Formalism, Grammatical Rules, and Normativity

Geoffrey K. Pullum

School of Philosophy, Psychology and Language Sciences, University of Edinburgh

1 Introduction

Formalism as a school of thought in logic and mathematics does have some connections to what linguists sometimes call formal linguistics and sometimes call generative linguistics. Both have some connections to the notion of a formalized theory. After clarifying the issues somewhat, I turn to the question of whether rules of grammar (formalized or not) have normative force — whether they are concerned with what is right or wrong in some sense, rather than with the way things are in the world. I take the view that the subject matter of grammar is indeed normative, though that is not at all the same as having a prescriptive motivation, and I hold that it is perfectly sensible to attempt to render grammatical rules precise through formalization, but that the generative systems nearly always assumed by those who call themselves formal linguists probably represent the wrong mode of formalization. I then attempt to argue that the normativity of grammar is not in conflict with the idea that linguistics is an empirical science, and point to a framework that shows promise of reconcile those two apparently counterposed ideas.

2 Formalism

The school of thought known as 'formalism' in logic and mathematics takes these disciplines to be concerned solely with procedures for manipulating strings over a finite inventory of meaningless symbols. Put like that, it sounds pointless: surely logic and mathematics were supposed to be **about** something, not just meaningless games? But though this be madness, yet there is method in it. The point is to ensure that proofs are constructed in a way that is scoured clean of any hint of meaning or truth. Then it can be shown that proof truly accomplishes something.

Ideally we want everything logically provable from true premises to turn out to be true, and everything that is logically true to be provable. To show that this has been achieved, without circularity, we need first a method of proof that pays no regard to meaning or truth, and second a way of demonstrating that it proves all and only the logical truths. In technical terms, we want a consistent and complete method of proof.

A crucial contribution to the formalist program in logic, just over a century old, is presented in section III of Chapter 6 of *A Survey of Symbolic Logic* (first edition, 1918)[1] (1918) by Clarence Irving Lewis (1883–1964). Lewis clearly saw that the foundational work *Principia Mathematica* (1910–1913, henceforth *PM*) by Alfred North Whitehead (1861–1947) and Bertrand Russell (1872–1970), despite its renown as a centrally important attempt at reducing mathematics to logic, had not separated syntax from semantics in the way that was needed. The

---

[1]The edition is crucial. When the second edition by Dover was authorized, Lewis stipulated that Chapters 5 and 6 of the book had to be omitted; he thought that whatever value the book had did not depend on those two somewhat more speculative and heterodox chapters. Clearly he was being too modest.

distinction between axioms and inference rules had not yet emerged (Whitehead and Russell subsumed them both under "primitive propositions"), and in consequence Modus Ponens was framed in a semantically contaminated way that from the formalist viewpoint is shockingly confused. It says "Anything implied by a true elementary proposition is true" and, where x is a real variable, "When φx can be asserted . . . and $\varphi x \supset \psi x$ can be asserted . . . then ψx can be asserted."

This means we cannot legitimately make use of the claim that p implies q unless we can be sure that p is true. But p might of course be a proposition we are by no means sure of. Using logic to see what follows from a false assumption like this is an important technique of discovery, but Whitehead and Russell's statement of Modus Ponens doesn't allow that. Lewis understood that if we want to be sure that our symbolic reasoning is trustworthy, we must have a purely syntactical method of deriving strings, one that does not in any way depend on meaning, and we must then show that the strings it derives are the right ones — the ones that are true iff the initially adopted premises are true.

The program Lewis had informally urged was taken up in real earnest by a PhD student in the Department of Mathematics at Columbia University around the time when Lewis's book was published. Emil Leon Post (1897–1954) aimed to set Whitehead and Russell's logic on a really sound foundation of pure mathematics. His proposed a plan for a doctoral dissertation, under the supervision of the philosopher and mathematician Cassius Jackson Keyser (1862–1947), was an ambitious one. He aimed to construct

(I) a way of testing a formula of the propositional calculus used in *PM* to see whether it was a tautology (i.e., a logical truth);

(II) a system for deriving certain formulas (intuitively, the theorems) from given formulas (intuitively, the axioms) that was totally independent of semantic or logical categories like 'constant' or 'variable' or 'connective', working on strings of symbols without reference to their potential meaning; and

(III) a proof that the set of tautologies as determined by (I) coincided with the set of derivable strings defined by (II).

For the limited portion of Whitehead and Russell's logic that he tackled, Post actually achieved that goal (see the abridged version of his PhD dissertation published as Post (1921). He planned to go on and complete the job of dealing with the whole of *PM*'s logic, including the quantificational reasoning, in a postdoctoral year at Princeton, where he had obtained a Procter fellowship. In pursuit of that goal, he generalized his syntactical proof system further, and created a type of formal system that would revolutionize theoretical linguistics nearly 40 years later.

The system Post worked out in 1921 was not described in a publication until 1943, but there were reasons for that. The reasons had to do with a long battle against severe manic-depressive mental illness (now usually called bipolar disorder). Suffice it to say that the system was general enough that it could express any reasonably envisageable set of rules for deriving strings from other strings. Indeed, Post rapidly came to believe that it fully captured the intuitive notion 'set for which it is possible in principle to imagine how one might systematically enumerate the membership'. Today such a set is usually referred to as **recursively enumerable (r.e.)**.

Formalizing the inference rule Modus Ponens is a very simple application of Post's system. Assume a symbol inventory containing '$\supset$', ')', '(', and letters like p and q, and P_i are to be interpreted as indexed variables over unbounded strings. Assume you have already been given some strings that don't need to be proved (intuitively those are axioms), and you are trying to

build a list of other strings using them (those are the theorems). Then Modus Ponens says that if you can find a string on your list that has a '(' followed by some other stuff which we'll call P_1, followed by the symbol '⊃', followed by some further stuff which we'll call P_2, followed by a ')', and the stuff that we called P_1, on its own, is also on the list, then the stuff that we called P_2 can be added to the list.

The way Post put it was that a string of the form (P_1 ⊃ P_2) together with a string of the form P_1 **produce** a string of the form P_2.

Modus Ponens is an extremely simple application of the idea of building strings systematically on the basis of already obtained strings (though it was the crucial for Post's PhD dissertation project). But in 1920–1921 as a Procter fellow at Princeton, Post started working on doing the rest of Whitehead and Russell's logic — the reasoning that involved quantifiers — and in connection with that he worked out a radical generalization of the notion of a rule of inference. There could be any finite number of 'premise' lines, and both the premises and the conclusion could be of any finite length and contain any number of specified symbol strings and/or the variables over strings (P_1, P_2, etc.). He presented his generalized superschema in a truly bewildering tableau; as restated in a slightly modified form due to Davis (1982), which (believe it or not) makes things slightly clearer, it looks like this (where the g_i and h_i stand for specific strings of symbols that would be given in the production, and the down arrow '⇓' means 'produce' in Post's sense):

$$
\begin{array}{ccccccc}
g_{1,0} & P_{1,1} & g_{1,1} & P_{1,2} & ... & P_{1,n_1} & g_{1,n_1} \\
g_{2,0} & P_{2,1} & g_{2,1} & P_{2,2} & ... & P_{2,n_2} & g_{2,n_2} \\
& & \vdots & & ... & & \vdots \\
g_{k,0} & P_{k,1} & g_{k,1} & P_{k,2} & ... & P_{k,n_k} & g_{k,n_k} \\
& & & \Downarrow & & &
\end{array}
$$

$$
h_1 \quad P_{r_1,s_1} \quad h_2 \, P_{r_2,s_2} \quad ... \quad h_j \quad P_{r_j,s_j} \quad h_{j+1}
$$

Each r_i (which tells us which premise line the variable comes from) is required to be a number between 1 and k (where k is the total number of premises), and each s_i (which tells us which variable we are talking about, counting from the left) has to be between 0 and n_{r_i} (where n_{r_i} is the total number of variables in the relevant line). This guarantees that the last line will call only P_i variables that are present somewhere in the earlier lines. Thus everything in the conclusion must be either an explicitly specified string or something drawn from the material covered by the P variables of the premises. This means that the conclusion can say 'put the content of variable x from premise number y into the conclusion at this point', but not 'put in some random stuff at this point'.

My only purpose in exhibiting the above tableau is to make the point that it represents a schema fully general enough to express arbitrary string edits. It is more than general enough to state anything from simple phrase structure or categorial rules to Chomsky's most elaborate generalized transformations.

Thus a chapter of the history of mathematical logic turns out to form a crucial part of the prehistory of generative linguistics. For Post's specific design of a formalist proof system with axioms as inputs was to emerge later in Noam Chomsky's work under the new name: **generative grammar**.

Chomsky hit upon the idea of rewriting systems as a mathematical technique for giving syntactic descriptions of human languages some thirty years after Post developed his production systems. Late in 1951, in the revised version of his MA thesis, Chomsky (1951: 3) used the verb 'generate' for the relation between a grammar and a string of symbols — for the first time in linguistics, as far as I know. By 1954 both Zellig Harris (1909–1992) and Charles Hockett (1916–2000) had used 'generate' in the same way (see Harris 1954: 260 and Hockett 1954: 390), but they were probably not influenced by Chomsky's usage, for his December 1951 revision of the MA thesis was hardly circulated at all. By 1955–1956 Chomsky had completed a first draft of *The Logical Structure of Linguistic Theory*, which proposed a theory of 'generative grammars' in detail (though of course it was very little read, and did not appear in print until 1975.

From 1959 on, Chomsky has occasionally mentioned Post's name as an earlier source for the prior use of 'generate' in the mathematical literature. Chomsky (1959: 137n) notes that he is "following a familiar technical use of the term 'generate'," and cites Post (1944). However, that paper says almost nothing about how production systems work. The locus classicus for production systems is Post (1943). a paper which Chomsky has never cited, probably because (as conjectured by Urquhart 2009: 471) he learned about Post's work mainly or entirely from secondary sources like the mathematical logic text by Paul Rosenbloom (1950). In any case, in the six decades that followed, generative grammar became the overwhelmingly dominant framework for the study of syntax, and it stems directly from the formalist program of the first two decades of the 20th century as adumbrated by C. I. Lewis and turned into mainstream mathematics by Emil Post.

3 Rules

With generative grammars firmly established as mainstream in linguistics, both linguists and philosophers commonly speak as if generative grammars of the sort that Chomsky advocates contain something like rules of grammar in the traditional sense. Chomsky even had an early paper called 'On the notion "rule of grammar",' which would tempt anyone to think that. I am not aware of anyone who has pointed out that it is simply not true. Linguists have completely overlooked a key fact about generative grammars: that they do not consist of rules in any sense that would be recognized by traditional grammarians or non-linguists like philosophers.

The ordinary intuitive understanding of a rule is something that you can follow, i.e. comply with, or break, i.e. violate. It defines a regular pattern or practice, a way of going on in the same way. But the component parts of a generative grammar do not do this. Chomsky (1986: 243) recognizes this when he says,"we would not say, as scientists, that a person follows the rule of phrase structure" formulated as 'VP → V NP Clause'. A rule of this sort (a context-free phrase structure rule, to be exact) is generally thought to be interpretable as saying 'a verb phrase may consist of a verb followed by a noun phrase followed by a complement clause', but in fact it means nothing of the kind. The presence of such a "rule" in a generative grammar neither says nor implies that a VP always contains a V followed by a NP and a Clause in that order. It does not even say that this is possible. It does not entail that a VP always contains a V, or even that it may contain a V. These things may be true in a grammar with such a "rule", or they may not. Everything depends on what the rest of the grammar says.

There could be a transformational "rule" that always shifts V to the end of VP (as with

the "universal base" analyses of the early 1970s that derived even SOV languages from VSO underlying structures), or the V could be shifted out of the VP altogether (as in much more recent transformational analyses). In either case there would never be a VP containing a V with an NP following it. Nothing is fixed by any individual statement in the grammar. Only the entire grammar, taken holistically, does anything at all; and what it does is to provide an instantaneous description of the entire set of well-formed sentences. No part of the grammar expresses any generalization about the shape of expressions in the language.

Through all of the last 60 years of linguistics, and especially the discussion of linguistics among philosophers, there has been talk of Chomsky-style generative grammars containing 'rules' that is completely counter to the way generative grammars actually work. If we take a rule to be a statement that expresses some generalization about the form of linguistic expressions, then no proper subpart of a generative grammar, of any scope or size, is a rule or contains a rule.[2]

What I have said about phrase structure rules holds also for transformations. A transformation saying "move a *wh*-marked phrase to the beginning of the clause" does **not** say that the language has clause-initial *wh*-words. The language might or might not exhibit them: there could be another transformation that moves them back again (the device known as "reconstruction" in post-1980 transformational grammar does exactly that when mapping s-structures to logical form), or a transformation that moves them to the very end of the sentence, or a rule that simply expunges them completely (which is what happens in bare relatives like *the one I want*). Everything depends on the rest of the grammar and how it all interacts.

It is very important, therefore, not to imagine when we talk about people following particular rules, or languages having particular rules, that 'rule' does not refer to anything found in a generative grammar. The non-technical and informal notion of a rule is valuable, even essential; but it simply cannot be equated with any pieces or elements of generative grammars.

Normativity

What I have just said entails that we are in a certain amount of difficulty when we come to consider the issue of whether rules of grammar are (or are not) **normative**. A normative statement is one that deals not with how things are but how they ought to be, or how it is appropriate for them to be given some set of values. Nothing in a generative grammar has that property.

Some philosophers doubt that normativity can arise in a physical universe. How could any physical distribution of elementary particles constitute a situation in which some things (drawn from an indefinitely large range) are objectively 'good', or 'beautiful', or 'right'? Such reflections lead to moral antirealist views under which ethical statements like 'That's morally wrong' or 'You should apologise' are regarded as having more in common with grunts or cries of pain than statements like 'That's gold'.

---

[2]We can see a partial exception in the case of Chomsky (1981), *Lectures on Government and Binding* (*LGB*), but that is precisely because it is not fully generative. In *LGB* modules of grammar of a completely different sort are introduced. When the "binding theory" says that anaphors (like reflexive pronouns) must be bound in their governing category, it actually is talking about something that has to hold within the structure of any expression. It says that an NP node with a reflexive pronoun as its lexical realization always has a coindexed node that c-commands it in the tree (roughly, is closer to the root, and dominated by a node that also dominates the reflexive pronoun). This use in in *LGB* of what linguists often call 'constraints' is a departure that Chomsky made from his earlier theoretical work. It disappeared again after 1990 with the appearance of his 'minimalist program'.

My own views in metaethics incline toward moral realism. But at least one philosophically inclined linguist assumes we have to accept antirealist error theories of ethics. Replying to an article in which I argued that claims about grammaticality are normative ((Pullum 2007)), Geoffrey Sampson responded (2007: 112):

> I was at least assuming that grammatical description consists of statements that are correct or incorrect: but correctness is not a concept applicable to the domains of ethics or aesthetics. (As it is often put in the case of ethics, 'you cannot derive an ought from an is'.)

Sampson is assuming that claims like 'Torturing children is wrong' or 'Bach's music is beautiful' are not even truth-apt, and he thinks that my passing mention of ethics and aesthetics has committed me to the view that claims about grammaticality aren't either. This is of course nothing like what I believe. But it is instructive to read Sampson's views, and the later re-statement of them in Sampson and Babarczy (2014: 96–99), because the extreme divergence of views makes clear how difficult the philosophical clarification of descriptive linguistics is going to be.

Let me start by attempting to be clear about what I think normativity is. Normativity is generally taken primarily as a property of statements, and then derivatively as a property of domains or subject matters in which normative statements are the appropriate mode of discourse (see Millar 2004: 93–96 for a careful discussion of how the two are related).

The claims of geology are not normative; the system of table manners is. Number theory is not a normative discipline; ethics is. Aeronautical engineering is not normative; aesthetics is. (Logic is a rather interesting case, since on the one hand we want to say that it is a plain and undeniable fact that $P \rightarrow Q$ is logically equivalent to $Q \vee \neg P$: every state of the world where if you forget your debit card PIN you have to pay cash is a state of the world in which either you have to pay cash or you don't forget your PIN; but on the other hand it also seems right and proper to reason logically rather than illogically. Regrettably, exploring the case of logic is unfortunately beyond the scope of this paper.)

Millar (2004: 92–99) points out that all the classic cases of normativity involve normative statements providing reasons for doing, feeling, believing, desiring, or intending something. I believe grammatical normativity falls together with the classic cases. 'It is not good table manners to lick your knife' offers a reason for not licking your knife; 'Torturing children is wrong' implies a reason for not torturing children; 'Bach's music is beautiful' suggests a reason for planning to attend a Bach concert; 'Attributive adjectives expressing colour always follow the noun in French' provides a reason for positioning colour adjectives after the noun when seeking to be regarded as using normal French.

4   Prescriptivism

Touching on a rule of grammar that tells you where to put your adjectives brings us inevitably to a consideration of prescriptive grammars. Some discussion cannot be avoided, though in fact I will not have much truck with them here.

I make a terminological distinction that is not standard: although many have referred to old-fashioned 'don't-do-this' grammar and usage books as 'normative grammar', and many

have said that normative statements are prescriptive rather than descriptive, I am going to use 'normative' and 'prescriptive' quite differently.

With respect to the grammatical rules for a human language (especially one with a high-prestige standard variety), there is a crucial distinction between two stances or attitudes:

— **descriptive** grammar involves the identification and statement of the rules or constraints that define the linguistic system (rules and constraints that I am going to argue are normative);

— **prescriptive** grammar involves the issuing of injunctions or opinions or exhortations about what system ought to be used, or judging how well or poorly some use of language complies with a given system.

Prescriptivist grammarians certainly see language as a normative domain, but not in the way I am interested in. What primarily marks out prescriptivists is that they see their role as advising or instructing or cajoling other language users to alter their linguistic behaviour. They want to change the way you speak and right, to lead you out of error and toward the correct path.

We can set aside here the fact that prescriptivists often have the rules wrongly conceived or wrongly formulated. They often doggedly maintain the validity of rules that do not match what they profess to regard as excellent usage, e.g. the usage of people they explicitly profess to admire (Orwell, Strunk, White, whoever). Often it can be shown that they defend a rule which they unknowingly and constantly violate in their own writing, which one might have thought was a knock-down drag-out argument that the rule cannot be right, at least for their own English. They never accept such arguments, preferring to insist, irrationally, that even their own usage is to be condemned if it does not comply with the fictive rule. And they invariably ignore grammatical differences between dialects, treating non-standard English *He don't never come here no more* as simply incorrect standard English, as if it were a poorly executed attempt at saying *He doesn't ever come here any more*, when in truth languages or dialects that have negative concord working-class and low-prestige dialects of English around the world fall together with standard Italian, standard Polish, and other languages in which repeated morphological expressions of negation reinforce each other rather than cancelling out.

But all of this is basically a side issue, because even if the prescriptivists had all the rules exactly right, their enterprise would still be quite distinct from that of descriptive linguistics They are in the critical and advisory business of evaluating language use as good or bad. I'm not.

John Searle draws a relevant distinction (in his book *Speech Acts*, 1969) between **constitutive** and **regulative** rules. Constitutive rules define or set up the activities to which they apply; regulative rules are established to govern an activity that can proceed independently of them and in defiance of them. Knocking other marathon runners down in order to get ahead of them is still clearly running in a marathon, because the rule that you shouldn't use physical violence against other runners in a marathon is regulative. But moving a knight six squares directly toward the other end of the board is not playing chess; the rule that a knight moves to a second-nearest square of opposite color is constitutive.

Prescriptivists take grammatical rules to be regulative. Criticizing other people's linguistic behaviour and attempting to get them to modify it is the goal. That has nothing to do with my topic here, so I want to simply set the prescriptive stance aside.

The question I am concerned with is whether the **descriptive** view of grammar also involves takes a normative perspective of what the subject matter is.

What suggests normativity in the subject matter of grammar, more specifically syntax, is the fact that there is (or at least, linguists assume there is) a distinction between well-formed and ill-formed expressions, and it holds over an indefinitely large range, certainly far too large for it to be a matter of list membership.

What a grammar has to do is not to summarize some finite set of observations or facts, but to use all available evidence to discover a definition, over an indefinitely large class of candidate objects (potential expressions), of the difference between those that are good or properly structured in the language under study and those that are bad or improperly structured.

This does not mean that linguistics fails to be empirical (contrary to Sampson's assumption). Its task is to find out what the right constraints are, and that is not an a priori matter. It can only be done empirically, ultimately by reference to the usual behaviours and reactions of the native users of the language when distractions and irrelevant extraneous factors do not intrude. This is true despite the fact that both intuitions and corpus attestations are fallible sources of evidence. The epistemology is therefore subtle. I have suggested elsewhere (see Pullum 2017) that it should be seen as based on the method of reflective equilibrium.

The way generative linguists usually view it, the grammar has to cover all the expressions of the language, and only those expressions, and it must do it in a way that tells us the status of novel expressions — expressions we have never encountered before. That means making a description that is fully explicit about how the expressions of the language — all of them, however many there may be — are structured. And that calls for some kind of formalization of both the representation mode and the and grammatical rule system.

5 Formalization

What I mean when I refer to formalization in syntax is simply **the use of mathematical and logical tools to make theoretical claims more explicit**. Talk about formalization is therefore not essentially connected to the 'formalist' progam. It has nothing to do with de-emphasizing the semantic, pragmatic, rhetorical, or aesthetic aspects of human languages, or with assigning more importance to form than to function, or with Carnap's project of eliminating meaning from the language of science, or with Hilbert's (doomed) project of reducing all of mathematics to questions of logical truth in some decidable formal logic.

The tools that formally-inclined linguists have borrowed from logic for use in framing syntactic theories over the last few decades have included rewriting systems, automata, graphs (most importantly trees), and model theory. More recently, 21st-century linguistics has been increasingly employing tools from statistics and probability theory.

But there is truth in the familiar remark about how to a three-year-old with a hammer everything looks like a nail. Syntacticians have become so completely engrossed in working with generative grammars that they see everything in terms of derivations, and cannot conceive of what life would be like in any other terms.

They have paid very little attention to the fact that a generative grammar of L says absolutely nothing about the structural properties of any non-sentence of L. They have ignored the fact that the sharp boundaries of any set defined by a generative grammar fly in the face of the

widely accepted intuitive view of ill-formedness as gradient — the fact that one ungrammatical sentence can be more ungrammatical than another.

They have also paid little or no attention to the fact that a generative grammar makes syntactic properties depend crucially on the contents of a finite lexicon: a derivation that does not terminate in a string of items belonging to the relevant lexicon is not a derivation at all, so there is no way for a generative grammar to represent an example like Carnap's *Pirots karulize elatically* as grammatical unless the lexicon contains a noun *pirot*, a verb *karulize*, and an adverb *elatically* — which for standard English it does not.

These are not problems for formalized syntax in general. For one thing, when we are talking about the invented languages of logic and computer programming, the worries I just expressed about generative grammars turn into virtues. For proving theorems about logical systems — completeness, consistency, compactness — it is absolutely crucial that the formulas of the logic should be sharply defined to form a specific set with a known cardinality. And for proving correctness of a computer program, the same is true. There can be no gradient levels of ill-formedness, or potential tolerance of minor deviance, or uncertainty about the finite list of allowable symbols, when we are talking about logics or programming languages. This is the grain of truth in Michael Tomasello's capsule summary of 'Chomskian generative grammar' (Tomasello 2003: 5), which he says is 'a "formal" theory, meaning that it is based on the supposition that natural languages are like formal languages.' In a way, he is correct, though I suspect not in the way he intended.[3]

The sense in which generative grammars do treat natural languages like formal languages has to do with the origins of their formal machinery, as already outlined. It was developed for purposes within mathematical logic, and it is perfectly suited to the description of the invented languages for logic, metalogic, and computer programming. But it is important that there ways of making grammars mathematically explicit that are quite distinct than the generative one. Chomsky has sometimes confusingly denied this point, claiming (as in e.g. Chomsky 1966: 12) that for him the term 'generative grammar' means nothing more than 'explicit account of sound/meaning correspondences', but this does not square at all with his actual usage (Ney 1993 argues this point at length). Pullum and Scholz (2001) use the term 'generative-enumerative syntax' to stress that the referent is syntax formalized in terms of nondeterministic random enumerators; their paper discusses certain types of explicit grammar that are not generative in this sense.

The non-generative mode of formalizing grammars that Pullum and Scholz discuss uses model theory rather than rewriting systems to formalize syntactic description. Grammatical rules are taken to be constraints on the structure sentences, in a straightforward and informally comprehensible sense: The constraints in a model-theoretic grammar for English would say things that for convenience we can readily paraphrase in English. A few examples:

— A preposition phrase has a preposition as head.

— A lexical head is the initial subconstituent of its parent.

---

[3]The rest of his summary is inaccurate and confused. He says generative grammar uses 'a unified set of abstract algebraic rules' (they are actually of diverse types, not at all unified); and they 'are both meaningless themselves and insensitive to the meanings of the elements they algorithmically combine' (but algebraic operations always need interpretations, as McCawley (1968) carefully shows, and grammar rules can be written to build semantic representations simultaneously with syntactic ones, as in Montague (1973) or Gazdar et al. (1985). Finally he says the rules come with 'a lexicon containing meaningful linguistic elements that serve as variables in the rules' (but I see no sense in which lexical items serve as variables).

— A pronoun subject of a finite clause takes its nominative case form.

Such constraints can be stated more precisely as formulas of a logic; a grammar can be defined as a finite set of such formulas; structures of sentences can be taken to be the models for the interpretation of that logic; and grammaticality can be reconstructed as satisfaction of the constraints in the grammar, in the model-theoretic sense.

For additional concreteness, let's look very briefly at the syntax of preposition phrases (PPs) in English. A typical old-fashioned generative grammar would include a phrase structure rule like this:

PP → P NP

Under its standard interpretation this licenses derivational steps in which the symbol 'PP' is replaced by the sequence 'P NP', and derivatively licences the building of (part of) a tree diagram that looks like this:

```
   PP
  /  \
 P   NP
```

You might easily think that the rule entails that prepositions always have NP complements. It does not. There could be (and in fact for English there will need to be) other rules in the grammar saying things like this:

PP → P PP

PP → P Clause

So some PPs will not have NP right branches. And you might easily think that the rule at least says that those prepositions that do take NP complements precede their NP complements. But it doesn't entail that either. There could be another rule in the grammar saying this (where ε is a symbol representing the null string):

PP → ε

In that case there might be no prepositions appearing in the language at all. Or there could be a transformational rule like this:

```
X  -  P  -  NP  -  Y
1  -  2  -  3  -  4  ⇒  1  -  3  +  2  -  4
```

In that case prepositions would always be suffixed to their NP complements — unless some other rule in the grammar tampered with things further. This is what I am referring to when I say that the grammar provides its definition holistically: in the same way that we are told that in the Brexit negotiations nothing is agreed until everything is agreed, with a generative grammar you don't know anything about what any part of the grammar says about any part of a sentence until you know what the entire grammar yields.

You might certainly have thought that the rule 'PP → P NP' does says that a PP always contains a P. But it does not do that either. It doesn't guarantee anything about the interior of PPs. There could be another rule in the grammar saying 'PP → A B C'.

Under the view I favour, the conditions on PPs would be stated directly as constraints — and for concreteness let's take them to be constraints on the structure of trees.

Assume a predicate logic in which we quantify over nodes with variables $x, y, z \ldots$ and have a vocabulary of monadic predicates for category labels and binary relations for grammatical functions. We write 'B(x)' to mean 'node x is labelled with category label B' and 'F(x, y)' to mean 'the F of node x is node y' (i.e., 'node y bears the grammatical function F to its parent node x'). Then the constraint saying that PPs have P heads would be expressed precisely in this way:

$$\forall x[\text{PP}(x) \rightarrow \exists y[\text{Head}(x, y) \wedge P(y)]]$$

And if we write Lexical(x) to mean 'node x bears a lexical category label', 'Parent(y, x)' to mean 'x immediately dominates y', and '$x \prec y$' to mean 'node x is to the left of node y', then the second constraint can be stated like this:

$$\forall x, y[\text{Head}(x, y) \wedge \text{Lexical}(y) \rightarrow \forall z[\text{Parent}(z, x) \rightarrow x \prec z]]$$

Any set of trees characterized by a set of first-order logic formulas in this sort of way will be a regular set of trees (recognizable by a finite-state tree automaton) and its string yield will be a context-free stringset (these results are corollaries of theorems now found in textbooks of finite model theory like Ebbinghaus and Flum 1999 and Libkin 2004). By giving a finite set of first-order logic statements interpreted on tree models in this way we explicitly characterize a set of trees and thereby a context-free set of strings. We are in effect giving a formally explicit for a context-free language without using a context-free generative grammar. I argue elsewhere (Pullum 2013) that this yields significant advantages.

I point all this out merely in order to establish the point that normative principles like those stated informally in English above can be made fully precise and become, without change, a formalized grammar with precise consequences, known parsability results, etc. I am in no doubt that it is worth pursuing the goal of making the syntactic structural principles for human languages fully explicit — formalizing them using the tools of logic and mathematics.

On the other hand, I do not believe that by formalizing some version of the grammar rules for a language we are thereby defining a hypothesis about the mind — ultimately the brain structure — of an idealized native speaker of that language. Kripke (1982) believes there are profound difficulties for this basically Chomskyan view, because there seems to be no way to identify the unique set of rules that guides a given speaker's grasp of their native language: indefinitely many grammars would account for all of the utterances they have produced (or judged acceptable) in the past, and nothing identifiable about the speaker in question can be said to fix the structure that will be revealed in further utterances as the speaker goes on.

I think Scholz (1990) was right to elaborate on Kripke's worry, and argue that we do not obtain an explanation of S's linguistic capabilities simply by saying 'S has a mentally inscribed representation of the generative grammar G in his brain'. And I think Chomsky is wrong in responding to Kripke by denying that there is any normative aspect to grammar (see Chomsky 1986: Ch 4).

What I am suggesting is that we are better placed to see how there can be a normative conception of grammar that is not prey to Kripke's metphysical worries if we conceptualize grammar in model-theoretic terms, as approximate compliance with certain structural constraints on the form of sentences.

It is just possible that it might also assuage some of the worries expressed by Riemer (this volume) concerning a kind of authoritarianism that he sees in the ideas of generative grammar. Imagining that there is some generative mechanism that is the mental and ultimately neurophysiological reality underlying the capacity to use English, and teaching students about it as if it were unique, he feels, is inimical to the idea that this complex world can be viewed from many divergent perspectives. It might even militate against our students feeling that they have the intellectual freedom to explore alternatives. (I am not sure I capture Riemer's worry with this brief paraphrase; I urge the reader can turn to his chapter.)

The relevance of the model-theoretic approach to Riemer's problem is that I do not regard it as remotely plausible that there is one unique correct set of constraints defining English, or one correct way of formally stating that set. What is needed for someone to be a competent user of English is not that they have some ideal and perfect generative system implanted in their brain chemistry, but merely that they have developed a set of constraints on grammatical form that, to a good approximation, leads them to structure their utterances in ways roughly like the way other speakers of English structure theirs. There will be indefinitely many ways to do this, and indefinitely many ways to represent formally what has been done. Let a hundred flowers blossom and a hundred schools of thought contend. It is overwhelmingly plausible that among the hundreds of millions of native speakers of English, many slightly different systems of constraints contend. The reason this is not problematic is that all we need for linguistically expressed communication to be possible is substantial overlap in the consequences of the different systems. (The slight differences between them will of course be the seeds of future linguistic change.)

One further thing I should say about normativity concerns the notion of universal grammar, where the issue of normativity may not be relevant at all. Modern linguistics since Chomsky (1965) has stressed the goal of formulating a theory of **universal** grammar (UG). This theory has been taken to be not just a systematization of the facts that have been discovered to hold for all human languages, but a kind of ideal model of the human infant's capacity for learning languages. And it is important that UG could in principle be an entirely non-normative domain: it could be a description of a set of neurophysiologically instantiated devices or of a psychological organization.

I don't regard this as plausible or well supported, because of the lack of any account of mechanisms. How does UG constrain the growth of grammars? We know how the curved horns of a ram grow (faster-growing cells on one side and slower-growing on the other), but nothing at all about how UG works or develops in the brain. And more recent work tells us nothing that helps (the literature on what some have been calling 'biolinguistics' over the past decade seems to me increasingly to be parodying itself).

But we can set aside the issue of whether there might be a serious theory of the biological aspects of the human language acquisition capacity. I am concerned here simply with understanding grammars of individual languages. This will be necessary regardless of whether the human capacity to form and use them is constrained by a built-in UG. It is by no means clear clear to me that any non-normative, naturalistic, neurophysiological account of the parochial grammars of specific languages makes sense.

It is uncontroversial that the mentally inscribed grammar that Chomsky posits does not (i) describe all the utterances that have occurred in the past, or (ii) predict what utterances will occur in the future, or (iii) identify the probabilities of occurrence for future utterances, or (iv) make you do whatever it is you do when you produce an utterance or understand one. Assuming that *Fetch quickly it* is ungrammatical, for example, the correct statement of the grammatical rules of English does not imply (i) that no one has ever said it, or (ii) that no one will ever say it, or (iii) that the probability of its being uttered is low, or (iv) that you ought not to utter it.

The implication of a grammar defining something as ungrammatical is closer to being that no one **should** say it if they want to be taken as speaking English as it is usually spoken. What kind of 'should' is that? Well, it should be clear enough that we are not talking about anything closely analogous to the moral sense of 'should'. Moral philosophers standardly take morality to be universal. Certainly for a moral realist (and moral realism is the metaethical view that I would subscribe to), that torturing a child is not immoral just for certain people in certain circumstances though possibly moral in other cultures or under other circumstances; it is morally wrong for everyone. Someone who disagrees is simply mistaken. Societies and cultures can evolve, and come to see that moral views they held earlier were mistaken,

But essentially none of the rules in the grammar of a particular language can be taken to hold universally. There are some 7,000 extant languages, differing quite radically in grammar as well as lexicon. no one of them is more grammatically correct in its structures than any of the others. In matters of grammar we have to be radically relativist.

Morality also relates to behaviour in a way that has consequences for human actions. From the fact that torturing a child is immoral it follows that you ought not to torture a child. But from the fact that *Fetch quickly it* is ungrammatical in English nothing at all follows concerning what anyone ought to do or not do. Whether you ought to utter it will depend entirely on the circumstances. During a minute of silence at a funeral, you ought not to utter it, or anything else. But if you are playing a foreign character in a play and the script has your character saying *Fetch quickly it*, then at the relevant point you ought to utter it.

This issue has sometimes been discussed in the philosophy literature, in the context of the normativity of meaning. Philosophers of language are in fact maddeningly uniform in their habit of talking only about reference of words — as if all we ever do with our language were pronounce the word 'cat' and successfully achieve reference thereby to a creature of the species *Felis silvestris catus*, and as if that were the deepest and most interesting thing about language. But even there, saying that *cat* refers to a certain animal species in the Felidae family doesn't mean that you ought to describe cats by using that word. Even if you do want to refer to the creatures, you might want to use *moggy* or *pussy* or *foul razor-clawed mewing beast from hell*. And you have no obligation to talk about cats or refer to them at all if you don't choose to.

My concern here is not with elementary lexical meaning but with syntax. However, it just as clear that no syntactic rule or constraint or principle conveys any obligation or presumption about how you should act.

For one thing, the constraints of syntax mostly involve categories and restrictions and classifications of which you have no conscious knowledge whatsoever, so typically you could never know whether you had violated them or not. And almost everyone agrees that 'ought' implies 'can', of course: it is generally not taken as coherent to assert that you have a duty to do something that you are totally incapable of doing.

This point offers a clue to some understanding of what the prescriptivists are up to. They peddle rules that, while not a correct reflection of the actual syntax of the language, are fairly easy for even linguistically unsophisticated people to check: never place an adverb between infinitival *to* and a plain-form verb (*to boldly go*), never end a sentence with a preposition (*What are you looking at?*), never begin a sentence with *however*, avoid the passive, etc. Such rules are actually hopeless as a guide to how to actually use the language like a normal person, but they have become cultural markers of attention to grammar. Most people aren't aware of how inaccurate they are or how much they are ignored by truly accomplished writers, and it is moderately easy to identify violations.

Prescriptivists fail in some of their identifications: they mistake particles for prepositions, and mistake existentials or predicative adjective constructions for passives, and so on; but they THINK they are correctly applying genuine rules, and they see a kind of quasi-moral force to the rules: they see the people who do not respect the rules as falling short of the standards of behaviour that society ought to maintain and enforce. In other words, they see the rules of grammar as regulative.

Domains involving obligations on agents to act (or not act) in certain ways are not the only domain for normativity. Neither constitutive nor regulative rules of a game have any connection to anything about obligations. You are not obliged to play the game at all, for one thing, so the rule defining a knight's move certainly doesn't say you should or should not move any particular knight.

Moreover, making an illegal move is not something you 'shouldn't do' in absolute terms; whether you should do it depends on the circumstances. You might be playing against a sadistic jailer, with the life of a fellow prisoner forfeited if you lose, in which case you SHOULD make the illegal move if you can get away with it and it will ensure a win.

What you should **do**, if 'should' has anything like the sense it has in connection with morality or similar kinds of obligation, has nothing whatever to do with the normative force of linguistic rules.

I should also note that the rules of syntax do not have normative force in any **instrumental** way: it is not that they should be obeyed because bad things will happen if they are not obeyed, or because following them will enable one to get things that one needs. There is a sometimes hint of this in naive talk about language and how we learn it: people talk about needing to make oneself understood in order to interact satisfactorily with other people. But in fact there is very little pressure to get things right grammatically. There are people who live most of their lives speaking mostly in a language that is not their native language, and speaking it very badly. They still find ways to get what they want.

And perhaps the starkest and most obvious refutation of the instrumental view comes from looking at the actual facts of human infants' experience: while they are incapable of speech they are constantly looked after, and all their needs are met, but once they are 4 or 5 and can speak and understand they start being expected to do things that other people want them to do. The idea that we speak the way we do because we have to on instrumental grounds is nonsense.

6 Understanding

The normative rules of games and linguistic systems do not define anything as unethical, or contemptible, or inadvisable, or evil. But I think what they do can be elucidated in terms of the very interesting work of Alan Millar in *Understanding People* (2004) cited above.

Millar stresses the notion of a **commitment** to follow rules: he separates the **following** of rules from the **commitment** to follow them.

It is important that there can be tacitly acknowledged rules that no one has set down in detail. Management of phone calls is an example. Nowhere is it set down that when a call is connected, the recipient of the call is supposed to speak first, or that normally the maker of the call is supposed to instigate the ending of the call, or that the recipient speaks last (echoing the goodbye of the maker of the call). The fact that a certain set of rules is tacit does not preclude the existence of a practice based on them.

This is the potential answer to the problem with Crispin Wright's observation about effortless first-person authority: we don't appear to have that for syntax, in the general case. We can have tacit command of a rule that one cannot state, and the fact that we cannot state it does not mean we cannot recognize departures from what it requires.

What the tacitly acknowledged rules of grammar do is to define certain ways of structuring expressions and associating them with meanings. Millar put it this way in the handout for a talk at the University of Edinburgh in 2010:

> The commitments incurred by participation in a practice have a closely analogous structure to those incurred by beliefs and intentions. Participating in a practice, G, incurs a commitment to following the rules of G. To a first approximation that amounts to it being the case that one ought to avoid continuing to participate in G while not following the rules of G. ... [T]here are two ways to discharge the commitment—by withdrawing from the practice and by complying with the rules.

Rules of grammar that define expressions as being structured in certain ways do not entail that you ought to structure your expressions in those ways.

You do incur a commitment to structure your expressions in the English way when you decide to speak English. But incurring a commitment to do something does not entail that you ought to do it. Sometimes the way to deal with a commitment that you incur by engaging in a certain practice is to cease engaging in the practice.

Millar uses the example of offensive ethnic terms to illustrate the point that the existence of a practice says nothing about what you ought to do: *Chink* in English is an offensive epithet meaning "Chinese person". But the existence of a practice of saying *Chink* to mean 'Chinese person' does not imply that anyone ought to use the term: most would say the opposite is true.

It is not by any means as easy to illustrate the same sort of thing in syntax, but I think it is possible. Consider the fact that is usual for the direct object to follow the verb in what CGEL calls canonical clauses:

The old tree shades the house.

*The old tree the house shades.

That rule codifies a certain practice, applying the generalization to an indefinitely large range of expressions, not just recording properties of previous utterances. It has normative force in that it defines it as syntactically incorrect to put the object before the verb. If you do not put the object before the verb, other things being equal, you are not respecting the constraints of English syntax. But it does not follow that anyone ought to place any direct object before any verb. A person might have reason not to respect the constraints of English syntax. For example, someone might be imitating the poetry of the 18th or 19th centuries, where the long-extinct order with object before verb was often employed.

One might well ask what point there could be to having rules that do not have to be obeyed. Isn't this like having rules of the road that nobody has to follow? The rules of the road at least have instrumental motivation relating to safety, but in the case of grammar we do not even have that. Millar in effect answers that question when he says:

> On the view I am promoting it is our participation in practices and our ability to recognize what these require in particular situations that enables us both to make reasonable predictions about, and make sense of, what people think and do. The basic idea is simple. Because we know how people are supposed to act we can often make sense of how they do act and how they are likely to act.

In talking about making sense Millar is not necessarily to be interpreted as referring to the understanding of meaning, either semantic or pragmatic. The point has nothing inherently to do with the correct apprehension of literal meaning or perceiving the utterer's intended meaning, though it may apply to those accomplishments, and it may not even imply those abilities. I am breaking with the usual practice among philosophers of language in wanting to talk about syntax.

The understanding I am alluding to is simply a matter of making sense of what is going on linguistically: what sort of system the interlocutor is using, what to expect, what to infer about the intent of a speaker or the likely form of further utterances.

Understanding of meaning is neither necessary nor sufficient for the syntactic ability I am talking about. It is not necessary because it is often possible to grasp the structure of an uttered sentence without having the vaguest clue as to its meaning. Faced with something like *I doubt whether that is not necessarily not untrue*, most people are aware that they have heard something grammatical but they cannot work out the truth conditions of a quadruple negation. And you can immediately perceive the grammaticality of *Appearances are not deceptive, it only seems as if they are*; working out from its meaning whether it is sensible or contradictory is much harder. But it is also not sufficient: being able to extract the meaning does not imply grasping the syntactic structure. We often correctly identify intended meanings despite massive non-compliance with syntax: if a worried stranger says '*Me need you help! Me house go fire!*' you would know what was meant.

However, it remains true that if you follow the practice of ordering words and structuring phrases and clauses in the English manner, and I understand you to be following that practice, I can make sense of certain aspects of your intentions that otherwise would be inscrutable to me, and I can make much better guesses at why your linguistic behaviour is the way it is, and what to expect in future utterances. I can assume that if you begin an utterance with an auxiliary verb and continue with a nominative pronoun (*Can we...*; *Is he...*; *Will they...*) you are beginning a closed interrogative, and thus that you will probably continue with a verb phrase, and are probably about to ask me a question with a yes/no answer, and are likely to have made the

assumption that I know the answer to the question, and so on.

If you depart from the usual practices, that doesn't necessarily mean that you shouldn't have, or that you have made a mistake. If I detect that you have departed from the usual grammatical practices, I can simply regard you as having ceased to operate in accord with the usual rules, and I can reason about why that might be and what I should assume from now on. A normal context is one where you appear to be respecting the same constraints on sentence structure as I would, and using words I have encountered before, so I assume you are using my language, and I turn out to be right. But there are also anomalous contexts, such as one in which a native speaker who is required to draft a forced confession by the agents of a foreign dictatorship puts subtle syntactic mistakes into the text to signal that he is under duress, intending people back home to spot that he is not complying with the usual constraints.

There are any number of reasons why you might depart from normal syntax: — you might be half asleep, drunk, delirious, brain-damaged, interrupted, distracted, foreign, playful, joking, impersonating someone. . . But if I can see **why** you have departed from the usual grammatical practices, I may be able to understand what is going on with you. Otherwise, you become mysterious. But it is of course your right to be mysterious if you want to: the fact that grammatical constraints have normative force doesn't impose on you, even as a default, any obligation to obey.

7   Conclusion

What I have tried to do in this rather wide-ranging survey is to make the following points. I have stressed that the formalist movement in logic had an important rationale: representing logical proof in a totally syntactic way so that the completeness and consistency of a logical language (the correspondence between its resources for proof and its efficacy in preserving truth under inference) could be mathematically verified. It was in pursuit of this goal that Emil Post invented the systems we now call generative grammars the systems that were repurposed three decades later for use in linguistics.

Though talk of 'rules of grammar' persisted in linguistics, even in Chomsky's work, it did not really make sense within a generative perspective (as Chomsky himself noticed): rewriting operations are **not** rules in the sense that you could follow them, behave in accordance with them, be guided by them, or violate them. Yet we do need a conception of rules with normative force in the ordinary sense of rules that can be respected or violated, for the point of linguistic grammars is not to compactly represent a corpus or to characterize a mental organ, but to define well-formedness over an indefinitely large class of sentences.

The claim that rules of grammar have normative force should not be confused with the bid to change people's linguistic habits and practices represented by the prescriptive grammar tradition. And it should also not be taken to be at odds with the empirical character of linguistics (there is a fact of the matter about whether English is a prepositional or a postpositional language), or to be in conflict with the goal of making the predictions of grammars fully clear and explicit, with help from tools from logic and mathematics.

I believe that generative grammars (the systems that Post invented) were the wrong tool to pick, but there are alternatives. Repurposing model theory for syntactic description, for example, is a better idea. Under that view grammars are finite sets of constraints on structure, and well-formedness is construed as model-theoretic satisfaction. There is a small but growing

minority of syntacticians who are developing theories of grammar along these lines (see e.g. Rogers 1998, Pullum and Scholz 2001, Pullum 2013).

What remains to be accounted for is how formalized constraints on sentence structure can have normative force without in any sense implying anything about what anyone should do, or implying any judgment on a person that they have done something wrong (or right). And I believe the answer to that can be found in the work of Millar (2004), who relates the normativity of rules in certain kinds of systems not in what you ought to do but in what can make your words or actions more intelligible to others.

References

Chomsky, N.: 1951, *Morphophonemics of Modern Hebrew*. Typescript of a radical revision of Chomsky's MA thesis, dated December 1951; retyped and published by Garland, New York, in 1979.

Chomsky, N.: 1955–1956, *The Logical Structure of Linguistic Theory*, Manuscript, mimeographed 1955; revised 1956 and distributed on microfilm by MIT Library, Cambridge, MA; ultimately published as Chomsky 1975.

Chomsky, N.: 1959, On certain formal properties of grammars, *Information and Control* **2**(2), 137–167. Reprinted in *Readings in Mathematical Psychology*, Volume II, ed. by R. Duncan Luce, Robert R. Bush, and Eugene Galanter, 125–155, New York: John Wiley & Sons, 1965 (citation to the original on p. 125 of this reprinting is incorrect).

Chomsky, N.: 1965, *Aspects of the Theory of Syntax*, MIT Press, Cambridge, MA.

Chomsky, N.: 1966, *Topics in the Theory of Generative Grammar*, Mouton, The Hague.

Chomsky, N.: 1975, *The Logical Structure of Linguistic Theory*, Plenum, New York. Published version of Chomsky (1955).

Chomsky, N.: 1981, *Lectures on Government and Binding*, Foris, Dordrecht.

Chomsky, N.: 1986, *Knowledge of Language: Its Origins, Nature, and Use*, Praeger, New York.

Davis, M.: 1982, Why Gödel didn't have Church's thesis, *Information and Control* **54**(1), 3–24.

Ebbinghaus, H.-D. and Flum, J.: 1999, *Finite Model Theory*, second edn, Springer, Berlin and New York.

Gazdar, G., Klein, E., Pullum, G. K. and Sag, I. A.: 1985, *Generalized Phrase Structure Grammar*, Basil Blackwell, Oxford.

Harris, Z. S.: 1954, Transfer grammar, *International Journal of American Linguistics* **20**, 259–70.

Hockett, C. F.: 1954, Two models of grammatical description, *Word* **10**, 210–231. Page references are to the reprinting in Joos (ed.) 1966, 386-399.

Kripke, S.: 1982, *Wittgenstein on Rules and Private Language: An Elementary Exposition*, Harvard University Press, Cambridge, MA.

Lewis, C. I.: 1918, *A Survey of Symbolic Logic*, University of California Press, Berkeley, CA.

Libkin, L.: 2004, *Elements of Finite Model Theory*, Texts in Theoretical Computer Science, Springer.

McCawley, J. D.: 1968, Concerning the base component of a transformational grammar, *Foundations of Language* **4**, 243–269. Reprinted in James D. McCawley, *Grammar and Meaning*, 35–58 (New York: Academic Press; Tokyo: Taishukan, 1973).

Millar, A.: 2004, *Understanding People: Normativity and Rationalizing Explanation*, Oxford University Press, Oxford.

Montague, R.: 1973, The proper treatment of quantification in ordinary English, *in* J. Hintikka, J. Moravcsik and P. Suppes (eds), *Approaches to Natural Language*, D. Reidel, Dordrecht, pp. 221–242. Reprinted in Montague (1974), 247–270.

Montague, R.: 1974, *Formal Philosophy*, Yale University Press, New Haven, CT. Edited by Richmond Thomason.

Ney, J.: 1993, On generativity: The history of a notion that never was, *Historiographia Linguistica* **20**(2/3), 441–454.

Post, E. L.: 1921, Introduction to a general theory of elementary propositions, *American Journal of Mathematics* **43**, 163–185. Reprinted in van Heijenoort (1967: 264–283) and reproduced in Davis (1994b: 21–43).

Post, E. L.: 1943, Formal reductions of the general combinatory decision problem, *American Journal of Mathematics* **65**(2), 197–215. Reproduced in Davis (1994b: 442-460).

Post, E. L.: 1944, Recursively enumerable sets of positive integers and their decision problems, *Bulletin of the American Mathematical Society* **50**(5), 284–316. Reproduced in Davis (1965: 305-337) and in Davis (1994b: 461-494).

Pullum, G. K.: 2007, Ungrammaticality, rarity, and corpus use, *Corpus Linguistics and Linguistic Theory* **3**, 33–47.

Pullum, G. K.: 2013, The central question in comparative syntactic metatheory, *Mind and Language* **28**(4), 492–521.

Pullum, G. K.: 2017, Theory, data, and the epistemology of syntax, *in* M. Konopka and A. Wöllstein (eds), *Grammatische Variation: Empirische Zugänge und theorische Modellierung*, Institut für Deutsche Sprache, Jahrbuch 2016, Walter de Gruyter, Berlin, pp. 283–298.

Pullum, G. K. and Scholz, B. C.: 2001, On the distinction between model-theoretic and generative-enumerative syntactic frameworks, *in* P. de Groote, G. Morrill and C. Retoré (eds), *Logical Aspects of Computational Linguistics: 4th International Conference*, number 2099 in *Lecture Notes in Artificial Intelligence*, Springer, Berlin and New York, pp. 17–43.

Rogers, J.: 1998, *A Descriptive Approach to Language-Theoretic Complexity*, CSLI Publications, Stanford, CA.

Rosenbloom, P.: 1950, *The Elements of Mathematical Logic*, Dover, New York.

Sampson, G.: 2007, Reply, *Corpus Linguistics and Linguistic Theory* **3**(1), 111–129.

Sampson, G. and Babarczy, A.: 2014, *Grammar Without Grammaticality: Growth and Limits of Grammatical Precision*, number 254 in *Trends in Linguistics: Studies and Monographs*, Walter de Gruyter, Berlin.

Scholz, B.: 1990, *Kripke's Wittgensteinian Paradox*, PhD thesis, The Ohio State University, Columbus, OH. Published by University Microfilms, Ann Arbor, Michigan.

Searle, J.: 1969, *Speech Acts: An Essay in the Philosophy of Language*, Cambridge University Press, Cambridge.

Tomasello, M.: 2003, *Constructing a Language*, Harvard University Press, Cambridge, MA.

Urquhart, A.: 2009, Emil Post, *in* D. Gabbay and J. Woods (eds), *The Handbook of the History of Logic, Volume 5*, Elsevier, Amsterdam, p. 617–666.

Whitehead, A. N. and Russell, B.: 1910–1913, *Principia Mathematica*, Cambridge University Press, Cambridge.