linguists who inherited its intellectual history.

For Formalists, like Hilbert, mathematics was not a pre-interpreted theory of some extra-physical or mental reality but rather an uninterpreted calculus of symbols, the manipulation of which yields structures capable of later interpretation.

Every science takes its starting point from a sufficiently coherent body of facts given. It takes the form, however, only by *organizing* this body of facts. This organization takes place through the *axiomatic method*, i.e. one constructs a *logical structure of concepts* so that the relationships between the concepts correspond to relationships between the facts to be organized (Hilbert, 1899 [2004]: 540).

Hilbert’s axiomatic method was based on a concept of implicit definition. For example, unlike the axiomatics of Euclid which involved explicit definition of geometric terms such as a point being defined as “extensionless” or the like, Hilbert’s axioms introduced implicit definition directed toward the goal of divorcing theory from intuition (although intuitions might still play a motivating role for the axioms). As Shapiro puts it “geometry was becoming less the science of space or space-time, and more the formal study of certain structures” (2005: 63). The idea is that lines and points are to be defined purely in terms of the axioms of geometry and furthermore anything that fulfills the conditions set by the axioms will do equally well (what Shapiro calls “free-standing”). Nothing logico-conceptual is supposed to be given in advance of theory.

What separated Hilbert’s programme from other versions of Formalism was a focus on finitary methods, an aspect which took on special significance for linguists in the early 20th century. A discussion of finitism or strict finitism would take us too far afield. Suffice to say, the beginnings of proof theory in mathematics incorporated a notion of consistency paired with existence and a distinct axiomatic approach coupled with the implicit definition of core concepts. The full fruition of this project could not be achieved since the consistency of arithmetic could not be proven through finitary means (thanks to Gödel’s incompleteness result).<sup>1</sup>

Nevertheless, Hilbert’s axiomatic method had a profound effect on the scientific community. An effect which did not go unfelt within the linguistic community at the time. Early 20th century linguistics saw the anthropological goals of figures such as Sapir and Whorf take on a secondary role to the rigour of mathematical methodology.

[D]uring the 1930s and 1940s other developments in the theory of logical syntax occurred, which were ultimately to have profound implications for linguistic research, and the starting point was usually Hilbert’s proof theory, which seemed to imply the meaning-less syntactic manipulations could suffice to resolve a whole range of epistemological problems (Tomalin, 2006: 89).

Bloomfield was among the first to embrace this approach to linguistics. In the spirit of the clarification of confusion, the identification of errors and the general precisification of the field, Bloomfield proposed the axiomatic method as its chief tool of investigation. Unfortunately, logical positivism and its circumscribed philosophical agenda also crept into Bloomfieldian linguistics. Meanings and any mental characterisation were anathema to early linguists. Formal syntax, just as the propositional or predicate calculus, was a more secure footing upon which to base the scientific study of language. Thus, the mathematical foundations became firmly entrenched.

One of the first sights of the application of recursive techniques or proof-theory in modern linguistics was in a paper by Bar-Hillel in 1953. Bar-Hillel attempted to extend the use of formal recursive techniques beyond the purely mathematical. By taking English as the metalanguage and French as the object language, he recursively redefined the basic parts of speech (noun, verb, etc) in order to establish a mathematical two-part recursive definition of a “proper” or grammatical French sentence.

Bar-Hillel’s use of recursive definitions to analyse the structure of sentences in natural language can be viewed as one manifestation of this pervasive desire for the mathematisation of syntactic analysis, which became such a characteristic feature of certain kinds of linguistic research in the mid-twentieth century (Tomalin, 2006: 67).

<sup>1</sup> However, the incompleteness proof did not destroy proof theory itself. Gentzen developed a proof-theoretic approach without the limitations of the Hilbert programme and finitism in full view of Gödel’s incompleteness. His more specific aim was to prove the consistency of logical deduction within arithmetic.

This result did not go unnoticed by Chomsky. With this methodology came a movement away from empirical discovery procedures (the likes of which his mentor Harris had been after) toward mathematical precision and specification of evaluation procedures. This move was similar to one attributed to Hilbert in his axiomatic treatment of geometry, based not on spatial considerations, but pure mathematical structure arrived at through implicit definition. In the same way, geometry and linguistics can be thought to be motivated by intuition, but their study is ultimately tied up with mathematical investigation.<sup>2</sup> Continuing with the theme of mathematisation, Peregrin (1995) describes Chomsky’s contribution in the following way.

Chomsky’s *novum* was that he proposed organizing the rules into a hierarchical system allowing the systematical generation, and basing all this upon setting up of the grammar as a real mathematical structure. Such a mathematization entailed an exceptional increase of rigour and perspicuity and, moreover, it led to the development of a metatheory, investigating into the formal properties of grammars (e.g. their relative strengths) (88).

Although, Hilbert, Bloomfield and Bar-Hillel all had an influence on the mathematical trajectory of the field. Chomsky’s “novum” was more directly inspired by the work of Emil Post and the mathematisation of syntactic structure. I will briefly touch on this aspect of early generative grammar in the next section. The metatheory alluded to by Peregrin is captured in the field of formal language theory (FLT) which has since Chomsky’s early work been dominated by generative (proof-theoretic) mathematics.

We do not have the space to enter into a protracted discussion of FLT here. A few details should suffice. This issue is relevant for why cardinality concerns have shaped the foundations of linguistics as they have. FLT involves the mathematical characterisation of classes of formal languages. A formal language, in this sense, is a set of sequences of strings over a finite vocabulary. The members of this set vary according the field to which we apply formal language theory, i.e. words if we are talking about natural languages or states if we are talking about programming languages etc. Furthermore, in formal language theory we are concerned with the finite ways in which these languages can be described, “FLT deals with formal languages (= sets of strings) that can be defined by finite means, even if the language itself is infinite” (Jäger and Rogers, 2012: 1957). This is usually done by means of formal grammars (i.e. sets of rules by which we construct well-formed sentences or answer the membership problem for a class of structures). The linguistic side of FLT is directed towards the goal of describing the various constructions of natural language syntax in terms of the class of formal languages in the hierarchy of such languages that best captures it.

### 2.3 Mathematisation in the other sciences

The practice of mathematisation is by no means idiosyncratic to the development of linguistic theory. In the natural sciences too, problems are often approached with a certain sort of mathematical transmogrification. As early as Galileo, the natural world has been considered amenable to precise mathematical characterisation. Considering the observable phenomenon of free falling objects, a pattern emerges between distance and time (i.e. distance is proportional to time squared). It does not, however, suffice to merely describe this pattern in plain language.

The mathematisation of the problem consists in our being able to specify the relation between distance and time in a precise way, a specification that is not possible using qualitative language. But note here that the relation between the *qualitative* concepts of distance and time plays an important role in what we call the ‘mathematisation’ of the problem [...] What is interesting, however, is that from the uses of mathematics as a type of tool for reconstruction emerges a representational framework with a life of its own (Morrison, 2015: 2-3).

Other examples include the technique of the thermodynamic limit in particle physics. In order to explain the breakdown of electromagnetic gauge invariance, physicists help themselves to the

<sup>2</sup> I will return to this analogy in section 8. On the philosophical side, I think that mathematisation is related to the treatment of the rules posited initially for the sake of modelling a phenomenon and then eventually for their own sake. A similar Wittgensteinian diagnosis of the issue within the context of the foundations of linguistics can be found in Wright (1989). Simply put, it is a problem of rule-following. The problem of mathematisation can essentially be viewed as the progression and instantiation of the myth of the autonomy of rules or “the image of a rule as a rail laid to infinity” (Wright, 1989: 238).

notion of a phase transition. Phase transitions involve a thermodynamic limit or “in other words, we need to assume that a system contains infinite particles in order to explain, understand, and make predictions about the behaviour of a real, finite system” (Morrison, 2015: 27). Similar techniques are used in population genetics in which the mathematisation of finite real populations results in models of infinite populations and their properties. In some of these cases, such as the field of mathematical physics, the line between mathematics and the natural world is irrevocably blurred.

One key difference, however, between mathematisation in the natural sciences and linguistics is that the mathematics is treated as a modelling tool in the latter and an explanandum in the former, especially in the case of infinity. Let this serve as an introduction to the concept which will be developed through the grades of involvement in section 6. In the next section, we will move on to the mathematical approach which underlies Chomsky’s seminal *Syntactic Structures* and the methodological traces it left behind in the ensuing field of linguistics.

### 3 Infinity and The Legacy of *Syntactic Structures*

Though often unappreciated, the work of Emil Post has had a profound effect on the field of linguistics (see Pullum 2011 for further discussion). Chomsky’s *Syntactic Structures* (1957) (and other papers at the time such as his 1956b) offers a distinctly proof theoretic approach to the idea of a grammar or rule-based production system. A Post canonical production system is a just a tuple  $\langle A, I, R \rangle$  with a finite vocabulary  $A$  or “axioms”, a set of initial words  $I$  (disjoint from  $A$ ) and a finite set of transformations or production rules  $R$  (these are binary relations) (each of which has an antecedent  $x$  and consequent  $y$  such that  $(x, y) \in R$  which ensures that there are no free variables in the consequent that are not in the antecedent). This system resembles familiar natural deduction systems in propositional and predicate logic. If you want to prove a specific conjecture, you start with the members of  $A$  and derive the conjecture *via* repeated application of  $R$ . “In particular, [Post] developed a generative characterization of the recursively enumerable (r.e.) sets, and later laid the foundations of recursive function theory” (Pullum, 2011: 280). In so doing, he provided a formalism for modelling the concept of a logical proof. In a sense, this provides a proof procedure for discovering the strings or formal language “generated” by a given system of rules (on a finite alphabet or vocabulary). In addition, Post canonical systems are Turing complete i.e. belong to the same class as Turing machines.

It is not surprising that the mechanism of a Post production system became central to the concept of a generative grammar used in linguistics. Chomsky (1957: 22) defines  $[\Sigma, F]$  grammars in the following way:

$[\Sigma, F]$ Grammar : Each such grammar is defined by a finite set  $\Sigma$  of initial strings and a finite set  $F$  of “instruction formulas” of the form  $X \rightarrow Y$  interpreted: “rewrite  $X$  as  $Y$ ”.

From these (and earlier) insights, formal language theory was born. As we saw above, FLT is the abstract theory of syntax in which languages are viewed as sets of strings without semantic content. For Chomsky, a  $[\Sigma, F]$  grammar (or generative grammar) is a system of rewrite rules on sets of terminal (or words) and non-terminal strings (or phrasal categories). *Syntactic Structures* suggests a proof that the syntax of natural language cannot be captured by a specific kind of formal language (a regular or finite-state language), although this result is technically never proven (and later developments such as Shieber (1985) prove that the more complex phrase-structure grammars are also inadequate for some languages).

A central insight, namely that the linguistic capacity of language users is unbounded, is what led Chomsky to develop the mathematical analogue of a computational system in order to represent this phenomenon. Lobina (2014) claims that Chomsky adopted this position at a time when the terms “computation” and “recursion” were used interchangeably and this might explain his insistence of the centrality of recursion within linguistics itself.<sup>3</sup> The idea that natural language syntax could be represented by recursive rule systems or as computational devices was not completely

<sup>3</sup> Although as early as *Syntactic Structures*, Chomsky distinguishes between recursion and computation in constructing a non-recursive grammar of a small fragment of English. I thank an anonymous referee for pointing this out to me.

novel as we saw with Bar-Hillel.<sup>4</sup> Chomsky’s early work emphasised the need for greater precision which ultimately lead to linguists or mathematical linguists taking more interest in formal language theory, i.e. the mathematisation of syntax. However, mathematical linguists tend to restrict themselves to linguistically interesting or motivated investigation into the infinite classes of formal languages and their respective complexities. For instance, context-sensitive languages and regular languages are generally of little interest given that it is unlikely that natural language constructions can be found within these formal parameters (oversimplifying, the latter are too restrictive and the former too complex).

Whereas in formal language theory the use of sets can be viewed as an abstraction or convenience in some sense, within recursion theory sets (in terms of functions) are not optional. If recursion is assumed to be a feature of the natural landscape as opposed to merely a feature of our models, we move toward a more pervasive mathematisation.<sup>5</sup>

It is not my purpose to detail the developments of formal language theory or formal syntax here. The claim, which I hope to impress upon the reader, is that the beginnings of the generative or biolinguistic movement have nontrivial logical and mathematical foundations. The historiography of linguistics is a much more complex matter than I have shown in the previous two sections, where my intention was illustrative rather than comprehensive. Needless to say, Carnap, Goodman, Quine, Harris all deserve mention within a more complete story. However, such a task is beyond the scope of the present work (see Newmeyer (1996) for a generative approach and Tomalin (2006) for a more objective attempt).

It is important to mention, at this juncture, that the generative tradition in linguistics offered an approach to central aspects of natural language comprehension and production related to infinity that were previously unaccounted for, namely productivity, learnability and creativity. The tools of this restricted class of Post-canonical systems, i.e. generative grammars, allegedly provided insight into something that could not be approached without this mathematical apparatus, specifically the question of how a finite system could generate an infinite output (a key aspect of productivity, learnability and creativity on this account). Chomsky (2000) claims that the explanation of natural language creativity only became available with the advent of computability theory in the 20th century (and before then seemed like a contradictory property for a physical system to possess), or “[a]dvances in the formal sciences provided that understanding, making it feasible to deal with the problems constructively” (Chomsky, 1995: 4). Again, recursion commanded a central explanatory role in linguistic theory.

Hinzen and Uriagereka (2006) draw even stronger conclusions concerning the connection between linguistics and mathematics.

[T]he human language faculty poses much the same explanatory problems for contemporary physicalism as the mathematical faculty does (72).

and,

Chomsky’s technical correlation between language and mathematics is also well-taken, given the biological isolation they both share as systems instantiating discrete infinity. However, if the latter can be abstracted from FL [faculty of language], just how much of mathematics is FL using to begin with? [...] Surely the successor function fits naturally into syntagmatics (84).

We shall return to this issue below, but for now suffice to say that the analogy between linguistics and mathematics goes deeper than just that of recursion and the successor function. The ensuing movement only added to the mathematical foundations of linguistics. The use of sets and functions have become ubiquitous in linguistic theory but the mathematical apparatus does not stop at these relatively benign tools of characterisation or formalisation. Along the way, recursion and discrete infinity somehow became in need of physical interpretation.

<sup>4</sup> The history of the term “recursion” in linguistics is extremely messy. Briefly, the idea is that recursive functions introduce a property of self-reference. This involves two steps. One which specifies the condition of termination of the recursion or the base case and the recursive step which reduces all other cases to the base. For some early attempts at formal definition, see Chomsky (1959) and Chomsky and Miller (1963) for an elaboration of the term language and recursion in linguistics. For a development of these ideas within recursive characterisations of particular linguistic constructions see Langendoen (2008).

<sup>5</sup> Notice, even Chomsky’s famous (1956b) disavowal of the relevance of stochastic grammar formalisms, in which approximation through continuous mathematics is the goal, can be seen as motivated by mathematisation. The statistical methods of continuous mathematics do not generally make a mathematical object of the target domain but rather treat it as physical process capable of “approximate” characterisation.

On the other side of the fence, semantics was also undergoing mathematisation. Following the pioneering work of Lambek (1958), Montague famously stated that he “reject(s) the contention that an important theoretical difference exists between formal and natural language” (1970a: 188). In fact, this is a point (reiterated below) upon which he believed himself to be in agreement with Chomsky as evinced in ‘Universal Grammar’.

There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians; indeed, I consider it possible to comprehend the syntax and semantics of both kinds of language within a single natural and mathematically precise theory. On this point I differ from a number of philosophers, but agree, I believe, with Chomsky and his associates. (Montague, 1970b [1976]: 222).

Indeed, according to Thomason (1974), Montague held syntax and semantics to be branches of mathematics. It is not clear that Chomsky and his associates explicitly held the view Montague attributes to them. Nevertheless, the idea which does seem attributable to generative linguists is that structures of natural language not only lend themselves to mathematical characterisation but moreover there is some “special” connection between the faculty of language and that of mathematics as they both “instantiate” a mathematical property, namely discrete infinity.

All approaches agree that a core property of FLN is recursion, attributed to narrow syntax in the conception [...] This capacity of FLN yields discrete infinity (a property that also characterizes the natural numbers) (Chomsky, Hauser and Fitch, 2002: 1571).

This aforementioned connection gets to the heart of the distinction between *formalisation* and *mathematisation* which I propose and hope to illuminate in the rest of the paper.

Whereas *Syntactic Structures* could be read as an application of discrete mathematics to language, opening up new vistas in the study of formal grammar, *Aspects* made clear that formalisation, and indeed the study of languages/grammars was not an end in itself. The goal was rather to reverse-engineer the structure of language to discover the mind that made it possible in the first place (Boeckx, 2015: 128).

Thus, if all of the grammars of natural language imply a discrete infinity of expressions, then this property is attributed to the language faculty by “reverse-engineering”. Statements like these are the hallmark of mathematisation. We will return to these issues in later sections.

Before delving into the mathematical grades of involvement, which aim to make sense of mathematisation and infinite generalisation in linguistics, some brief details of contemporary linguistics are in order.

#### 4 I-Language and Grammar

On the dominant Conceptualist (or mentalist) view, linguistics is a branch of cognitive psychology and the foundations of *linguistics proper* (a term introduced in Katz and Postal 1991 to describe the job of the working linguist) are eventually to be subsumed by biology or neuroscience. A more general understanding is that linguistics constitutes the study of a biological system which is responsible for language generation. Therefore, it incorporates a physicalist ontology.

So what kind of thing is a language on this view? For a Conceptualism or mentalist it is psychological in nature, linked to the state of the so-called “language faculty”.

We can take a language to be nothing other than a state of the language faculty [...] So let’s take a language to be (say, Hindi or English or Swahili) a particular state attained by the language faculty. And to say that somebody knows a language, or has a language, is simply to say their language faculty is in that state (Chomsky, 2000b: 8).

Thus, all there is to a natural language is a state of the mind/brain. Here, language is identified with the linguistic competence in that language.

More importantly, a grammar is supposed to be a scientific theory of the mental behaviour or of the state that the language faculty is in (so different states for English, Mandarin, isiZulu etc.).

A better usage would be to restrict the term “grammar” to the theory of language, and to understand the language as what we may call “I-language” where “I” is to suggest “intensional” and “internalized” (Chomsky, 1990: 678).

The grammar thus describes a language which is “internalised” in the sense of being located in the mind and eventually the brain of the ideal speaker-listener and “intensional” in terms of a function that determines a restricted set of expressions or a “grammar” and not the entire discretely infinite faculty of language itself. Another word for the former characteristic is “individualistic” or non-relational in the sense of possessing properties that depend on or are related to only internal mental features of the language-user (see Ludlow 2011 for further discussion).

#### 4.1 Formal Aspects of the Theory of Syntax

The question which most concerns this paper (and special issue) is that of the formal aspects of linguistics, such as discrete infinity and recursion in early generative grammar and the operation of “merge” in minimalism (to be defined below), and their relation to the mathematical realm. Whether through recursion or merge, Conceptualism seems to advocate for an extra-biological, or at least special, claim that language is somehow mathematically unique. This claim starts with the idea that the human language capacity (and cognitive capacity in general) is supposed to be understood as finite in its resources, yet one apparent aspect of natural language is its creative nature, assumed to be capable of (discretely) infinite expression.

The most striking aspect of linguistic competence is what we may call the ‘creativity of language’, that is, the speaker’s ability to produce new sentences that are immediately understood by other speakers although they bear no physical resemblance to sentences that are ‘familiar’ (Chomsky, 1966: 74).

Most linguistics textbooks start with the claim that natural language is infinite (for examples, see Lasnik 2000, Sag *et al* 2003, Yang 2006). Some linguists even go as far as to claim that infinity is the only linguistic universal (Epstein and Hornstein 2005). This aspect of the tradition has led to some criticism (see Pullum and Scholz 2010) but it has also led to the connections with the realm of mathematics as we saw in section 3 and can glean from these rather speculative comments in Chomsky (2010: 48):

The “gigantic development of the mathematical capacity is wholly unexplained by the theory of natural selection, and must be due to some altogether distinct cause,” if only because it remained unused. One possibility is that it is derivative from language. It is not hard to show that if the lexicon is reduced to a single element, then unbounded Merge will yield arithmetic.

In ‘On Phases’, Chomsky is more explicit on how this procedure is to be accomplished. Even though as Tomalin (2007: 1795) notes, if the lexicon contains a single element, then merge cannot be applied without some sort of indexation.

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 (Chomsky, 2005: 6).

Talk of recursion, arithmetic, discrete infinity and the set-theoretic operation of merge seems to suggest a deeper analogy with the formal sciences. In a lecture in 2011 at Carleton University, Chomsky claimed that “perhaps the most elementary property of human language is that it consists of a discrete infinity of interpretable expressions – so there’s five-word sentences, and six-word sentences, no five-and-a-half words sentence, so it goes on indefinitely like the integers. That’s kind of unusual, there’s nothing like that known in the biological world.” In my view, it is this alleged divergence from other aspects of the natural or biological world which leads the linguist from a purely naturalistic endeavour to a partly formal one. The problem is that discrete infinity requires the apparatus of discrete mathematics for characterisation. This is not mere formalisation but rather a mathematisation of an object (or state) claimed to be biological in nature.

Let us consider the “merge” postulate for a moment. MERGE is an arbitrary operation on sets of syntactic items, essentially it takes two objects and combines them into one (labelled) object. It is *internal merge* which performs the role which recursion performed previously. *External merge* takes two distinct objects as input and *internal merge* allows embedding and thus allows for recursion. Furthermore, *internal merge* involves duplicating items within the operation (see Chomsky 1995 and Langendoen 2003 for more details). For instance, if we merge syntactic objects (modelled as a set)  $\alpha$  and  $\beta$  to form the unordered set  $\{\alpha, \beta\}$  and there is a  $\gamma$  such that  $\gamma \in \alpha$  and we merge this object with  $\{\alpha, \beta\}$ , we would have two copies of  $\gamma$  in the resulting structure. In this way, we can account for all movement with minimal operations in the syntax (and various constraints on the operations). It is in explaining the “arbitrariness” of merge that we once again see a parallel with arithmetic.

Within the framework just outlined, there is also no meaningful question as to why one numeration is formed rather than another – or rather than none, so that we have silence. That would be like asking that a theory of some formal operation on integers – say, addition – explain why some integers are added together rather than others, or none (Chomsky, 1995: 208).



Here we are dealing with an arbitrary set-theoretic function which yields a discrete infinity of natural language expressions, i.e. biological output. In the case of arithmetic, the output of the successor function is not usually considered to be physical in any strict sense (although nominalists in the philosophy of mathematics might disagree). And yet both language and arithmetic are alleged to have sprung from the same well.<sup>6</sup> Whether it is Turing machines, discrete infinity or merge, my claim is that the Conceptualist approach involves a level of mathematisation of the “natural object” of language (what Platonists such as Postal (2003) call “incoherent”). Linguistics thus seems to view itself as a “special” science in a sense divorced from other empirical sciences, and wedded to aspects of the formal sciences. As I suggested in the first section, the move seemed to involve a departure from mere *formalisation* to a distinctive *mathematisation* of recursive elements specifically.

[I]n the earliest work, although recursive components were considered useful formal procedures that simplified the basic analytical framework, no strong claims were made concerning their biological status. Gradually, though, as the theory of GG developed [...] the role of recursion within the GG framework began to acquire cognitive connotations, with the eventual result that [...] it has been hypothesised that recursion is a genetically-embedded computational procedure (Tomalin, 2007: 1785).

Given the conceptualist and more so the biolinguistic agenda (of integrating the study of language with other biological systems), the onus is on the conceptualist to provide an evolutionary story for how such a purportedly “extra-biological” discrete infinity creating operation such as merge emerged from the physical world. In an attempt to do so, Hauser, Chomsky and Fitch (2002) go as far as to state that recursion or merge is the core property of the faculty of language (narrowly construed). They go further to state that merge is an evolutionary mutation which gave rise to human linguistic abilities (and perhaps arithmetic ones too). This story has met with large-scale criticism and an in-depth discussion of this postulate is beyond the current scope but it does serve as an example of the centrality of the recursion and discrete infinity claim within the Conceptualist paradigm. At the very least, Conceptualists (and biolinguists alike) seem to have to account for the role of such set-theoretic operations in their science which is alleged to be an empirical one. Is talk of recursion and infinity merely descriptive or modelling of a physical feature of our biological makeup or is it to be taken more literally as an actual feature of linguistic knowledge and generation? And if so, how? In section 7, I will offer such an explanation on their behalf but for various reasons they may be less inclined to adopt it. In section 9, I will discuss these issues in somewhat more detail.

## 5 Just How Big is NL?

An interesting offshoot of much of the discussion of infinity and recursion in linguistics has led to a subsidiary question as to the exact cardinality of natural language (Collins 2010, Hinzen and Uriagereka 2006, Kornai 2014, Pullum and Scholz 2010). This ancillary debate has in turn led to controversy between Conceptualists and Platonists as to the correct foundations for the science (Langendoen and Postal 1984, Katz 1994).

The standard assumption is that certain structural elements of natural language such as embedded clauses, adjectival modification (i.e. seemingly endless iteration of “very” and the like) and conjunction allow us to produce a proof of discrete infinity along the lines of the integers.

By the same logic that shows that there are an infinite number of integers - if you ever think you have the largest integer, just add 1 to it and you will have another - there must be an infinite number of sentences (Pinker, 1994: 86).

I will not go into the details of this “logic”. Suffice to say that no such analogous proof has been given to date.<sup>7</sup> Nevertheless, the corollary of this claim is that if natural language is infinite,

<sup>6</sup> As pointed out to me by an anonymous reviewer, Merge itself should not be thought of as yielding infinite output, as a source of recursion or link to arithmetic but rather recursion must be discovered in the atoms and inherited by the objects constructed by the Merge operation. There could be two reasons for this caution. Firstly, in allowing infinitely many Merges you can generate infinite objects, but  $\aleph_0$  is not an integer itself. Secondly, if we have a constraint such that  $\text{Merge}(\alpha, \beta)$  can only be applied if there is neither an  $\alpha$  nor a  $\beta$  present in either  $\alpha$  or  $\beta$ , Merge will run out of options resulting in a finite number of constructable objects.

<sup>7</sup> And as Pullum and Scholz (2010) convincingly argue, no such proof can be given *sans* the induction axiom and successor function in the case of natural language.

it is equal in cardinality to that of the set of positive integers or natural numbers, i.e. discretely infinite or  $\aleph_0$ . The two main assumptions that lead to this conclusion is (1) that every sentence of natural language is of finite length and (2) that there is no longest sentence.

The reason I mention this cardinality debate (which many linguists take to be profoundly uninteresting) is that it presented a very clear case of mathematisation. In fact, some Platonists challenged the Conceptualist view of cardinality not based on its inappropriate analogy with arithmetic but rather its allegedly erroneous conclusion thereof. Specifically, Langendoen and Postal (1984) produced a formal proof which argues for the nondenumerable infinity of natural languages (i.e. strictly greater than the set of natural numbers). This possibility was meant to attack the heart of the competence model of grammar. Although it has been largely neglected in the literature, I think it is pertinent to the issue at hand since it involves explicit mathematisation of infinite generalisation. According to this result, the cardinality of natural language is not even a set but a “megacollection” (or “proper class” in set theory). If this is the case, it is argued, then the Post canonical system procedure or the generative grammar one is inadequate in determining or enumerating the set of natural language sentences since sets have fixed cardinalities (and language does not). Thus, these procedures cannot capture the magnitude of natural language.

I cannot reproduce this proof in any detail here which involves the construction of coordinate compounds (unbounded conjunctions) in an attempt to mirror the subset relation and allow for a Cantorian style proof by contradiction on the set of sentences (I refer the reader to Langendoen and Postal (1984) or for a shorter overview in Katz (1985)). They achieve this result by rejecting (1) above or the claim that sentences are finite in length which in turn allows for infinite sentences and conjunctions of infinite sentences (or clauses). Later amendments to the proposal (Langendoen 2010) attenuate this stance to claim that any “cap” on the cardinality of NL from the possibility of transfinitely many sentences can only be countered by linguistic evidence.<sup>8</sup>

What is significant for the present topic is the presumed consequence of the so-called “Vastness Theorem”. The main consequence of this theorem is supposed to be that all generative (constructive or proof-theoretic) grammars are rendered useless in characterising NL from the onset since they assume a denumerable infinity (like the natural numbers). As a result, a large research paradigm in linguistics should be abandoned. Non-constructive grammars would have to supplant the generative methods (see Pullum and Scholz (2001) for a discussion of the cardinality neutrality of model-theoretic approaches to grammar). In addition, grammars cannot be about physical brain-states as *per* the biolinguistic paradigm since this would be mathematically impossible. Brain-states are finite (whatever this means exactly) and without generative procedures for arriving at the requisite infinity (since these max out at  $\aleph_0$ ), unlikely objects of our grammars, i.e. generative grammars can never describe linguistic reality fully. Put in another way, generative grammars supposedly provide us with a bridge from the finite to the infinite in linguistics. If the full complexity of natural language constructions exceed the reach of generative procedures, then linguistic competence cannot be identified with the target of grammatical theory since competence is only a proper subset of natural language.

We will see why the above reasoning is flawed in the first grade of involvement. Let this section suffice to show the extent to which the infinity postulate has been taken in linguistics.

## 6 Three Grades of Mathematical Involvement in Linguistics

So far, we have seen a gradation (perhaps descent) of mathematical involvement in linguistics or *mathematisation* as I have called it. In what follows, I hope to impose some order and argue for sober methodological reflection on these issues. I shall approach the mathematisation of linguistics *via* a strategy of identifying grades of mathematical involvement for the grammars of linguistic theory, following a similar strategy proffered for modality by Quine (1976). The hope is that this will provide compelling argument as to which grade offers the best home for linguistic theory while avoiding the pitfalls of both conceptualism and Platonism. With each grade a further methodological burden is introduced. Importantly, however, as we shall see, this progression does not follow the traditional landscape of the debate in theoretical linguistics. The first grade is not just another

<sup>8</sup> This situation then moves into the realm of different cardinalities for different languages, some transfinite others finite. I thank an anonymous reviewer for directing me towards this possibility.

label for nominalism, grade two for Platonism and so on. Although, the motivations behind Platonism seem to find a good place within the bounds of the second grade, this level of involvement by no means entails the Platonistic positions of Katz or Postal, as I hope to show. One way to think about the grades is that they represent a cluster of theories each with more commitments (in terms of connections to the target system) than the last.

Within the first grade of involvement, the mathematics involved in grammar construction is merely a helpful aid and not directly structurally committing to the target system. The first grade of involvement proposes a quasi-instrumentalist picture of linguistic methodology (although this is not necessarily the case, a structural realist interpretation of linguistic modelling might also be feasible, see Nefdt (2016b)). On this account, grammars are indeed scientific in nature (à la Chomsky) but more akin to models than theories (*contra* Chomsky). On the second grade, grammars are theories. However, on this level of mathematisation, grammars are mathematical or formal theories on the level of set theory, arithmetic or universal algebra. In essence, the mathematics involved in various grammar formalisms is enough to establish the reality or existence of linguistic objects without further empirical consideration. The mathematics is part of the linguistic target system itself. On the last grade, grammars are representational devices and directly represented (cognised, known or embodied) by speakers of the language. They are theories of linguistic competence. This view goes beyond mere modelling, for various reasons I will show. For now though, it is enough to understand that this grade does impose systematic structural constraints on the reality which the grammar describes, in addition to material preservation and correspondence to internal mechanisms.<sup>9</sup>

## 6.1 Ontology and Methodology

It is not my purpose to deny that there is a relationship between ontology and methodology in the sciences. It seems reasonable to expect that what something is has an important effect on how it should be studied. However, I think the relationship between the two is much looser than assumed in much of the linguistics literature. The problem has been diagnosed elsewhere (George 1989, and more recently McDonald 2009). The idea is that what determines the separation or subject matter of a discipline is not only the ontological status of its objects but also its approach or methodology.

For instance, some Platonists argue that the claim that sentences (unlike utterances) are not located in space or time (when and where is sentence “The dog ate my homework”?) entails that linguistics is a science of *abstracta* (Postal 2003). However, a linguist could grant that sentences are not located in space or time, and are therefore abstract objects or types, without conceding the point that linguistics is like mathematics. The reason for this is that the question of with which discipline linguistics is aligned is not only a matter of ontology. It is at least in part a methodological question. In other words, how the abstract objects, i.e. sentences, are studied and employed also determines the science, e.g. as abstract objects in themselves or as convenient theoretical entities for scientific study. “It is not clear that having abstracta in the domain of a science is sufficient to make a science formal and nonempirical” (McDonald, 2009: 294). Put in another way (following George 1989), astronomy involves attributing numbers to the planets, for instance there are nine planets in the solar system (if you still have faith in Pluto). In a sense, astronomy involves *abstracta* since numbers are abstract objects but it is not concerned with abstract objects *qua* abstract objects as in mathematics. Neither astronomers, nor linguistics (I shall argue), are directly interested in the properties of abstract objects but instead their concerns lie with how abstract objects might be related to natural phenomena or best *model* such phenomena, which is an empirical question (like the number of planets). George (1996) goes on to state that mathematical entities (such as grammars) can be identified empirically. He states that “[i]f I want to know the trajectory of a particle, I am engaged in an empirical inquiry whose goal is to identify a particular function” (George, 1996: 300). Similarly, identifying grammars is an empirical matter. I think that

<sup>9</sup> Yablo (2013) offers a related account of the three grades of mathematical involvement for scientific explanation. Yablo’s three grades are roughly and respectively defined as follows: on grade one, mathematics has a descriptive role (something like the first grade on my view but more limited), on grade two it has a structural role and on the third grade it has a substantive role. He attempts to capture the substantive role in terms of a modal notion of *extricability*. We can think of extricability in terms of logical subtraction. “Logical subtraction sometimes yields a well-defined remainder, surely. Snow is cold and white - Snow is cold = Snow is white, I assume” (Yablo, 2013: 1014).

this perspective is too narrow for a number of reasons. For one thing, it suggests a bottom-up perspective on scientific and linguistic investigation. However, in many cases mathematical entities (or models) were already developed independently (such as Post canonical systems) and then applied to natural phenomena. Thus, the situation is more like using a particular function to model the empirical target of the trajectory of a particle. Nevertheless, George's analysis does suggest a useful perspective on the applied nature of linguistics and the status of mathematical entities within this enterprise.

Linsky and Zalta (1995) take the above suggestion further in claiming that abstract objects are not only a convenience for scientific theorising but a necessity. Their notion of abstract object is divorced from the usual Platonic concept but the idea remains that most sciences require some sort of abstract level of interpretation, whether it is the relatively benign use of numbers for characterisation in astronomy (and everywhere else) or the idea of species or types of animals in biology, *abstracta* of some sort seem to be a conceptual necessity.

Before entering into the discussion of the grades, I shall briefly outline the idea by means of a popular map analogy in the philosophy of applied mathematics.

## 6.2 Finding your way in New York City

Let us imagine that some person, let's call her Sarah, is planning a trip to New York City. Sarah is the kind of traveller who enjoys consuming as much knowledge about a city before arriving as possible. Let us imagine further that she is on a budget and would like to see as many places across the five boroughs as she can during her visit. Thus, she is particularly interested in getting around by means of the New York subway system. So, she directs all of her efforts to understanding it prior to her arrival.

Now, there are different aspects of the system which might be of interest to a traveller. Someone could want to know the average time it takes to get from one point in the system to another, which stations are closer to one another and which routes need to be taken to get to different stations. One could also be interested in what the subway trains are made of, how they achieve their maximum velocity, what kind of mechanisms are involved in the braking system etc. One might say that there are different grades of answers which Sarah might aim for in her understanding of the transit system.

A convenient way of viewing the grades of involvement are along the lines of the following (nested) claims.

- $C1'$  : Structures of the representational device (phone GPS, subway grid map, picture etc.) are weakly structure preserving to the structures of the actual subway system.
- $C2'$  : Structures of the representational device are ontologically committing or substance equivalent to the subway system, i.e. made of the same stuff.
- $C3'$  : Structures of the representational device include the actual mechanisms involved in the workings of the subway system (trains, sliding doors, ticketing etc.).

Now in terms of our example, Sarah could be asking a number of kinds of questions relating to the claims above. For instance, she could want to know the average time it takes to get from Central Park in Manhattan and 81st (the National History Museum) to Flushing Avenue in Brooklyn. In terms of  $C1'$  this would require some sort of map of the subway grid with a scale corresponding to real distances. If she were interested in an answer along the lines of  $C2'$ , she might require a 3D model of some sort (ideally made of the same substances as the real subway) as can be found at the New York Transit Museum (incidentally located in Brooklyn).<sup>10</sup> Asking a question would then involve running a real world simulation and viewing the result.  $C3'$  would require information about how the trains actually get from Manhattan to Flushing Avenue, i.e. the inner working of their electrical makeup corresponding to how they traverse the system.  $C3'$  might also involve the exact routes and paths (e.g. orange line, green line and how they connect) needed to get to and from these places. In other words,  $C3'$  requires a mapping between the distances of the stations

<sup>10</sup> See Weisberg (2013) for a discussion on how the large scale Bay Area Model of San Francisco assisted scientists in rejecting the proposal to build a dam in the Bay Area.

and the workings of the trains and routes in getting from one to the other which could include some physics or engineering information.

There is of course a way of answering the question about the distances of this route and all the possible routes without direct recourse to any of the claims above. There could be a graph which accurately represents the distances between all the stations in the NYC subway system without respecting their “actual” routes, directions or interconnections. This graph would not be a map in the normal sense, since you couldn’t use it to find your way from one station to another. You could use it to accurately know the distance from Central Park to Flushing Avenue or from any place in the five boroughs to any other (within the subway system). Similarly, if you wanted to know the average time from one station to the next. All that needs to be preserved in this graph are the relative spatial and temporal relations between the stations and that structure is multiply realisable. For instance, getting to and from our designated place in Manhattan to the place we specified in Brooklyn could (and does) involve changing lines. Our graphs would neglect this detail. Taking the idea even further, there could be a permutation of the transit system of New York such that if we map NYC stations to stations in, say, the Kiev metro system according to which we would preserve the distances between the stations and average time between them. Now consulting the graph specified (or the Kiev metro grid) will not give us an answer in terms of  $C2'$  or  $C3'$  but it will *indirectly* track the spatial and temporal information (or structures) in which Sarah might be interested, in this case distances and times (abstracting away from delays, commuter congestion etc.). I hope to show in section 7 that grammars operate in this indirect manner and provide a wealth of information despite their indirectness.

Specifically, the grades, I wish to propose, can be characterised in terms of the following claims.

- C1 : Structures of the grammars *model* the structures of natural language (or the linguistic competence thereof), i.e. weakly structure preserving.
- C2 : Structures of the grammars are material preserving to the structures of natural language, i.e. the grammar is comprised of the same substance as the language.
- C3 : Structures of the grammars track actual mechanisms involved in language processing and comprehension.

The first grade of involvement is committed to  $C1$  in some weak form (while in principle being compatible with all of the constraints). The second grade of involvement is committed to  $C1$  and  $C2$  while the third is committed to all three constraints.

Notice that being a grade one advocate for the NYC subway system can yield genuinely useful insights. Sarah will do fine in planning her trip with using a graph or map which only represents the distances, average time, and perhaps directions (maybe just a graph containing vectors) whether or not this graph actually shares the spatial relations of the system or its material makeup. Nor was her graph arrived at by magic. Such a representation could have been devised for various reasons, perhaps it involves a simpler representation than an isomorphic structure would do or the pattern it uses is more user-friendly pictorially, perhaps Sarah knows the Kiev metro system better.

In the following, I will show that going according to the grades of involvement tells a story about mathematisation in linguistics. In other words, the grades track methodological claims rather than purely ontological ones but in so doing shed light on how the ontological claims developed from mere physicalism to abstract Platonism.

## 7 The First Grade: Models and Linguistic Reality

There is a growing literature on the nature of scientific modelling in philosophy. Although the role of models and their connection to the scientific enterprise has been much less explored than various accounts of or against scientific realism or the demarcation problem etc., modelling should be of particular interest to the linguist who, I will argue, faces a similar task to the empirical scientist in attempting to account for a natural phenomenon fraught with complexity by means of smaller more tractable representations of it.

Simply put, the first grade of involvement places our linguistic grammars at the level of scientific models. Their core aim is to capture salient features of linguistic reality, not necessarily to represent it in its entirety. This is achieved by various abstractions and idealisations, one of which is the

notion of a “generative grammar” for modelling linguistic creativity. But let’s not jump ahead. Perhaps unsurprisingly, this position is much less controversial within the computational linguistics literature. “It is clear that to the extent that linguistic theories, i.e. grammars, aim to capture human knowledge of language, these theories are formal models” (Tiede and Stout, 2010: 147).

When a computational linguist provides a stochastic model or attempts to represent the next word in a grammatical chain as a finite Markov process, she is not necessarily making a claim about how human beings actually parse expressions (i.e. the exact route and train between two stations in the subway system). The model could have other evaluative benefits, such as predictive capabilities or efficient parsing complexity. Similarly, the model could be implemented in machine translation or other natural language processing uses. Explanatory models also work for a number of reasons in ways that do not correspond to strict adherence to *C1*, *C2* or *C3* above. I attempt to show that this is a property of linguistic models or grammars in the following subsections.

My argument is that not only do grammars genuinely share a number of properties with scientific models but also that the only way to maintain the conceptualist approach of ascribing both infinity and a physicalist ontology to natural language is by accepting that grammars are formal models of a target system, in this case linguistic competence. However, this grade of involvement is technically compatible with both a nominalistic and Platonistic ontology, i.e. grammars could be modelling idealised linguistic tokens (the output of linguistic competence (Devitt 2006)), or an abstract mind-independent linguistic reality (Katz 1981, Postal 2003).

## 7.1 What are models?

So, what are models? And how do they relate to reality? One place to begin is by appreciating how modelling differs from other types of scientific theorising. The basic idea is that a model is an *indirect* representation of a target system or some aspect thereof, in this case natural language. The model bears certain resemblance relations to the target system such that stipulations within the model reflect aspects of the target system. For Godfrey-Smith, “the modeler’s strategy is to gain understanding of a complex real-world system *via* an understanding of simpler, hypothetical system that resembles it in relevant respects” (2006: 726). My claim in this section is that a grammar is precisely this sort of device and therefore that linguists find themselves in the modeler’s position with relation to natural language. By designing grammars which generate or constrain the grammatical output of a given language, linguists create small hypothetical systems which reflect or resemble structural descriptions of that language *via* rules that comprise the grammar. In this way, models, or grammars in this case, are theoretical intermediaries.

In order to see how this works, we should appreciate that mathematical models are essentially abstractions. They are abstract objects. They are designed to simplify a target system which otherwise would be too complex to approach scientifically (i.e. precisely). They might have various aims, simplification is one, explanation might be another, prediction yet another. Sometimes these aims can come apart. In his classical treatment of scientific modelling, Giere (1988) held that models were idealised structures (or abstract objects) aimed at representation of the target system in the real world. As previously mentioned, these structures or model systems bear resemblance relations to the target system. We might be interested in structural relationships such as various morphisms to capture this resemblance as in *C1* (although Giere preferred a less formal account of the relation).<sup>11</sup>

The above picture is not to be confused with the so-called “semantic view” in the philosophy of science (Suppes 1960, van Fraassen 1980).<sup>12</sup> We would do well to distinguish two different senses of the word “model” here. In mathematical logic, a model is a set-theoretic entity with a domain of elements (or universe) and a relation which holds between those elements. “A model, basically, is a set of objects (and relations between them) that functions as an *interpreting structure* for a

<sup>11</sup> Another way to think of what a model is involves an analogy with fictional worlds, pretenses or ways that the world *could* have been (Frigg 2010). This view breaks down the connection with model theory in mathematics. In this way models are akin to the fictional worlds of Sherlock Holmes or Luke Skywalker. Counterfactual analyses are also generally connected to the type of representation involved in modelling. For instance, Giere (1988) affirms that model systems are systems which would be concrete if they were in fact real.

<sup>12</sup> As opposed to the then popular “syntactic” accounts in which scientific theories were considered to be consistent sets of sentences in formal languages (“theories” in the logical sense).

set of sentences.” (Godfrey-Smith, 2006: 727). Suppes (1960) held that scientific models and the logician’s set-theoretic models were one and the same. It is, thus, possible to talk of “truth in a model” or under an interpretation. The central concepts here are “truth” and “satisfiability”.

In the ensuing discussion, however, we will follow the literature in conceiving of models as belonging to a heterogeneous class of objects which includes physical models (used in biology and chemistry), scale models (used in engineering), computational models (used in population studies and computer science more broadly) and mathematical models (used everywhere including linguistics). Mathematical models, of which grammars are a proper subset, can be conveniently conceived of as abstract objects (whether they are set-theoretic, logical theories or fictional worlds).

## 7.2 Multiple Models: Against Strong $C1$

There are a few ways in which a model can respect  $C1$  or a structure preserving mapping. Assuming a structure for the target system, we could require there to be a strong morphism, such as an isomorphism or homomorphism, between the model and the target. But as Frigg (2010) notes “[i]n order to make sense of the notion that there is a morphism between a model system and its target we have to assume that the target exemplifies a particular structure” (254). Thus, on this reading of  $C1$  we are forced to attribute a structure to the target system, i.e. natural language. This, however, is an unavoidable prerequisite for the modelling process.

The choice between structure mappings is also not arbitrary. For instance, establishing an isomorphism places a much stronger constraint on the relationship between a model and a target system than does a homomorphism or weak homomorphism, since an isomorphic relation requires a homomorphism and an inverse morphism. The point is that some kinds of morphisms will place an added burden on the model in terms of its relationship to the target system. For instance, if a strong morphism is required (such as an isomorphism), then all of the elements of the target system will have to be interpreted into the target system in a structure preserving way. Compare the structural relationship posited between two models connected by bisimulation in modal logic. In this case, the two models make the same modal formulas true despite potentially diverging greatly in their internal makeup (the simplest case involves a model with a one world cycle and a model with two worlds accessible to each other). I take a weaker version of  $C1$  to capture the “modelling” relation in which features of the model produce relevant answers to questions about the target system (similar to Bromberger’s (1989) idea of projectible questions to which the model and modelled receive the same answers but this view fails to surmount the problem of idealisation described below).

Standardly, as we have discussed above, models involve abstractions and omissions from the target system. Of course, we could have a circumscribed domain or in our case a proper subset of natural language, such as syntax, as a target. In this case we would need to establish a relation between the model structures and the syntactic structures of various natural language constructions. Again, the choice of mappings becomes important.

The considerations above point to conceiving of  $C1$  as introducing a range of relationships between the model and the target system. On one side of the range, are the strong structure mappings and on the other rather weak relations. In all of the cases, the theorist or modeller is attempting to establish a correspondence between the structures of their models and the structures of the target system.

The reason strong  $C1$  often fails for models is due to one of the most common properties of models, they are multiply realisable. This property goes in both directions. Different models with different structures can be used to model the same phenomenon and the same model can be used to model different phenomena. In terms of the latter scenario, consider a mathematical model of a pendulum and an identical model of a certain circuit.

The mathematical structures view seems committed to identifying both the pendulum model and the model of the electrical circuit with the mathematical structure they have in common and, thus, to insisting that the pendulum model and the model of the circuit are one and the same model (Thomson-Jones, 2012: 768).

Similarly, strong  $C1$  identifies grammars by their structural/mathematical properties. But one could conceive of the same aspect of a grammar modelling two distinct natural language constructions (as in the pendulum and circuit case). Philosophers of science are wont to find additional

means of model individuation. For Thomson-Jones (2012), models are constituted by sets of propositions. For Weisberg (2013), the modeller’s construal of the model differentiates between models with the same structures. Nevertheless, more than just structural equivalence is often needed to distinguish between models.

More pertinent to the case of linguistics is the scenario in which the same phenomenon can be represented by multiple models with non-equivalent structures. This is related to the famous problem of equivalent grammar formalisms, initially presented in Quine (1972). Take two weakly equivalent grammars, phrase-structure grammar and tree-substitution grammar, for instance. These two grammars generate the same sets of sentences or in Quine’s terms are “behaviourally equivalent”. The problem is that since they are both empirically adequate or generate the same sentences, there is no way of deciding which grammar is the correct description of the target (in Quine’s critique, the target would be mental states of language users). This argument serves to question the need for a model or a grammar to respect anything as strong as the constraint on structure preservation exemplified by isomorphisms or homomorphisms etc.. This is also not a merely theoretical worry. Recently, there have been a flurry of formal proofs of weak or expressive equivalence of various syntactic formalisms such as tree adjoining grammar (Joshi), generalized phrase structure (Pollard) and categorial grammar (Steedman). Furthermore, Chomskyan syntax can also be shown to be equivalent to these formalisms (see Michaelis, 2001; Mönnich, 2007) (or rather Stabler’s (1997) formalisation and interpretation of minimalist syntax has been shown to be equivalent to some of the above grammars).

There is another related worry stemming from the modelling literature. In Weisberg (2007b), a particular kind of modelling strategy is discussed, namely *multiple models idealisation*. This practice involves constructing many connected but incompatible models each of which focuses on one or more aspects of the target. This strategy differs from other kinds of idealisation “in not expecting a single best model to be generated” (Weisberg, 2007b: 646). Naturally, strong structure preservation or one-to-one correspondences are not appropriate within this practice. Since scientific theories can have diverse goals such as accuracy, simplicity, predictive power etc. and the construction of one model to fit all of these criteria necessarily involves “tradeoffs”, this approach offers the theorist a way of meeting all of these objectives separately. “If a theorist wants to achieve high degrees of generality, accuracy, precision, and simplicity, she will need to construct multiple models” (Weisberg, 2007b: 647). This practice is common in climatology, ecology, biology and population studies.

If we consider the various models used in the service of linguistic theory and aimed at natural language, this picture seems to further militate against adherence to a stronger interpretation of *C1*. Pragmatic models such as Stalnaker’s model of common-ground and the conversational context or Lewis’ scorekeeping in a language game are distinct from optimality theoretic formalisms of phonology and generative grammars for syntax. Even if we stick to syntax and semantics, generative enumerative syntax (based on proof-theory) and model-theoretic semantics (based on model-theory) embody distinct mathematical properties and formalisms, yet they both serve to capture an interconnected part of linguistic reality. Within syntax alone, we could conceive of the nested adequacy conditions of Chomsky (1965) as multiple models with distinct goals. Although these models might not be incompatible, there can be no notion of a strong structure preserving mapping (between model and target) when multiple models with distinct structures are being used to model a single target.

Thus, the idea that models or grammars have to preserve the (assumed) structure of the target system through strong mappings or morphisms is too strict and not a necessary condition on linguistic modelling. Adherence to a weaker notion of “model” is all that is required.

### 7.3 Idealisation: Against *C2*

In this section, we will see how the rejection of *C2* or material preservation can rescue the Conceptualist position from inconsistency.

Scientists often do not associate similarity of their models and the target system with “truth” or even approximately true descriptions. For instance, in biology, Fisher described a model of fictitious three-sex organisms to explain the emergence of two-sex organisms involved in sexual reproduction



(Weisberg, 2007b: 223). In physics, Boyle’s law is usually explained by assuming that gas molecules do not collide. This is not strictly true since low-pressure gases do collide. They do not tend to reflect the collisions in their behaviour with the result that these collisions are not admitted into the model (see Strevens (2007) for discussion). And in linguistics, we are asked to consider an idealised linguistic community of speakers and hearers who know their language perfectly and are never error-prone (Chomsky 1965).

The tools of models are abstractions and idealisations of various sorts (as in other types of scientific theorising). These terms are sometimes used interchangeably in the literature. The basic idea behind these techniques is simplicity or tractability. Abstraction generally involves the removal of extraneous or superfluous material of the target system in the model. Idealisation sometimes additionally involves distortions of the real world such as the cases mentioned above.

The details of these strategies are not of particular importance at this stage.<sup>13</sup> However, there are some important features of idealisation which are relevant to the discussion at hand. For one thing, it is not ontologically committing as *per* C2, especially in the non-alethic cases (strictly false models). Fisher’s three-sex model does not commit the target system to the existence of three-sex organisms, nor does Boyle’s law commit the physical world to non-colliding gas molecules and importantly nor does Chomsky’s model commit us to idealised speakers. Even more simply, a physical model of DNA does not commit the structure of DNA to the specific dimensions or the material composition of the model, e.g. a metre in size and plastic or styrofoam in composition. Secondly, as previously mentioned, models need not be directly related to the target system under study. In economics, von Thünen proposed a model of an isolated state on fertile land cut-off from all communication and contact with the outside world by a barren wilderness around its borders. Mäki claims that the false assumptions underlying this model serve the purpose of “neutralizing a number of causally relevant factors by eliminating them or their efficacy” (2011: 50). Thus, actually true and causally relevant aspects of a real economy are removed from the model.

In a similar fashion, grammars understood as models do not commit us to any specific ontology of natural languages. In the syntactic models of FLT, we treat sentences as sets of (uninterpreted) strings. This idealisation in no way commits us to natural language expressions being strings or sets of strings. In the semantics literature, meanings are designated by functional types. Meanings are not mathematical objects or functions in reality.

Tiede and Stout (2010) go further in claiming that we are not committed to natural languages being discretely infinite just because recursion is a feature of our generative grammars (this is true for technical reasons as well, as recursion does not guarantee discrete/denumerable infinity). Discrete or denumerable infinity is assumed or a “modelling choice” on their view (and mine). The features or properties of the target system which we want to represent in our grammars are productivity, systematicity and conciseness. The first two properties are familiar from the literature on compositionality (presumably they would hold that this principle too is a modelling choice). The last property is assumed to prevent overgeneration of grammatical expressions.<sup>14</sup>

Now it is clear to see how the Conceptualist can have his cake and eat it too. The “core” property of linguistic creativity is modelled as discrete infinity, the latter being the element of our models of natural language competence. Here we can freely employ idealisations such as sets of uninterpreted strings as sentences governed by recursive rules. The target system, however, is not committed to the ontology of the model, i.e. sets and functions etc. Therefore, the target system can still be a physical object or brain-states of individual language users.

Postal (2009) discusses this possibility briefly. He claims that to understand “infinite generation” or recursion as idealisations of some sort is to equivocate on both the terms ‘idealisation’ and ‘recursive’. In contrast to idealisations such as frictionless planes in physics, this idealisation is more akin to “one which claims the solar system has an infinity of planets” (2009: 110). Postal deems such idealisations “silly”. On my account, if such an idealisation were useful to a physicist or astronomer or helped understand some other property of the solar system, then it would be

<sup>13</sup> See Thomson-Jones (2005) for general discussion. Also see Stokhof and van Lambagen (2011) for discussion with relation to linguistics and Nefdt (2016a) for a response

<sup>14</sup> Here they borrow from Savitch (1993) who shows why we might assume that languages are (essentially) infinite despite having no evidence for them not being simply largely finite. Savitch’s paper is a formal attempt at capturing parsimony judgements in grammars, i.e. we treat finite sets as *essentially* infinite if this allows us to get simpler descriptions than we would if we treated them as finite.

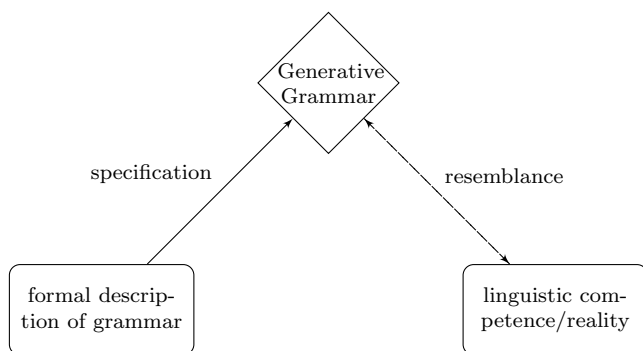


Fig. 1 \*

a perfectly acceptable aspect of a model (statistical cosmology is full of such idealisation). See footnote 14 for a reference and suggestion as to why an infinity assumption could be a simplifying tool for a linguist even if the target system is in fact finite. In addition, physics abounds with such idealisation. See the example of the thermodynamic limit in section 2.

If we maintain a separation between models and linguistic reality, the former being capable of the abstraction required for infinity statements and the latter being capable of physical description, then the Conceptualist movement can be rescued from alleged “incoherence”. See the figure below for an illustration of the current picture of linguistic modelling (adapted from the general scientific picture in Giere 1988: 83). There are two salient relations here, the first is between the formal description of the grammar and the grammar as a model itself, i.e. recursive phrase-structure rules or constraints on feature structures etc., and the second is between the grammar and the target system which it needs to resemble in some way (the dashed line indicates that this resemblance relation is intentionally left vague).

According to the above diagram, the formal descriptions of the grammar might involve things like MERGE or recursion thus committing the grammar to discrete infinity (or the capacity for such cardinality) but the target system is in no similar way committed. For example, in discussing the question of the size or cardinality of natural language(s), Langendoen (2010) claims that “from the fact that one’s grammatical model is closed under such an operation [iterative or recursive operations], it does not follow that the language it models is” (2010: 140). This would require an additional argument or proof (he does, however, go on to attempt to offer such an argument).

Langendoen’s claim corresponds to the picture above in which the grammar aims to resemble or model certain aspects of the target system, such as our ability to process and produce previously unheard utterances or the fact that there seems to be non-arbitrary cut-off point for creating distinct expressions. The resemblance relation could go in both directions, bottom-up or top-down, but in neither case does it commit the target system to formal features of the grammar in ontology as in *C2*. Hence, there is no incoherence here and this picture is derived from a similar view suggested for the rest of the sciences. In addition, if we return to the idea of “reverse-engineering” in the claim by Boeckx in section 3, we can appreciate the error of mathematisation in a new light. Attempting such a “reverse-engineering” assumes that the formal features of the model, such as recursion, requires an ontological interpretation in the target system such as an account of its evolution within the language faculty (we will return to this in section 9).

Grammars are abstract objects with an number of formal mathematical properties, these properties do not necessarily pertain to their linguistic targets in any ontologically significant way.

#### 7.4 Lewisian Modelling: Against *C3*

Much like *C1*, a range is introduced by *C3* which starts from tracking specific linguistic mechanisms to more general cognitive mechanism involved with language (but also possibly other cognitive processes such as memory, movement, planning etc.). It is not clear, however, that a grammar needs to track any mental happenings whatsoever. The target of linguistic grammars could be outward linguistic behaviour or patterns which emerge from communities of speakers. In a sense, the individual idiolects (or I-languages) could determine the nature of these patterns, at least in part,

but they could also be partly determined by external mechanisms such as linguistic conventions. Of course, different targets might in fact make different models desirable. For instance, on the Lewis' (1975) view, which will be discussed in this section, the explanation of subsentential elements of the grammar is of ancillary importance.

Before we discuss Lewis' view which has held sway amongst philosophers of language more so than theoretical linguists, let us briefly consider another example of how a model might not aim to track internal mechanisms responsible for production. A theorist might be interested in how people go about solving particular multiplication problems. There are two models which might be of interest here, standard "times table" multiplication or the more complex binary multiplication. There could be various reasons for preferring one model to the other such as respective processing times. The algorithm for binary is the same as that of standard times table or decimal multiplication but operates *via* three manoeuvres, namely  $0 \times 0 = 0$ ,  $1 \times 0 = 0$ , and  $1 \times 1 = 1$ . So in binary, you merely replace symbols with other symbols with no carrying over as in decimal multiplication. The results are inter-translatable. In a sense then, it does not matter with which you choose to model arithmetic performances, since despite the different methods they turn out to be equivalent. In fact, mental multiplication might be "truly" captured by neither method. Nevertheless, the models could represent not only the results of such a mental calculation but other features such as timing correlations or common error explanations. Therefore, the actual mechanisms involved in multiplication might be quite different from the structures of the model and yet the model might indirectly correspond to those mechanisms nonetheless. In a sense, both methods exist at Marr's first level of the description of computational processes in which the function computed and the reasons why are stated. His second level states which algorithm actually computes the function. We will consider positive views against interpreting grammars in terms of *C3* in section 9 but for now, we can simply appreciate that such a requirement seems unnecessary.

Furthermore, with relation the claim that grammars track actual mechanisms involved in natural language cognising, there is a more compelling reason to think that *C3* is superfluous. Lewis (1975) offers an account of how human beings use languages, construed as abstract objects, which is in the spirit of the first grade of involvement presented in this section. For Lewis, languages are abstract objects or functions which assign meanings to sets of strings (sentences). A language is then utilised by a community of speakers if and only if there is a convention in that community of *truthfulness* and *trust* in that language. The precise definitions of the terms are not important for present purposes.

In terms of Fig. 1, the reality which is being modelled is linguistic communication. The formal description is given in terms of functions and sets of sentences (a grammar) and the resemblance relation is provided by a notion of "convention" which allows for a given formal object (or grammar) and not another to model linguistic communication accurately.

For an abstract language to be "realised" on this view, is for a community of speakers to possess a convention of some sort in that language. Sounds and sentences get their meanings relative to such a community of speakers and a meaning function (abstract language). However, there are a number of issues with taking an uncharitable reading of this proposal, i.e. taking it to be literally about the realisation of nonspatio-temporal abstract objects in the real physical world. In terms of the dialectic of this chapter, one might wonder how exactly this "realisation" relation is to be construed on a literal reading in which abstract objects are considered to be outside of the causal nexus. In other words, how does a community of actual speakers access a function from sentences to intensions or *abstracta* in the necessary way?

On the modelling view, the above relation is unproblematic. Languages are merely modelled as functions from sentences to intensions. Yalcin has a similar interpretation in mind with relation to formal semantics.

Semantic theory is not interested in the semantic value of properties of these abstract objects *qua* abstract objects. Rather, it is interested in an aspect of the question which of these abstract objects well-models what it is one knows, when one knows a language (2014: 36).

Yalcin goes on to remonstrate against Lewis' impoverished notion of a language, especially its lack of the property of productivity (which he [Yalcin] considers to be a central desideratum of semantic theory). I think that the reason for this is that Lewis was rejecting *C3*-like reasoning and instead opting for an explanation of how a linguistic community uses or realises a public language

(assigns meanings to its sentences).<sup>15</sup> Therefore, the model was impoverished or rather simplified for this purpose.

In terms of grammars, Lewis (1980) opts for a more broadly construed sentential account which importantly casts doubt on *C3*.

I use the word ‘grammar’ in a broad sense. Else I could have found little to say about our assigned topic. If it is to end by characterizing truth-in-English, a grammar must cover most of what has been called syntax, much of what has been called semantics, and even part of the miscellany that has been called pragmatics [...] You might insist that a good grammar should be suited to fit into psycholinguistic theory that goes beyond our common knowledge and explains the inner mechanisms that make our practice possible. There is nothing wrong in principle with this ambitious goal, but I doubt that it is worthwhile to pursue it in our present state of knowledge (81).

This is compatible with grammars being scientific models (or abstract objects in Giere’s sense) aimed at representing only the class of linguistic objects or languages usable by various communities (and a few innocuous extras). The grammars do not seem to specify linguistic competence or *abstracta* in the traditional sense nor “inner mechanisms” of language users, rather they specify functions that can describe or model the linguistic behaviour of language-using communities.

Lewis’ position (as with my own) can be seen as a conciliatory intermediate position between the two different ontologies mentioned in the previous sections. In a sense, both positions, Conceptualism and Platonism, are correct. They both tell a part of the story. Understanding the whole story involves appreciating how these ontologies connect with one another. If we appreciate a language as an abstract object in the sense of it modelling the patterns of speakers/hearers, we can do so. As Yalcin mentions, we are not interested in languages as abstract objects *qua* abstract objects but rather as formal tools for modelling linguistic behaviour. So in this sense, a language is a formal object picked out by a linguistic community by means of the finite rules of the grammar. It might be somewhat misleading to describe the view thusly. In the parlance of scientific modelling, a natural language, which is an abstract object for Lewis, is used to model the particular linguistic conventions of a given community. Substitute Lewis’ *language* for grammar here and his *languages* for rule-governed linguistic behaviour or competence and you have the first grade of involvement for linguistics.

By taking the scientific modelling route, as I have, we can explain (in Lewisian terms) how a language can both be infinite and finite. The *Language* is given to us by the formal grammar which can involve recursive functions and thus discrete infinity while *Languages* we use in our daily lives, constitute the finite expressive needs of linguistic communication. The position also allows for some neutrality on whether we are modelling mental processes or states or the behavioural output of public languages as in Lewis (1975). We might discover that the features of our models do track the internal mechanisms of individual language-users but this is by no means entailed by the modelling process.

Conceptualists might be reluctant to accept the first grade of involvement for a number of reasons. Firstly, as previously mentioned, the first grade does not entail any specific ontological position. It is perfectly compatible with our grammars modelling abstract objects themselves, theories (in the logical sense) or bee dances for that matter and not only the human mind/brain. Thus, it is a weaker claim than the Conceptualist position in this sense. I think that this is an

<sup>15</sup> In fact, adherence to a strong interpretation of *C3* can lead us astray in some cases. There is a school of thought which takes infinity or fixed cardinality not only to be a modelling choice, as in the previous section, but to be a feature of the particular mathematical model used in linguistic theory, namely Post canonical production systems. Thus, infinity is an artefact of the model. It is obvious that not all the artefacts of models should receive interpretation in the target system. Especially if productivity facts can be captured by alternative formalisms which do not posit the putative property. For a specific example, Sampson (2001) criticises Chomsky’s problematic “undue preoccupation with strings” in *The Logical Structure of Linguistic Theory* (1975). He points out that treating the syntax *via* derivations of strings and sets of strings is an unnecessary detour when phrase-structure grammars could be characterised with well-formedness conditions on trees directly. Furthermore, the derivational alternative forces certain untoward consequences.

Chomsky’s approach forces him to impose two quite arbitrary restrictions on phrase-structure rules, namely, that no rule may rewrite any symbol *A* as either the null string, or as a sequence including *A*. Both of these forbidden types of rule frequently seem appropriate in describing real language, and under the alternative view of phrase-structure grammars there is no objection to them (Sampson, 2001: 156).

In this case, the model has features that the real world does not. In many other cases, the target has features which outstrip the models. Trying to find a home for every feature of the model as a mechanism or constraint might turn out to be deeply problematic.

advantage of the view but it might escape an important aspect of the Chomskyan project, namely representationalism. I will return to this point in the section 9.

## 8 The Second Grade: Grammars as Mathematical Theories

So far we have been trying to characterise the nature of the linguistic enterprise according to the mathematical involvement of its grammars. However, some theorists such as Postal (2003, 2009) argue that there is a tension in attempting to reconciling (1) the physical biological aspect of natural language (or the empirical scientific status of linguistics) with (2) the formal aspect of its description. In 7., I argued that this can be done successfully. However, my view took no essential position on the ontology of natural language and thus could be compatible with something other than (1). For those who hold something like (1), the third grade of involvement (the topic of the next section) is the next natural step. However, on that grade, it is not clear how to meet Postal's challenge among other things (at least as the view currently stands). Nevertheless, there is still another option available to those interested in a coherent ontology for linguistics. This option takes the form of rejecting (1) outright and placing (2) at the forefront of the linguistic agenda. So the second grade of mathematical involvement for grammars places linguistics at the level of a formal science. Importantly, however, I hope to show that this position is not exhaustively captured by the linguistic Platonism of Postal or Katz and specifically does not necessarily entail its infinite ontology.

A grammar, on this grade, is viewed as a scientific theory but of a specific kind, namely a mathematical theory. The modelling picture of 7 (and Fig. 1) is thus truncated and the formal descriptions of the grammar, such as the proof systems along the lines of the Post canonical system of section 3, specify linguistic reality directly. Another way to put this is that the structural relation between the grammar and the target, as per strong  $C1$ , is identity. On this view, grammar construction involves intuiting or deducing aspects of an abstract linguistic reality in similar fashion to proof construction in logic or mathematics.

On this grade of involvement, in order to describe or explain certain (constitutive) properties of natural languages at the appropriate level of abstraction (types in lieu of tokens) such as recursion or infinity, mathematics is not only structurally necessary as in  $C1$  but also materially so as in  $C2$  since both the grammar and the target are equally abstract (or non-spatio-temporal).

My argument against this construal of the relationship between mathematics and linguistics takes two forms. On the one hand, I will argue for a methodological distinction between linguistic and mathematical theorising. On the other hand, I will argue that even if this latter objection can be overcome, the second grade of mathematical involvement still does not entail linguistic Platonism or an actual linguistic infinity, since interpreting linguistics as a formal science opens up myriad possibilities within the foundations of mathematics.

Methodological distinctness might seem obvious. It might be argued that it is clear that linguists do not make similar claims to those of mathematicians nor use similar methods to establish those claims, namely *a priori* methods. Linguistic grammars are concerned with natural languages, use empirical data and are thus scientific theories (or models) not formal ones. However, doing this might be begging the question against the Platonist or the adherent of a second grade of involvement (not identical to the Platonist). The Platonist claim is precisely that natural languages are abstract in the same sense as natural numbers are, linguistic claims are true of an objective (necessary) acasual reality and grammars are proof systems or mathematical theories describing properties and relations of this reality.

One reason for the methodological discrepancy between linguistics and mathematics could be that linguistics certainly seems to use mathematical tools in identifying the properties of its objects (as do many sciences) but it does not mathematically define the objects of its inquiry or rather use mathematical methods. As I have argued in section 7, linguists are interested in grammars conceived of as abstract objects but not *qua* abstract objects. In mathematics, once you stipulate or prove the consistency or necessity of an object, its existence follows (consider restricted comprehension or 'separation' in ZFC set theory).<sup>16</sup> In linguistics, simply finding a consistent set

<sup>16</sup> Of course, this is not always the case. The consistency of the Continuum hypothesis and its negation do not establish existence.

of rules is not enough. These rules have to model the structures of real-world languages or linguistic competence, i.e. contingent facts. It is not sufficient to identify a possible sequence of syntactic object, O(object)S(ubject)V(erb) for instance, without investigating whether or not this sequence holds for any actual languages.<sup>17</sup>

In terms of the infinity debate. Many linguists, following Chomsky, Hauser and Fitch (2002), take recursion or discrete infinity to be a universal property of natural language (as we saw in section 4.1). However, this posit too has been challenged on the empirical level by Everett (2005) who infamously claimed that recursion is missing from the syntax of Pirahã, a language spoken in Amazonas Brazil, and thus cannot be a universal property of natural language.

It seems that in linguistics, even the most theoretically charged debates cannot escape the empirical. Thus, the fields of mathematics and linguistics are methodologically distinct.

Of course, a theorist on this grade of involvement could accept all (or most) of this reasoning and still maintain that linguistics is a formal science of a slightly different order. Katz (1981) anticipates some objections similar to the ones I have raised above. He holds that there is a “single faculty of intuition” responsible for competences of different *a priori* areas of knowledge and the varying abstract objects under their respective remits. As for the connection between linguistics and mathematics, he has the following to say.

It may be said, for example, that the practice of the grammarian and the mathematician are dissimilar in that the working mathematician, unlike the working grammarian, does not spend large amounts of time soliciting and collecting intuitions. Conversely, the grammarian does not make extensive use of formal deductive procedures (Katz, 1981: 215).

For Katz, this is all just a matter of degree and the comparatively short history of linguistics. Comparing linguistics as it is today (or in the 80’s) to mathematics as it is today is like comparing logic in the time of Aristotle to contemporary mathematical logic. Eventually as the science progresses, we will rely on intuition gathering less frequently. As for the lack of deductive procedures, he argues this is misleading. If we are talking about the proofs within metatheory (about systems) such as soundness, completeness, incompleteness etc. then linguistics indeed does not involve too much mathematics of this kind (with the notable exception of the Vastness proof). If, however, we are talking about first-order theories or proofs then there is an analogue in linguistics. The derivations of our Post canonical systems or generative grammars are such devices and these are ubiquitous. Once again, he thinks that the metatheory will also come along as more formalisation (or rather mathematisation) occurs in the study of natural language (a prerequisite for metatheory in formal systems).

I think that this is an interesting idea, even if it is speculative at best. Certainly, the history of geometry has shown a progression from concern with physical spatial intuitions and with the rise of non-Euclidean geometry in the 19th century culminating in the Hilbert programme in the 20th, to an abstract science not essentially informed by real-world constraints. I am not sure, however, how to imagine a similar scenario with relation to natural language in which linguistic intuitions no longer play any definitive role in grammar construction.

What underlies the methodological distinctness claim in my view is another confusion in terms of the nature of grammars which relates back to George’s (1989) characterisation. In the previous section, on the previous grade of involvement, grammars as models or abstract objects were confused with psychogrammars and the physiogrammars which underlie them. On this grade of involvement, grammars are confused with the abstract objects themselves. As mentioned in section 7, we are not interested in grammars as abstract objects *qua* abstract objects but rather as abstract objects *qua* models of linguistic phenomena.

This is my case for the methodological separation of mathematics and linguistics. We can now move on to the second part of the argument, namely that even if we do accept the second grade of mathematical involvement for the grammars of natural language, this in itself does not entail linguistic Platonism or the view that languages are infinite abstract objects not extended in space or time.

The reason is that the second grade of involvement, the claim that linguistics is a formal science akin to logic or mathematics (but perhaps not identical), does not necessarily entail the existence

<sup>17</sup> Incidentally, this rare sequence of word order can be found in some languages along the Amazon basin.

of abstract mind-independent objects. The philosophy of mathematics offers many different approaches to the ontology and interpretation of mathematics including nominalism. It is not clear to me why a Platonistic linguistics is considered to be the default position for a coherent ontology or for an analogy with the formal sciences. In fact, there are a number of reasons for opting for an alternative picture.

For one thing, if we accept a Platonistic ontology for linguistic objects we face Benaceraff's famous dilemma. Benaceraff (1973) argued that there is a tension between the semantics and epistemology for any theory of mathematical truth. If we attempt to offer a standard account of its semantics (in terms of our best truth-conditional theory) then we have reference to abstract objects which moves us further away from a standard (causal) account of its epistemology. Suffice to say, that with Platonism comes an added epistemological burden. Not only is reference to abstract objects difficult to explain but knowledge of an acausal non-spatio-temporal realm beyond the physical is highly problematic.

This problem *inter alia* has prompted many philosophers of mathematics to opt for nominalistic accounts of mathematics, which do not posit abstract objects (Field 1980, Azzouni 2004). There are also structuralist accounts, modal (Hellman 1989) which only require *possibilia*, eliminative, which similarly to nominalism, do not posit mathematical objects, non-eliminative, which do but in a “places-as-objects” within structures notion of object (Resnik 1997, Shapiro 1997). And then there are varieties of Platonism, hard-line (Gödel 1944) and more light-weight versions (Linsky and Zalta 1995). This list is not exhaustive by any means.

Therefore if linguistics is a formal science, as the second grade of involvement assumes it is, then there is no principled reason to opt for naive Platonism or the view that languages are actually infinite. As mentioned at the start of this section, the second grade of mathematical involvement for grammars, in which grammars are mathematical theories, does not entail any specific ontology for linguistic objects (or their necessary existence). Furthermore, we are still burdened with accounting for the empirical side of linguistic research and how abstract languages relate to everyday spoken languages.

## 9 The Third Grade: Representational Realism

On the last grade of mathematical involvement, the linguistic rules and the various posits (such as PRO, traces or copies depending on your generative persuasion etc.) of grammars are argued to have greater significance to the physical system represented than indirect representation or modelling of some sort. Linguists sometimes speak of the rules of a grammar being “internally represented” on this view. Chomsky (1986a: 243) describes a speaker “equipped with a grammar” as someone who “internalizes a system of rules”. On the basis of such an internalisation of a rule system  $R$ , the speaker's linguistic behaviour can be explained or predicted by the structure of  $R$  (i.e.  $C1$ ). To glean how this level of involvement is starkly different from the positions described in the previous sections, Pylyshyn is particularly illuminating.

[D]espite the uncertainties, none of us doubted that what was at stake in all such claims was nothing less than an empirical hypothesis about *how things really were inside the head of a human cognizer*. We knew that we were not speaking metaphorically nor were we in some abstract way describing the form of the data. (1991, 232)

We can see from this quote that the third grade of involvement, or “representational realism” as Pylyshyn calls it, is a much stronger claim than first grade (or the second). Grammars, on this view, are really like scientific theories and their posits are of the same nature as atoms and quarks are in physical theory, that is actual features of the physical system. If a grammar posits a mechanism of wh-movement or a recursive rule like adjectival modification (AP bar in X-bar theory), then these are features of a language user's actual brain-state when parsing these structures and languages are discrete infinite. This is in line with some form of  $C3$ . The claim that linguistics will eventually be subsumed by biology or neuroscience seems less vague on a this grade of involvement.

It is important to pause here to consider the difference between theories and models again. In the previous sections, I took models to be indirect representations of a target system. The mathematical structures involved in the building of models or grammars did not necessarily reflect any structural features of the target system in kind. For instance, positing recursive elements in the grammar

only modelled iterative constructions in natural language indirectly, therefore I held that natural language did not need to be committed to recursion or infinity. In fact, the move within generative grammar from the recursive structures of Post-canonical systems to the single set-theoretic merge operation is evidence of the fact that recursion is an aspect of the models or grammars which can change without the target system changing (presumably language didn't change when linguists moved from the Extended Standard Theory to Minimalism). Scientific theories, on the other hand, represent the target system or natural world *directly*. In other words, scientific theories tell us what there is in the world. If grammars are scientific theories, then the structures and posits within them are claimed to be actual features of natural language (which are brainstates on this view). Thus recursion and infinity are aspects of natural language competence and indeed Conceptualists often speak this way (as in the many examples shown in section 3). Such an interpretation of the role of grammars naturally lends itself to analogies with mathematical cognition which presumably involves similar structures. For instance, if merge is an evolutionary mutation, it cannot merely be a formal aspect of a model. In order for this claim to even begin to make sense, it has to be assumed to be a claim about actual features of linguistic competence or reality posited by the grammar. Thus grammars preserve the very structures of linguistic reality and the third grade of involvement is committed to a strong version *C1* as well as *C3* (we will see how it is committed to *C2* in section 9.1 below).

Evans (1981) is an interesting test case for representational realism. In response to a criticism (initially levelled at Chomsky by Quine (1972)) that weakly equivalent grammars (mentioned in section 7.2, with different internal structures but equivalent behavioural output) pose a problem for representational realism, Evans offers a dispositional account of tacit (semantic) knowledge. I will only focus on the structure of the argument with relation to infinite generalisations here.

Evans argues that given two weakly equivalent systems, one containing axioms or primitives and another containing composition rules and constituents, the former unlike the latter will be unable to predict the human speaker's ability to understand previously unheard or novel sentences. The second system might be considered generative but it is not necessarily infinite (in fact Evans cautions against confusing infinity and creativity).

The important aspect of Evans' dispositional account is that it has an empirical, testable component. Two weakly equivalent grammars create distinct dispositions, ones which have distinct explanatory power. In other words, there are ways in which we can divine which grammars are "internally represented" by an agent on the basis of certain dispositions elicited by the structures or posits of the grammar. Nevertheless, the property of discrete infinity is unlikely to be read off grammars even dispositionally. There might be no upper bound on the number of sentences (needed for creativity) but nothing indicating that language is a set containing a nondenumerable infinity of objects (see Pullum and Scholz (2010) for discussion). Accounts such as this one were part of the motivation behind early psycholinguistics. However, the failure of hypotheses such as the derivational theory of complexity or the claim that "the complexity of a sentence is measured by the number of grammatical rules employed in its derivation" (Fodor, Bever and Garrett, 1974: 319), showed that this project was not the sought-after evidence for representational realism or

In fact, as Devitt (2006) notes, even when there is some positive evidence for posits of a grammar such as constituent structure (like in the famous "click location" experiment discussed in Fodor *et al.* 1974), this offers no proof of the psychological reality of the rules unless we presuppose the truth of the second grade involvement (representational realism or Devitt's Representational Thesis (RT)). These experiments could show no more than that competence respects certain structural posits of our grammars (Devitt's minimal position M). Fodor *et al.* themselves could not find any place for internalised grammar rules in actual parsing, "[t]here exists no suggestions about how a generative grammar might be concretely employed as a sentence recognizer in a psychologically plausible model" (1974: 75).

A further interesting fact about psycholinguistic or physiological evidence is that it generally plays no real role in grammar construction. The primary data for the grammar on the second grade tend to be native speaker judgements (there are notable exceptions). Since Fodor *et al.* (1974), there has been a lot of research conducted on psychological and biological effects and interactions in language production and comprehension. For instance, Cowart (1989b) discovered that familial handedness can have an effect on grammaticality judgements involving subadjacency. "[R]ight-handed speakers without left-handed relatives are more sensitive to subadjacency violations (rate them as less



grammatical) than right-handers that have lefthanded relatives” (Keller, 1998: 7). If linguists were indeed in the business of developing grammars *qua* scientific theories of I-languages or brain-states, then such data would surely be more relevant to the task. Yet this and other types of physiological evidence seems to have no place in the grammars of the third grade. This fact is related to an argument presented in Soames (1984) to the effect that linguistics and psychology are empirically divergent or they require different sets of evidence for confirmation of their theories.

The main problem is that representational realism (strong *C3*), and the third grade of mathematical involvement generally, does not find a natural home for infinity claims and the like. If a grammar is embodied in a finite system such as the human brain, then claiming that discrete infinity is somehow also part of that system is problematic. Actual infinities belong in the realm of the formal not the natural (see Peregrin 2000 for one way of viewing the distinction). As we have seen, this forces linguists to claim the extra-biological nature of natural language.

Thus, even if we do accept this grade of involvement and its representational realism, we are still left in some confusion as to the biological underpinnings of the movement. This is the Postal (2003, 2009) problem of how the features of the grammars which involve sets, sentence types and discrete infinity are supposed to be captured by a physical biological system like a brain-state. In section 7., I offered a coherent picture of this relationship in terms of models and their idealisations. I also suggested that Chomskyans would be reluctant to accept this picture. The reason for this reluctance is that they usually have a stronger structural connection in mind, a combination of *C1*, *C2* and *C3*, as evinced by their adherence to representational realism, which takes the rules and posits of the grammars to be actual structures of the mind/brain of language users used during processing.

### 9.1 Core Grammar, *C2* and FLN

Lastly, the purported adherence to *C2* needs to be addressed. In the section 7, we saw how assuming material preservation or substance equivalence between models and their targets is not necessary. In the second grade of involvement, we saw how such a claim is problematic. Idealisations, false-models, even measurement models are all counterexamples to the claim that it is. Yet on the third grade of involvement, there is a position which seems to incorporate an aspect of *C2*. The idea is that linguistic competence and performance are separated by “core” and “peripheral” phenomena respectively.

(Chomsky 1981) distinguishes between two parts of the grammar itself, the core and the periphery. The core rules represent the deep regularities of language. The periphery represents marked exceptions, such as irregular verbs, for which there are no deep regularities.

The idea is that linguistic theory should focus on the core phenomena which are susceptible to more or less precise characterisation. While the periphery phenomena, laden with historical accident, dysfluency and irregularity, are to be ignored in a serious science of language. As Pullum (1983) describes the posit,

Chomsky does not assume that the grammars actually internalized by humans are (necessarily) defined as possible by universal grammar (UG). A basic “core” defined by UG is involved, but there is also a “marked periphery” of additional special constructions and exceptional cases that are learned on the basis of experience and not shaped in the same way by UG (448).

This proposal relies on the competence-performance distinction initially presented in Chomsky (1965). Competence is constituted by a generative grammar and indeed represented in the mind of the speaker. However, what psycholinguistic experiments are sometimes tracking are the heuristic and stochastic devices responsible for immediate parsing and performance needs (algorithms in George (1989)) hence the empirical divergence (Soames 1984). One way to think of the relationship between the two systems is that the performance system checks itself on the competence grammar as it processes language in real-time. Think of the competence grammar as the generative grammar box situated somewhere in the mind, around it are the various quick-fire linguistic responses to external stimuli. The box is responsible for checking whether or not a given input is well-formed. When we produce and interpret sentences on the fly, we generally do not rely on this box. But upon reflection we often consult it to “check” whether or not a given string of words is in fact grammatical (hence the divergence between grammaticality and acceptability judgements). To lend

some credence to this idea, at the beginning of *Aspects*, Chomsky seemed to base his idealisation of the true subject matter of linguistic theory on precisely this distinction (i.e. an idealised speaker-hearer in a homogeneous speech community with perfect knowledge of her language and immune from the vagaries of memory limitations and the like). As Stabler notes,

The linguistic idealization is apparently grounded in the empirical assumption that the mechanisms responsible for determining how phrases are formed in human languages are relatively independent of those involved in determining memory limitations, mistakes, attention shifts, and so on (2011, 70).

This is all very well. But the concept of a “core grammar” has never been precisely laid out in the literature nor has its separation from peripheral mechanisms. As Pullum notes, “it is not clear whether the word “grammar” should be replaced by “core grammar” at the appropriate points. Nor is it clear to me what difference it would make” (1983: 449). If we are to follow a traditional competence-performance divide, we run into problems. For one thing, it is not clear that the line between ideal competence and actual performance can be drawn as sharply as it is suggested here. Various aspects of performance have been shown to be highly systematic and context has been argued to have a significant effect on grammaticality or acceptability judgements (see Jackendoff 2002, Cann *et al.* 2012, Baggio *et al.* 2012). Furthermore, constructional approaches such as Goldberg (2003) and Culicover (2011) aim to show that restrictions which govern the so-called “core” apply equally on the “periphery”.

We see in Chomsky, Hauser and Fitch (2002) a more contemporary elucidation of the erstwhile core grammar postulate. In this work, the authors distinguish between the faculty of language narrowly construed (FLN), the locus of recursion and discrete infinity, and the faculty of language broadly construed (FLB), the supersystem involving various other cognitive processes at the periphery.

FLN is the abstract linguistic computational system alone, independent of the other systems with which it interacts and interfaces...All approaches agree that a core property of FLN is recursion....This capacity of FLN yields discrete infinity (a property that also characterizes the natural numbers) (Chomsky, Hauser and Fitch, 2002: 1571)

Again, the idea is that discrete infinity is an actual part of the uniquely human language faculty. The major problem for the third grade of mathematical involvement is that in the absence of any evidence of the psychological or neurophysiological correspondence required for such an analysis, i.e. how discrete infinity is realised in the brain, or a precise notion of “core” mechanisms, the success of its grammars as scientific theories is unclear (of course they could be indirectly related as per grade one). I tend to side with Higginbotham (1991: 559) in stating that at the current level of scientific knowledge in linguistics, the second grade of mathematical involvement of grammars is not *indefensible* but rather *inarticulate*, taking strong *C1*, *C2* and *C3* along with it. I go a step further, however, in claiming that it is therefore an unsound methodological position in light of better options, such as the first grade of involvement. I think that Katz (1981), Postal (2003, 2009) proceed in a similar fashion. The difference is that they offer something along the lines of second grade of mathematical involvement for grammars as a more sound footing for the foundations of linguistics and infinity claims. As shown previously, this latter possibility is no more satisfying on methodological but comes with an increased epistemological burden.

## 10 Conclusion

There are many advantages to carving up the positions on the foundations of linguistics and infinity claims in linguistics in the way that I have attempted to do in this paper. By appreciating the role of grammars on each grade of mathematical involvement, we can divorce certain methodological concerns from ontological ones. We can also restate the Conceptualist view on less “incoherent” or rather “inarticulate” grounds and resituate the debate outside the scope of metaphysical complications. I further argued that by understanding grammars as formal scientific models, we can resolve the objections posed by Katz, Postal and others of the linguistic Platonist persuasion as well as avoid talk of abstract objects in themselves. The grades also exposed and confronted the mathematisation of natural language by showing that with each grade additional claims (*C1-C3*) as to the significance of the mathematical apparatus were imposed.

In this paper, I have attempted to resituate (and hopefully reenergise) the debate on the ontological foundations of linguistic theory as well as the role of infinite generalisations within the field. I have placed the role of grammars at the center of my three grades of mathematical involvement for linguistics by describing them in terms of three methodological commitments, *C1*, *C2* and *C3*. Furthermore, I argued that the first grade of involvement offers the linguist the path of least resistance, drawing from insights in computational linguistics and scientific modelling. I argued that the second and third grades are more problematic but certainly not beyond redemption or merit. My hope is that, at the end of the day, whichever path a linguist chooses to take for the interpretation of her field, or infinite generalisations within it, will be marked more clearly by appreciating or at least considering some of the arguments presented above.

**Acknowledgements** I would like to thank Josh Dever, Ephraim Glick, Patrick Greenough, Geoff Pullum, Kate Stanton, Zoltán Szabó, Bernhard Weiss and two anonymous referees for their insightful comments on various drafts of this paper. I would also like to thank audiences at the Arché Research Centre and the joint Semantics Seminar at the linguistics and philosophy departments at Yale University for their excellent comments and suggestions on this research.

## References

- Azzouni, J. 2004: *Deflating Existential Consequence: A Case for Nominalism*. New York: Oxford University Press.
- Baggio, G., van Lambalgen, M., and Hagoort, P. 2012: 'Language, Linguistics and Cognition'. In Kempson, R., Fernando, T., and Asher, N. (eds). *Philosophy of Linguistics*. Elsevier B.V. North Holland.
- Bar-Hillel, Y. 1953. 'On Recursive Definitions in Empirical Science'. *11th Natural Congress of Philosophy* 5: 160-165.
- Benacerraf, P. 1973: 'Mathematical Truth'. *The Journal of Philosophy*, Vol. 70, No. 19. pp 661-679.
- Bloomfield, L. 1936: 'Language or Ideas?' *Language*, vol. 12, pp 89-95.
- Boeckx, C. 2015: 'Beyond Humboldt's problem: reflections on biolinguistics and its relation to generative grammar'. *Language Sciences*, Vol. 50, pp 127-132.
- Bromberger, S. 1989: 'Types and Tokens in Linguistics'. In *Reflections on Chomsky* George, A. (ed.), pp 58-90. Basil Blackwell.
- Cann, R., Kempson, R., and Wedgwood, D. 2012: 'Representationalism and Linguistic Knowledge'. In Kempson, R., Fernando, T., and Asher, N. (eds). *Philosophy of Linguistics*. Elsevier B.V. North Holland.
- Chomsky, N. 1956b: 'Three models for the description of language'. *IRE Transactions on Information Theory* IT-2: 113-123.
- Chomsky, N. 1957: *Syntactic Structures*. Mouton Press: the Hague.
- Chomsky, N. 1959: 'On certain formal properties of grammars'. *Information and Control* 2.137-167.
- Chomsky, N. 1965: *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Massachusetts.
- Chomsky, N. 1972: *Studies on Semantics in Generative Grammar*. The Hague: Mouton.
- Chomsky, N. 1986a: *Knowledge of Language: Its Nature, Origin, and Use*. New York Praeger.
- Chomsky, N. 1995: *The Minimalist Program*. Cambridge, Mass.: MIT Press. Reprinted 20th Anniversary Addition, 2015.
- Chomsky, N. 2000b: *The Architecture of Language*. New Delhi: Oxford University Press.
- Chomsky, N., Hauser, M., and Fitch, W. 2002: 'The Faculty of Language: What is it, who has it, and how did it evolve?' *Science*, Vol. 298. pp 1569-1579.
- Chomsky, N. 2005: 'On Phases'. MIT MS.
- Chomsky, N. 2006: *Language and Mind*. Cambridge, UK: Cambridge University Press.
- Chomsky, N. 2010: 'Some simple evo-devo theses: how true might they be for language?' In R.K. Larson, H. Yamakido, and V. Deprez (eds.), *Evolution of Human Language: Biolinguistic Perspectives*. Cambridge University Press, pp 45-62.
- Chomsky, N. 2012: *The Science of Language: Interviews with James McGilvray*. Cambridge University Press, Cambridge.
- Chomsky, N and Miller, G. 1963. 'Introduction to the Formal Analysis of Natural Languages'. In *The Handbook of Mathematical Psychology*, Volume II, edited by Duncan Luce, R., Bush, R., and Galanter, E., John Wiley & Sons.
- Cowart, W. 1989b: 'Notes on the Biology of Syntactic Processing'. *Journal of Psycholinguistic Research* 18(1): pp 89 103.
- Culicover, P. 2011: 'Core and periphery'. In *The Cambridge Encyclopedia of the Language Sciences*, Hogan, P. (ed.), 227-230. Cambridge: Cambridge University Press
- Davies, M. 1987: 'Tacit Knowledge and Semantic Theory: Can a Five Per Cent Difference Matter?' *Mind*, Vol. 96, No. 384, pp. 441-462.
- Devitt, M. 2006: *Ignorance of Language*. Oxford University Press.
- Epstein, S., and Hornstein, N. 2005: Letter on 'The future of language'. *Language*, 81:36.
- Evans, G. 1981: 'Semantic theory and Tacit Knowledge'. in *Collected Papers* 1996, pp 322-342. Clarendon Press.
- Everett, D. 2005: 'Cultural constraints on grammar and cognition in Pirahã: another look at the design features of human language'. *Current Anthropology*, 46, 621-634.
- Field, H. 1980: *Science without Numbers: A Defense of Nominalism*. Princeton, N.J.: Princeton University Press.
- Field, H. 1985: 'Comments and Criticism: On Conservativeness and Incompleteness'. *Journal of Philosophy*, Vol. 82, No.5, pp 239-260.

- Fodor, J., Bever, T., and Garrett, M. 1974: *The Psychology of Language*. McGraw-Hill Publishers.
- Frigg, R. 2010b: 'Models and Fiction'. *Synthese*, 172(2): 251–268.
- George, A. 1989: 'How not to become confused about linguistics'. In *Reflections on Chomsky*, pp 90–110. Basil Blackwell.
- George, A. 1996: 'Katz Astray'. *Mind and Language*, Vol. 11. No. 3, pp 295–305.
- Giere, R. 1988: *Explaining Science: A Cognitive Approach*. Chicago: Chicago University Press.
- Gödel, K. 1944: 'Russell's Mathematical Logic'. In Benacerraf, P., and Putnam, H. (1983) (eds.), *Philosophy of Mathematics: Selected Readings*. Cambridge University Press.
- Godfrey-Smith, P. 2006: 'The Strategy of Model-based Science'. *Biology and Philosophy*, 21: 725–740.
- Goldberg, A. 2003: 'Constructions: a new theoretical approach to language'. *TRENDS in Cognitive Sciences* Vol.7 No.5.
- Hellman, G. 1989: *Mathematics without Numbers: Towards a Modal-Structural Interpretation*. Oxford: Clarendon Press.
- Higginbotham, J. 1991: 'Remarks on the Metaphysics of Linguistics'. *Linguistics and Philosophy*, Vol. 14, No.5 pp 555–566.
- Hilbert, D. [1899]: *Grundlagen der Geometrie*. Leipzig, Teubner; *Foundations of Geometry*, tr. by Townsend, E, La Salle, Illinois, Open Court, 1959.
- Hinzen, W., and Uriagereka, J. 2006: 'On the Metaphysics of Linguistics'. In *Erkenntnis*, Vol. 65, No.1, pp 71–96.
- Jackendoff, R. 2002: *The Foundations of Language*. Oxford University Press.
- Katz, J. 1981: *Language and Other Abstract Objects*. Rowman and Littlefield, Totowa, New Jersey.
- Katz, J. 1985: 'An outline of platonist grammar'. In Katz, J (ed) *The Philosophy of Linguistics*. Oxford University Press.
- Katz, J. and Postal, P. 1991: 'Realism vs. Conceptualism in Linguistics'. *Linguistics and Philosophy* Vol. 14, No.5 pp 515–554.
- Kornai, A. 2014: 'Resolving the Infinitude Controversy'. *Journal of Logic, Language and Information*, Vol. 23, Issue 4, pp 481–492.
- Langendoen, T. 2003: 'Merge'. *Formal Approaches to Function in Grammar: In Honor of Eloise Jelinek*, ed. by Andrew Carnie, Heidi Harley & Mary Willie, 307–318. Amsterdam: John Benjamins.
- Langendoen, T. 2008: 'Coordinate grammar'. *Language* 84.691–709.
- Langendoen, T. 2010: 'Just how big are natural languages?' In van der Hulst, H. (ed.), *Recursion and Human Language*, pp 139–47. De Gruyter Mouton.
- Langendoen, T., and Postal, P. 1984: *The Vastness of Natural Languages*. Blackwell Publishers.
- Lasnik, H. 2000: 'Syntactic Structures'. In *Revisited: Contemporary Lectures on Classic Transformational Theory*. Cambridge, MA: MIT Press.
- Lewis, D. 1975: 'Languages and Language'. In Matinich, P. (ed) *The Philosophy of Language*. Oxford University Press.
- Lewis, D. 1980: 'Index, Context, and Content'. In Kanger, S., and Öhman (eds.) *Philosophy and Grammar*, 79–100.
- Linsky, B., and Zalta, E. 1995: 'Naturalized Platonism vs. Platonized Naturalism'. *Journal of Philosophy* xcii/10: 525–555.
- Lobina, D. 2014: 'When linguists talk mathematical logic'. *Frontiers in Psychology*, Vol. 5. 382, pp 1–3.
- Ludlow, P. 2011. *The Philosophy of Generative Linguistics*. Oxford University Press.
- Mäki, U. 2011: 'Models and the Locus of their Truth.' In *Synthese* 180, pp 47–63.
- Marr, D. 1982: *Vision*. W.H. Freeman and Company: New York.
- McDonald, F. 2009: 'Linguistics, Psychology, and the Ontology of Language'. *Croatian Journal of Philosophy* Vol. IX, No. 27.
- Montague, R. 1970a: 'English as a Formal Language'. In: Bruno Visentini (ed.): *Linguaggi nella società e nella tecnica*. Reprinted in R. Montague, *Formal Philosophy*. (ed.) R. Thomason. New Haven, CT: Yale University Press, 1974, pp 189–223.
- Montague, R. 1970b: 'Universal Grammar'. *Theoria* 36:37398. Reprinted in R. Montague, *Formal Philosophy*. (ed.) R. Thomason. New Haven, CT: Yale University Press, 1974, pp. 222–46.
- Morrison, M. 2015: *Reconstructing Reality: Models, Mathematics, and Simulations*. Oxford University Press.
- Nefdt, R. 2016a: 'Linguistic Modelling and the Scientific Enterprise'. *Language Sciences*, Vol. 54, pp 43–57.
- Nefdt, R. 2016b: 'Scientific Modelling in Generative Grammar and the Dynamic Turn in Syntax'. *Linguistics and Philosophy*, Vol. 39(5): 357–394.
- Newmeyer, F. 1996: *Generative Linguistics: A Historical Perspective*. London: Routledge.
- Peregrin, J. 1995: 'Structural linguistics and Formal Semantics'. *Travaux du Cercle Linguistique de Prague I*. Benjamins, Amsterdam.
- Peregrin, J. 2000: 'The 'Natural' and the 'Formal''. *Journal of Philosophical Logic* Vol. 29, No. 1, pp. 75–101.
- Postal, P. 2003: 'Remarks on the foundations of linguistics'. *The Philosophical Forum*. Vol. XXXIV, Nos. 3 & 4.
- Postal, P. 2009: 'The Incoherence of Chomsky's 'Bi-linguistic' Ontology'. *Biolinguistics* 3.1: 104–123.
- Pullum, G. 1983: 'How Many Possible Human Languages Are There?'. *Linguistic Inquiry*, Vol. 14, No. 3, pp. 447–467.
- Pullum, G. 2011: 'On the mathematical foundations of syntactic structures'. *Journal of logic, language and information*, 20(3), pp 277–296.
- Pullum, G. 2013: 'The Central Question in Comparative Syntactic Metatheory'. *Mind & Language*, Vol. 28, No. 4, pp. 492–521.
- Pullum, G. and Scholz, C. 2001: 'On the distinction between model-theoretic and generative-enumerative syntactic framework.' In *Logical Aspects of Computational Linguistics: 4th International Conference*. Springer Press.
- Pullum, G. and Scholz, B. 2010: 'Recursion and the infinitude claim'. In van der Hulst, H. (ed.), *Recursion in Human Language (Studies in Generative Grammar 104)*, pp 113–138. Berlin: Mouton de Gruyter.

- Pylyshyn, Z. 1991: 'Rules and representations: Chomsky and representational realism.' In Kasher, A. (ed.), *The Chomskyan turn*. Cambridge, MA : Blackwell, pp 231-51.
- Quine, W.V. 1972: 'Methodological reflections on current linguistic theory'. In Davidson, D., and Harman, G. (eds), *Semantics of Natural Language*. Dordrecht: D. Reidel, pp 442-454.
- Quine, W.V. 1976: *Ways of Paradox and Other Essays*. Harvard University Press.
- Resnik, M. 1997: *Mathematics as a Science of Patterns*. Clarendon Press: Oxford.
- Sag, I., Wasow, T., and Bender, E. 2003: *Syntactic Theory: A Formal Introduction. Second Edition*. CSLI Publications.
- Sampson, G. 2001: *Empirical Linguistics*. Continuum Press.
- Savitch, W. 1993: 'Why it might pay to assume that languages are infinite'. In *Annals of Mathematics and Artificial Intelligence* 8, 17-25.
- Shapiro, S. 1983: 'Conservativeness and Incompleteness'. *The Journal of Philosophy*, Vol. 80, No. 9, pp. 521-531.
- Shapiro, S. 1997: *Philosophy of Mathematics: Structure and Ontology*. Oxford University Press.
- Shapiro, S. 2005: 'Categories, Structures, and the Frege-Hilbert Controversy: The Status of Metamathematics'. *Philosophia Mathematica* III, 13, pp 61-77.
- Shieber, S. 1985: 'Evidence against the context-freeness of natural language'. In *The Formal complexity of natural language*. Springer Netherlands. 320-334.
- Soames, S. 1984: 'Linguistics and Psychology'. *Linguistics and Philosophy* 7, 155-179.
- Stabler, E. 1997: 'Derivational Minimalism'. In Retoré, C. (ed.), *Logical Aspects of Computational Linguistics*, 68-95. Springer.
- Stabler, E. 2011: 'Meta-meta-linguistics'. *Theoretical Linguistics*. Vol. 37, no. 1/2: 69-78.
- Strevens, M. 2007: 'Why Explanations Lie: Idealization in Explanation.' Unpublished manuscript, Department of Philosophy, New York University.
- Suppes, P. 1960: 'A Comparison of the Meaning and use of Models in Mathematics and the Empirical Sciences'. *Synthese*, 12, pp. 287-300.
- Thomason, R. 1974: *Formal Philosophy. Selected Papers by Richard Montague*. New Haven.
- Thomson-Jones, M. 2012: 'Modelling without Mathematics'. *Philosophy of Science*, Vol. 79, pp. 761-772.
- Tiede, H., and Stout, L. 2010: 'Recursion, Infinity and Modeling'. In *Recursion and Human Language*. (ed) van der Hulst, H. Berlin/New York: Mouton de Gruyter, pp. 147-158.
- Tomalin, M. 2006: *Linguistics and the Formal Sciences*. Cambridge University Press.
- Tomalin, M. 2007: 'Reconsidering Recursion in Syntactic Theory'. *Lingua* 117, pp 1784-1800.
- Weisberg, M. 2007a: 'Who is a Modeler?' *British Journal of the Philosophy of Science*, Vol. 58: 207-233.
- Weisberg, M. 2007b: 'Three Kinds of Idealization'. *The Journal of Philosophy*, Vol. 104, No. 12: 639-659.
- Weisberg, M. 2013. *Simulation and Similarity: Using Models to Understand the World*. Oxford University Press.
- Wright, C. 1989: 'Wittgenstein's Rule-following Considerations and the Central Project of Theoretical Linguistics'. In *Reflections on Chomsky*, pp 231-264. Basil Blackwell.
- Yablo, S. 2013: 'Explanation, Extrapolation, and Existence'. *Mind*, 121(484), 1007-1029.
- Yalcin, S. 2014: 'Semantics and Metasemantics in the Context of Generative Grammar'. In *Metasemantics: New Essays on the Foundations of Meaning* (eds) Burgess, A and Sherman, B. Oxford University of Press.
- Yang, C. 2006: *The Infinite Gift*. New York: Scribner.